





David C. Wyld  
Dhinaharan Nagamalai (Eds)

# Computer Science & Information Technology

The Fourth International Conference on Database and Data Mining  
(DBDM 2016)  
Dubai, UAE, April 23~24, 2016



**AIRCC Publishing Corporation**

## **Volume Editors**

David C. Wyld,  
Southeastern Louisiana University, USA  
E-mail: David.Wyld@selu.edu

Dhinaharan Nagamalai,  
Wireilla Net Solutions, Australia  
E-mail: dhinthia@yahoo.com

ISSN: 2231 - 5403  
ISBN: 978-1-921987-50-2  
DOI : 10.5121/csit.2016.60501 - 10.5121/csit.2016.60512

This work is subject to copyright. All rights are reserved, whether whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the International Copyright Law and permission for use must always be obtained from Academy & Industry Research Collaboration Center. Violations are liable to prosecution under the International Copyright Law.

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

## Preface

The Fourth International Conference on Database and Data Mining (DBDM 2016) was held in Dubai, UAE, during April 23~24, 2016. The Fourth International Conference on Instrumentation and Control Systems (CICS 2016), The Third International Conference on Signal Processing (CSIP 2016), The Fourth International Conference of Artificial Intelligence and Fuzzy Logic (AI&FL 2016), The Fourth International Conference of Soft Computing (SCOM 2016), The Fourth International Conference on Computational Science and Engineering (CSE 2016) and The Third International Conference on Computer Networks & Communications (CCNET-2016) were collocated with the DBDM-2016. The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The DBDM-2016, CICS-2016, CSIP-2016, AI&FL-2016, SCOM-2016, CSE-2016, CCNET-2016 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically. All these efforts undertaken by the Organizing and Technical Committees led to an exciting, rich and a high quality technical conference program, which featured high-impact presentations for all attendees to enjoy, appreciate and expand their expertise in the latest developments in computer network and communications research.

In closing, DBDM-2016, CICS-2016, CSIP-2016, AI&FL-2016, SCOM-2016, CSE-2016, CCNET-2016 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the DBDM-2016, CICS-2016, CSIP-2016, AI&FL-2016, SCOM-2016, CSE-2016, CCNET-2016.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

David C. Wyld  
Dhinaharan Nagamalai

## Organization

### General Chair

Jan Zizka  
Dhinaharan Nagamalai

Mendel University in Brno, Czech Republic  
Wireilla Net Solutions, Australia

### Program Committee Members

Abd El-Aziz Ahmed  
Abdelmadjid Allali  
Abdolreza Hatamlou  
Abe Zeid  
Aftab Alam  
Ahmed Hussein Aliwy  
Ahmed Taisser  
Ahmet Cinar  
Alejandro Regalado Mendez  
Aleksandar Sugaris  
Ali Abid D.Al-Zuky  
Ali Elkateeb  
Amar Faiz bin Zainal Abidin  
Anja Richert  
Ankit Chaudhary  
Ashraf A.Shahin  
Assem Abdel Hamied Moussa  
Ava Clare Marie O. Robles  
Ayad Ghany Ismaeel  
Ayad N.M.A  
Ayad Salhie  
Barbaros Preveze  
Chin-Chih Chang  
Chiranjib Sur  
Daniel D.Dasig  
Derya Birant  
Dongchen Li  
Edgardo Palza  
Elashiry M.A  
Farzad Kiani  
Fatih Korkmaz  
Fernando Zacarias  
Ghadeer Hassan  
Guoqing Xiao  
Hamza Zidoum  
Hoda Mokhtar Omar Mokhtar

Cairo University, Egypt  
Ben Bouali University, Algeria  
Islamic Azad University, Iran  
Northeastern University, USA  
King Khalid University, Saudi Arabia  
University of Kufa, Iraq  
Cairo University, Egypt  
Firat University, Turkey  
Universidad del Mar, Mexico  
ICT College, Serbia  
Mustansiriyah University, Iraq  
University of Michigan-Dearborn, USA  
Universiti Teknologi MARA, Malaysia  
RWTH Aachen University, Germany  
Truman State University, USA  
Cairo Univeristy, Egypt  
E commerce Manager, Egypt  
Mindanao State University, Philippines  
Erbil Polytechnic University, Iraq  
Atomic Energy authority, Egypt  
Australian College of Kuwait, Kuwait  
Cankaya University, Turkey  
Chung Hua University, Taiwan  
University of Florida, US  
Jose Rizal University, Philippines  
Dokuz Eylul University, Turkey  
Peking University, China  
University of Quebec, Canada  
Beni Suef University, Egypt  
Istanbul S.Zaim University, Turkey  
Cankiri Karatekin University, Turkey  
University of Puebla, Mexico  
Arab Academy for Science, Egypt  
Hunan University, China  
Sultan Qaboos University, Oman  
Cairo University, Egypt

Hossein Jadidoleslamy  
 Intisar Al-Mejibli  
 Jan Lindstrom  
 Jose Raniery  
 Kim Le  
 Kirtikumar Patel  
 Laya Ebrahimi  
 Lynne Grewe  
 Manal Abdullah  
 Manish Kumar Mishra  
 Maryam Rastgarpour  
 Marystella Amaldas  
 Mellal Mohamed Arezki  
 Mohammed AbouBakr Elashiri  
 Mohammed Amin  
 Mohsen Poor Arab  
 Moses Ekpenyong  
 Muhammad Sajjadur Rahim  
 Natasa Zivic  
 Natarajan Meghanathan  
 Nithya Rekha Sivakumar  
 Ognjen Kuljaca  
 Oluwatobi Olabiya  
 Omer Deperlioglu  
 Peiman Mohammadi  
 Rafah M. Almuttairi  
 Ramayah T  
 Rangiha Mohammad  
 Rhattoy A  
 Rim Haddad  
 Robert Burduk  
 Saad M. Darwish  
 Samadhiya  
 Selwyn Piramuthu  
 Seyed Ziaeddin Alborzi  
 Seyyed AmirReza Abedini  
 Shahab Shamshirband  
 Shuxiang Xu  
 Sidiq Syamsul Hidayat  
 Simi Bajaj  
 Tad Gonsalves  
 Tahere Royani  
 Terumasa Aoki  
 Vishal Zinjuvadia  
 William R Simpson  
 Wu Yung Gi  
 Yahya Slimani  
 Yuhanis Binti Yusof  
 Zaw Zaw Htike

MUT University, Iran  
 University of Essex, United Kingdom  
 MariaDB Corporation, Finland  
 University of Sao Paulo, Brazil  
 University of Canberra, Australia  
 Chemic Engineers Inc, United States  
 Islamic Azad University, Iran  
 California State University East Bay, USA  
 King Abdulaziz Uni KAU, Saudi Arabia  
 University of Gondar, Ethiopia  
 Islamic Azad University, Iran  
 Saigon International University, Vietnam  
 M'Hamed Bougara University, Algeria  
 Beni Suef University, Egypt  
 Higher Colleges of Technology, UAE  
 University of Tehran, Iran  
 University of Uyo, Nigeria  
 University of Rajshahi, Bangladesh  
 University of Siegen, Germany  
 Jackson State University, USA  
 Qassim Private Colleges, Saudi Arabia  
 Alcorn State University, USA  
 Digital Compression Technology, USA  
 Afyon Kocatepe University, Turkey  
 Islamic Azad University, Iran  
 University of Babylon, Iraq  
 Universiti Sains Malaysia, Malaysia  
 City University London, UK  
 Moulay Ismail University, Morocco  
 Innov'com Laboratory, Tunisia  
 Wroclaw University of Technology, Poland  
 Alexandria University, Egypt  
 National Chiao Tung University, Taiwan  
 University of Florida, USA  
 Universite de Lorraine, France  
 Technical and vocational University, Iran  
 University of Malaya, Malaysia  
 University of Tasmania, Australia  
 Semarang State Polytechnic, Indonesia  
 Western Sydney University, Australia  
 Sophia University, Japan  
 University of Sabzevar, Iran  
 Tohoku University, Japan  
 Dell, United States  
 Institute for Defense Analyses, USA  
 Chang Jung Christian University, Taiwan  
 Faculty of Sciences of Tunis, Tunisia  
 Universiti Utara Malaysia, Malaysia  
 International Islamic University, Malaysia

## **Technically Sponsored by**

**Networks & Communications Community (NCC)**



**Computer Science & Information Technology Community (CSITC)**



**Digital Signal & Image Processing Community (DSIPC)**



## **Organized By**



**Academy & Industry Research Collaboration Center (AIRCC)**



## TABLE OF CONTENTS

### **The Fourth International Conference on Database and Data Mining (DBDM 2016)**

<b>Predictive Analytics in Healthcare System Using Data Mining Techniques.....</b>	<b>01 - 09</b>
<i>Basma Boukenze, Hajar Mousannif and Abdelkrim Haqiq</i>	

<b>Multicriteria Decision Aided System for Ranking Industrial Zones (RPRO4SIGZI).....</b>	<b>11 - 27</b>
<i>AissaTaibi and BaghdadAtmani</i>	

### **The Fourth International Conference on Instrumentation and Control Systems (CICS 2016)**

<b>A Comparative Overview to DTC for PMSM Driven Pure Electric Vehicles.....</b>	<b>29 - 36</b>
<i>Fatih Korkmaz</i>	

<b>A New Approach to DTC Method for BLDC Motor Adjustable Speed Drives.....</b>	<b>37 - 44</b>
<i>Fatih Korkmaz</i>	

### **The Third International Conference on Signal Processing (CSIP 2016)**

<b>Detection of Automatic The VOT Value for Voiced Stop Sounds in Modern Standard Arabic (MSA).....</b>	<b>45 - 56</b>
<i>Sulaiman S. AlDahri</i>	

<b>Approaches in Using Expectation Maximization Algorithm for Maximum Likelihood Estimation of the Parameters of a Constrained State Space Model with an External Input Series.....</b>	<b>57 - 66</b>
<i>Chengliang Huang, Xiao-Ping Zhang and Fang Wang</i>	

<b>A Novel ENF Extraction Approach for Region-of-Recording Identification of Media Recordings.....</b>	<b>97 - 108</b>
<i>Majed El Helou, Abdel Wahab Turkmani, Rawan Chanouha and Samer Charbaji</i>	

<b>Automatic Translation of Arabic Sign to Arabic Text (ATASAT) System.....</b>	<b>109 - 122</b>
<i>Abdelmoty M.Ahmed, Reda Abo Alez, Muhammad Taha and Gamal Tharwat</i>	

## **The Fourth International Conference of Artificial Intelligence and Fuzzy Logic (AI&FL 2016)**

<b>MLTDD : Use of Machine Learning Techniques for Diagnosis of Thyroid Gland Disorder.....</b>	<b>67 - 73</b>
<i>Izdiyar Al-muwaffaq and Zeki Bozkus</i>	

## **The Fourth International Conference of Soft Computing (SCOM 2016)**

<b>Web Service Composition in Dynamic Environment: A Comparative Study.....</b>	<b>75 - 84</b>
<i>Aram AlSedrani and Ameer Touir</i>	

## **The Fourth International Conference on Computational Science and Engineering (CSE 2016)**

<b>First-Order Mathematical Fuzzy Logic with Hedges.....</b>	<b>85 - 96</b>
<i>Van-Hung Le</i>	

## **The Third International Conference on Computer Networks & Communications (CCNET-2016)**

<b>E-Learning in Nigeria : Current Implementations and Challenges.....</b>	<b>123 - 130</b>
<i>Oluwatosin Temitope Adierinola, Ojonukpe Ekwuche and Abimbola Afolayan</i>	

# PREDICTIVE ANALYTICS IN HEALTHCARE SYSTEM USING DATA MINING TECHNIQUES

Basma Boukenze<sup>1\*</sup>, Hajar Mousannif<sup>2</sup> and Abdelkrim Haqiq<sup>3</sup>

<sup>1</sup>Computer, Networks, Mobility and Modeling laboratory  
FST, Hassan 1st University, Settat, Morocco  
basma.boukenze@gmail.com

<sup>2</sup>LISI Laboratory, FSSM Cadi Ayyad University, Marrakech 40000, Morocco  
mousannif@uca.ac.ma

<sup>3</sup>Computer, Networks, Mobility and Modeling laboratory  
FST, Hassan 1st University, Settat, Morocco  
e-NGN Research Group, Africa and Middle East  
ahaqiq@gmail.com

## ABSTRACT

*The health sector has witnessed a great evolution following the development of new computer technologies, and that pushed this area to produce more medical data, which gave birth to multiple fields of research. Many efforts are done to cope with the explosion of medical data on one hand, and to obtain useful knowledge from it on the other hand. This prompted researchers to apply all the technical innovations like big data analytics, predictive analytics, machine learning and learning algorithms in order to extract useful knowledge and help in making decisions. With the promises of predictive analytics in big data, and the use of machine learning algorithms, predicting future is no longer a difficult task, especially for medicine because predicting diseases and anticipating the cure became possible. In this paper we will present an overview on the evolution of big data in healthcare system, and we will apply a learning algorithm on a set of medical data. The objective is to predict chronic kidney diseases by using Decision Tree (C4.5) algorithm.*

## KEYWORDS

*big data, big data analytics, machine learning, Healthcare, learning algorithm, C4.5*

## 1. INTRODUCTION

There is an almost universal definition shared with proponents of the ideology of big data, it is that "Big Data sets a situation in which data sets have increased at such huge sizes that conventional technologies of information, can no longer manage them effectively, either the size or the extent and the growth of the data set" [1].

The world has become submerged by a large amount of data. Every moment is equivalent to the generation of thousands data. All sectors and all their activities are involved due to digitization,

David C. Wyld et al. (Eds) : DBDM, CICS, CSIP, AI&FL, SCOM, CSE, CCNET-2016

pp. 01–09, 2016. © CS & IT-CSCP 2016

DOI : 10.5121/csit.2016.60501

the introduction of information technology as an effective tool, and the Internet which is becoming a very important user interface for daily interactions [2]. However, these generated data become more and more difficult to manage in terms of volume, variety and velocity [3]. This gave birth to a new domain named big data. In 2008, Gartner used for the first time the term "Big Data", in reference to the explosion of digital data and quoted it will impact the way we work [4].

"Big Data" and "analysis of big data" are inseparable. This reflects the common opinion that "Big data" does not refer to the problem of information overload, but refers also to the analytical tools used to manage the flow of data and transform the flood in a source of useful information.

The medical field has its great contribution in this deluge of data because of some technological innovations in the field like cloud computing which has relocated the tests of care beyond the four walls of the hospital, and has made them available anywhere and anytime [5], laparoscopic surgery and robotic surgery, which replaced classical surgery [6], also smart homes which allow patients self-care and monitoring using simple devices that deliver results on specific physiological conditions. There are also smart applications or software that can analyze the body signals using integrated sensors with the aim of monitoring [7], as well as mHealth technologies that support new methods of biological, behavioural and environmental data collection. These include sensors that monitor the phenomena with a higher accuracy [8].

All these innovations participated to the explosion of medical data, by multiplying data sources and electronic medical records containing diagnostic Images, lab results, and biometric information that are generated and stored [8.9.10].

Researchers have deduced that this explosion of medical data has the potential to improve clinical decisions at the point of care. Doctor will become able to extract relevant knowledge for each patient, which gives better decisions and results [11].

On most of this, the term "analyzing medical data" and "predictive analytics" in Google Trends showed an impressive growth of interest from 2011 [12], because the process of analysis in the medical sector does not stop just at the level of the ability to manage large databases, but it exceeds this to the ability to retrieve future knowledge, which is encouraged by many researchers and experts. Seen that an analysis of the big data shows itself as the only solution able to solve all the problems of the medical sector [13] by:

- ✓ Providing best Service
- ✓ Monitoring quality in hospital
- ✓ Improving treatment processes
- ✓ Detection of disease earlier

There are many algorithms for classification and prediction applied to predict the most killer diseases like breast cancer, heart disease, motor neuron, and diabetes. In this present paper, we apply a decision tree classifier (C4.5) [14], which is among the most influential data mining algorithm in the research community and among the top 10 data mining algorithms. Our aim is to predict chronic kidney disease by this learning algorithm.

The rest of this paper is organized as follows:

- Section 2 is about related work.
- Section 3 presents the context of the experiment, metrics and research hypothesis.
- Section 4 presents Experimental results
- Section 5 discusses the results

Finally, section 6 concludes the paper.

## 2. RELATED WORKS

**Andrew Kusiak et al** [15] have used data pre-processing, data transformations, and a data mining approach to elicit knowledge about the interaction between many of measured parameters and patient survival. Two different data mining algorithms were engaged for extracting knowledge in the form of decision rules. Those rules were used by a decision-making algorithm, which predicts survival of new unseen patients. Important parameters identified by data mining were interpreted for their medical significance. They have introduced a concept in their research work have been applied and tested using collected data at four dialysis sites. The approach presented in their paper reduces the cost and effort of selecting patients for clinical studies. Patients can be chosen based on the prediction results and the most important parameters discovered.

**Abhishek et.al** [16] have used two neural network techniques, Back Propagation Algorithm (BPA), Radial Basis Function (RBF) and one non-linear classifier Support Vector Machine (SVM) and compared in accordance with their efficiency and accuracy. They used WEKA 3.6.5 tool for implementation to find the best technique among the above three algorithms for Kidney Stone Diagnosis. The main purpose of their thesis work was to propose the best tool for medical diagnosis, like kidney stone identification, to reduce the diagnosis time and improve the efficiency and accuracy. From the experimental results they concluded, the back propagation (BPA) significantly improved the conventional classification technique for use in medical field.

**Andrew Kusiak et al** [17] have used data pre-processing, data transformations, and data mining approach to elicit knowledge about the interaction between many of measured parameters and patient survival. Two different data mining algorithms were employed for extracting knowledge in the form of decision rules. Those rules were used by a decision-making algorithm, which predicts survival of new unseen patients. Important parameters identified by data mining were interpreted for their medical significance. They have introduced a new concept in their research work, it has been applied and tested using collected data at four dialysis sites. The approach presented in their paper reduced the cost and effort of selecting patients for clinical studies. Patients can be chosen based on the prediction results and the most significant parameters discovered.

**Ashfaq Ahmed K et.al**, [18] have presented a work using machine learning techniques, namely Support Vector Machine [SVM] and Random Forest [RF]. These were used to study, classify and compare cancer, liver and heart disease data sets with varying kernels and kernel parameters. Results of Random Forest and Support Vector Machines were compared for different data sets such as breast cancer disease dataset, liver disease dataset and heart disease dataset. The results with different kernels were tuned with proper parameter selection. Results were better analyzed to establish better learning techniques for predictions. It is concluded that varying results were observed with SVM classification technique with different kernel functions.

**Sadik Kara et.al** [19] had concentrated on the diagnosis of optic nerve disease through the analysis of pattern electroretinography (PERG) signals with the help of artificial neural network (ANN). Implemented Multilayer feed forward ANN trained with a Levenberg Marquart (LM) back propagation algorithm. The end results were classified as healthy and diseased. The stated results shown that the proposed method PERG could make an effective interpretation.

With respect to all related work mentioned above, our work is predicting disease using chronic kidney failure datasets by C4.5 algorithm.

### 3. EXPERIMENT

In this work, we will apply C4.5, a learning algorithm that will make classification and prediction on a database to extract knowledge and classify patients into two categories: chronic kidney disease (ckd) and not chronic kidney disease (notckd).

#### 3.1 Experiment environment:

In this study, we use the Waikato Environment for Knowledge Analysis (Weka). It is a comprehensive suite of Java class libraries that implement many algorithms for data mining clustering, classification, regression, analysis of results. This platform gives to researchers a perfect environment to implement and evaluate their classification model comparing to TANAGRA or ORANGE [20]

#### 3.2 Chronic kidney disease dataset

We used the database Chronic Kidney Disease Dataset from UCI Machine Learning Repository [21]. This database contains 400 instances and 24 integer attributes, two class (chronic kidney disease (ckd), not chronic kidney disease (notckd)). Table 1 describes the attributes of the database, while Table 2 describes the distribution of classes.

Table1 : Information attribute

Attribute	Representation	Information attribute	Description
Age	Age	Numerical	Years
Blood pressure	Bp	Numerical	Mm/Hg
Specific gravity	Sg	Nominal	1.005,1.010,1.015,1.020,1.025
Albumin	Al	Nominal	0.1.2.3.4.5
Sugar	Su	Nominal	0.1.2.3.4.5
Red blood cells	Rbc	Nominal	Normal, abnormal
Pus cell	Pc	Nominal	Normal, abnormal
Pus cell clumps	Pcc	Nominal	Present, notpresent
Bacteria	Ba	Nominal	Present, notpresent
Blood glucose random	Bgr	Numerical	Mgs/dl
Blood urea	Bu	Numerical	Mgs/dl
Serum creatinin	Sc	Numerical	Mgs/dl
Sodium	Sod	Numerical	mEq/L
Potassium	Pot	Numerical	mEq/L

Haemoglobin	Hemo	Numerical	Gms
Packed cell volume	Pcv	Numerical	
White blood cell count	Wc	Numerical	Cells/cumm
Red blood cell count	Rc	Numerical	Millions/cmm
Hypertension	Htn	Nominal	Yes, no
Diabetes mellitus	Dm	Nominal	Yes, no
Coronary artery disease	Cad	Nominal	Yes, no
Appetite	Appet	Nominal	Good, poor
Pedal edema	Pe	Nominal	Yes, no
Anemia	Ane	Nominal	Yes, no
Class	Classe	Nominal	Ckd notckd

Table 2 : Class distribution

	Class	Distribution
<u>1</u>	Ckd	250 (62.5%)
<u>2</u>	Notckd	150 (37.5%)

### 3.3 Metrics and research hypotheses

To understand classifier's behaviour we should calculate those metrics, we use hypothesis below:

- True positive (TP) is the number of positive samples correctly predicted.
- True negative (TN) is the number of negative samples correctly predicted
- False negative (FN) is the number of positive samples wrongly predicted.
- False positive (FP) is the number of negative samples wrongly predicted as positive.

Table 3. Metric and research hypotheses

Metric	Description	Formula	
Accuracy	Number of correct predictions from all predictions made.	$\frac{TP + TN}{TP + FP + TN + FN}$	(1)
Sensitivity	Proportion of positives predictions that are correctly identified.	$\frac{TP}{TP + FN}$	(2)
Specificity	Proportion of negatives predictions that are correctly identified	$\frac{TN}{FP + TN}$	(3)
Precision	Positive predictive values	$\frac{TP}{TP + FP}$	(4)
Mean Absolute Error (MAE)	Comparison between forecasts or predictions and the eventual outcomes	$\frac{FP + FN}{TP + FP + TN + FN}$	(5)
F-measure	Combination of precision and recall.	$\frac{2 * Precision * Sensitivity}{Precision + Sensitivity}$	(6)

Another important metric Confusion Matrix is taken into account. It is a visualization tool which is commonly used to present the accuracy of the classifiers in classification. The columns represent the predictions, and the rows represent the actual class as shown in Table 4

Table 4. Confusion matrix description.

		Predicted	
		Positive	Negative
Actual	Positive	TP	FN
	Negative	FP	TN

## 4. EXPERIMENTAL RESULTS

In order to apply our classifier and evaluate its performance, we apply the 10-fold cross validation test which is a technique that splits the original set into a training sample to train the model, and a test set to evaluate it. After applying the pre-processing and preparation methods, we try to analyse the data visually and figure out the distribution of values in terms of performance and accuracy of the model.

Table 5 : C4.5 performance

<b>Evaluation criteria</b>	C4.5
<b>Time to build model (s)</b>	0.08
<b>Correctly classified instances</b>	396
<b>Incorrectly classified instance</b>	4
<b>Accuracy</b>	63%
<b>Error</b>	0.37

Table 6 : Simulation error

<b>Evaluation criteria</b>	C4.5
<b>Kappa statistic</b>	0.97
<b>Mean absolute error</b>	0.02
<b>Root mean squared error</b>	0.08
<b>Relative absolute error %</b>	4.79
<b>Root relative squared error %</b>	16.66



Table 7 : Accuracy measures by class

	<b>TP</b>	<b>FP</b>	<b>precision</b>	<b>recall</b>	<b>F-measure</b>	<b>Class</b>
C4.5	0.99	0.02	0.98	0.99	0.99	Ckd
	0.98	0.004	0.99	0.98	0.98	Notckd

Table 8 : Diffusion Matrix

	<b>Ckd</b>	<b>NotCkd</b>	
C4.5 (J48)	249	1	<b>Ckd</b>
	3	147	<b>Notckd</b>

## 5. DISCUSSION

we can say that the C4.5 classifier is powerful, according to the number of correctly classified instances (396) and just 4 instances misclassified, This is mentioned with a low error rate (0.37 ) (see Table 5) . We can also deduce that this algorithm is excellent because of the value of KS = 0.97, which reflects the performance and accuracy of the classifier (see table 6). Fromtable 7, we deduce that C4.5 achieves best results regarding precision (0.98 ckd and 0.99 notckd), and sensitivity or (recall) (0.99ckd and 0.98 notckd ).

C4.5 has proved its performance as a powerful classifier in term of accuracy and the minimum execution time, which makes it a good classifier to be used in the medical field for classification and prediction.

## 6. CONCLUSION

As conclusion, the application of data mining techniques for predictive analysis is very important in the health field because it gives us the power to face diseases earlier that threaten the human being; child, young and old people, through the anticipation of cure and helping in decision-making. In this work we used a learning algorithm C4.5 to predict patients with chronic kidney failure disease (ckd), and patients who are not suffering from this disease (notckd). The classifier used proved its performance in predicting with best results in terms of accuracy and minimum execution time.

This is a major challenge in the medical field and pushes us to increase our efforts to develop machine learning methods, to exploit information intelligently and extract the best knowledge.

## REFERENCES

- [1] franck Ohlhorst, January 2013 ' Big Data Analytics: Turning Big Data into Big Money', ISBN: 978-1-118-14759-7, pp 176 .
- [2] Samson Oluwaseun, F., Serdar , S., and Vanduhe,V ., (2014)," Advancing big data for humanitarian needs ", Procedia Engineering, vol . 78,N., pp 88-95

- [3] Amir, G., Murtaza, H., (2015), "Beyond the hype: Big data concepts, methods, and analytics», International journal of Information Management, vol . ,pp 137-144.
- [4] H., Chen, H. L., Chiang, C., Storey, (2012), 'BUSINESS INTELLIGENCE AND ANALYTICS: FROM BIG DATA TO BIG IMPACT', MIS Quarterly, Vol. 36 ,No. 4, pp. 1165-1188.
- [5] Jonathan Northover, Brian McVeigh, Sharat Krishnagiri. Healthcare in the cloud: the opportunity and the challenge. MLD.  
Available at [http://www.sunquestinfo.com/images/uploads/CMS/445/mlo\\_02-12014\\_healthcare\\_in\\_the\\_cloud.pdf](http://www.sunquestinfo.com/images/uploads/CMS/445/mlo_02-12014_healthcare_in_the_cloud.pdf)
- [6] Gabriel I. Barbas, Sherry A. Glied, (2010),'' New Technology and Health Care Costs — The Case of Robot-Assisted Surgery''; the new England journal of medicine , N°, 363, pp 707-704 . Available at <http://www.nejm.org/doi/full/10.1056/NEJMp1006602>
- [7] Marianthi Theoharidou, Nikos Tsalis, ''Smart Home Solutions for Healthcare: Privacy in Ubiquitous Computing Infrastructures''.  
Available online at <http://www.cis.aueb.gr/Publications/Smart%20Home%20-%20Site%20TR.pdf>
- [8] Steve G. Peters, James D. Buntrock,(2014),''Big Data and the Electronic Health Record'', Ambulatory Care Manage , Vol. 37, No. 3, pp. 206–210
- [9] R. Weil, (2014),'' Big Data In Health: A New Era For Research And Patient Care Alan R. Weil'', Health Affair, Vol. 33, N° 7, pp 1110.
- [10] Peter Groves; Basel Kayyali, ( 2013),'' The 'big data' revolution in healthcare'', McKinsy and Company. Center for US Health System Reform Business Technology Office. Available at <http://digitalstrategy.nl/wp-content/uploads/E2-2013.04-The-big-data-revolution-in-US-health-care-Accelerating-value-and-innovation.pdf>.
- [11] T., Huang, L., Lan, (2015), ''Promises and Challenges of Big Data Computing in Health Sciences'', Big Data Research vol. 2, pp 2-11 available at <http://dx.doi.org/10.1016/j.bdr.2015.02.002>
- [12] Khurshid R., G., Kai, Z., John T., W., and Charles P., F., (2014), ''Harnessing Big Data for Health Care and Research Are Urologists Ready? '' , Journal of European Urology, vol. N., pp 1-3
- [13] Wullianallur Raghupathi,Viju Raghupathi, (2014),''Big data analytics in healthcare: promise and Potential'', Health Information Science and Systems.  
Available at <http://www.biomedcentral.com/content/pdf/2047-2501-2-3.pdf>
- [14] Rashedur M. Rahman, Fazle Rabbi Md. Hasan 'Using and comparing different decision tree classification techniques for mining ICDDR, B Hospital Surveillance data'', Elsevier, Vol. 38, pp 11421–11436
- [15] AndrewKusiak, Bradley Dixonb, Shital Shaha, (2005),'' Predicting survival time for kidney dialysis patients: a data mining approach'', Elsevier Publication, Computers in Biology and Medicine, Vol. 35, pp 311–327
- [16] Abhishek, Gour Sundar Mitra Thakur, Dolly Gupta, (2012) "Proposing Efficient Neural Network Training Model for Kidney Stone Diagnosis", International Journal of Computer Science and Information Technologies, Vol. 3 (3), pp 3900-3904

- [17] Andrew Kusiak, Bradley Dixonb, Shital Shaha, (2005) “Predicting survival time for kidney dialysis patients: a data mining approach”, Elsevier Publication, Computers in Biology and Medicine ,Vol.35, pp 311–327
- [18] Ashfaq Ahmed K, Sultan Aljahdali and Syed Naimatullah Hussain, (2013) “Comparative Prediction Performance with Support Vector Machine and Random Forest Classification Techniques”, International Journal of Computer Applications Vol. 69, No.11, pp 12-16
- [19] Sadik Kara, Aysegul Guvenb, Ayse OztUrk Onerc, (2006) “Utilization of artificial neural networks in the diagnosis of optic nerve diseases”, Elsevier Publication, Computers in Biology and Medicine, Vol. 36, pp 428–437
- [20] M Hall, E Frank, G Holmes, B Pfahringer, ( 2009), ‘The WEKA data mining software: an update’, Volume 11, Issue 1, pp 10-18
- [21] “UCI Machine Learning Repository: Kidney failure Data Set [Online].  
Available: [https://archive.ics.uci.edu/ml/datasets/Chronic\\_Kidney\\_Disease#](https://archive.ics.uci.edu/ml/datasets/Chronic_Kidney_Disease#)

## AUTHOR

Basma BOUKENZE, PhD student at the Faculty of Science and Technology in Settati, Morocco. after obtaining a degree in computer Science in 2009, and a master degree in engineering network and System in 2011, has continued studies and currently registered in doctoral science and technology formation at the Faculty of Science and technology Settati, Morocco, member of The Mathematical Research structure and applied computing. Laboratory Computer Networks, Mobility and modelling.



*INTENTIONAL BLANK*

# MULTICRITERIA DECISION AIDED SYSTEM FOR RANKING INDUSTRIAL ZONES (RPRO4SIGZI)

AissaTaibi, BaghdadAtmani

Laboratoire d'Informatique d'Oran – LIO  
Université d'Oran 1 Ahmed Benbella  
taibiissa@yahoo.fr  
atmani.baghdad@gmail.com

## **ABSTRACT**

*Integration of Geographic Information Systems (GIS) and multi-criteria decision analysis (MCDA) is a privileged and indispensable way to evolve GIS into real decision support systems. RPRO4SIGZI, the system proposed in this paper allows, from a detailed study of geographical, environmental and socio-economic criteria to cooperate GIS and multi-criteria decision analysis method for spatial choosing of the right site for installing industrial projects. The result obtained by RPRO (Ranking PROMETHEE) for ranking industrial zones in western Algeria is refined by a viewing SIGZI (Geographic Information System for Industrial Zones). The RPRO unit rank industrial zones using the outranking PROMETHEE II method issue from European school and SIGZI module to the visualization of these zones on the map. RPRO4SIGZI system was designed for the evaluation of a new methodology of multi-criteria analysis guided by data mining. The objective is to show how data mining is used to model the preferences of the decision maker tainted with subjectivity and hesitance to generate suitable performance tables. Only RPRO4SIGZI system is presented in this paper.*

## **KEYWORDS**

*Geographic Information System (GIS), multi-criteria decision analysis (MCDA), Industrial Zones, MCDA-GIS integration, Cartography, PROMETHEEII.*

## **1. INTRODUCTION**

The geo-decisional study of zonal aptitude for selecting the location of new housing sites, industries and services is essential and is a real spatial decision problem. Policy makers should act early based on depth analysis of the environmental, socioeconomic and other criteria (factors, constraints) to carefully carry out their decisions to end without risks. This work is to rank the industrial zones of western Algerian using the outranking method PROMETHEE II [1]. It follows a preliminary choice based on a zonal aptitude analysis using non-compensatory Aggregation methods. Each zone is a space action since action to take is spatial if it is defined by its geographical location, shape and spatial relations [2]. Most judgment criteria have a geographical character. The specifics of this kind of problems is in favor of integration between GIS and

MCDA where the adoption of this approach. The researchers focused on this approach since 1999, hundreds of articles have been published [3]. The conceptual idea on which is based MCDA-GIS integration work is to use the functions of GIS to prepare inputs necessary for the MCDA methods and GIS presentation potentialities to visualize the results of the analysis on the map [4]. In literature there are many definitions for GIS, a coherent definition with this study is that of Marc Souris [5]: "A geographic information system (GIS) is primarily a database management system capable of managing localized data, and therefore able to capture, store, extract (especially on geographical criteria) to query and analysis, and finally to represent and map. The displayed objective is essentially for synthesizing, and allowing data management as decision support". The input for the PROMETHEE II method is a performance table which contains values (score) for each action (industrial zone) relative to the set of criteria plus the additive information about criteria necessary to use this method.

Evaluation of an action relative to geographic criteria is based on an important feature of GIS: Mapping, this discipline is the first step of spatial analysis; a map is a model of reality that contains the geometric representation of objects and categories of objects with graphical and semiotic logic [6]. For example the seismic value of a zone derives from its geographical position on the seismic map of Algeria. As a result of total ranking, the best zone is obtained with visualization on map before and after ranking. The adoption of GIS-MCDA approach in this case has confronted us several problematic such as the choice of the appropriate MCDA method? The subjectivity and hesitance of decision makers? ... To solve the second problem we will engage data mining to model the preferences of the decision maker and generate performance tables. The rest of this article is presented as follows. Section 2 is devoted to the problem and some related works, Section 3 and 4 is devoted to the GIS-MCDA approach and PROMETHEE II method, the proposed model is in section 5, a case study is illustrated in section 6, we end with a conclusion and perspectives.

## **2. PROBLEMATIC AND RELATED WORK**

Problems related to the evolution of urban fabric, building new cities, and creation of new industrial zones is analysis problems of zonal aptitudes in a larger context of decision support. The spatial decisional study of zonal aptitude to select the location of new housing sites, industries and services is essential and is a real problem of spatial decision. Anarchic zoning to solve such problems can cause epidemiological change and deterioration in the health of citizens. The linear model of Simon and its extensions are insufficient to respond to the complexity of these problems [7]. Geographic information systems (GIS) are important for the analysis of decision problems where the geographic components of the data are considered. GIS is primarily an explanatory help tool for decision. Multi-criteria decision analysis methods MCDA provides the techniques necessary to structure and evaluate alternatives in decision-making problems according to defined set of criteria and proposed weighting. GIS research areas and multi-criteria decision analysis methods for decision aids are distinct but they help each other to get the best spatial decision problems solutions.

MCDA-GIS integration works have increased since 1990. Most of this work since 1990 until 2004 were identified and categorized in [3]. In [8] there is recognition of the variety and complexity of multi-criteria analysis methods, to remedy this, the authors have made scanning and classification of all methods. The classification leads to the following classes:

- Non-compensatory Aggregation methods
- Weighting methods (AHP...)
- Compensatory aggregation methods
- Outranking methods (ELECTRE, PROMETHEE ...)
- Mathematical Programming Method.
- Heuristic methods (MOLA, GA, SA ...)

More recent work as in [9] where comparing AHP and PROMETHEE II for selecting best techniques used in building conclude that PROMETHEE II is the appropriate method since its results are consistent, easy to understand and requires less information from decision makers. In [10] the objective was to estimate the ecological values of the Piedmont region of northern Italy and generate maps for use as assistance with the decision variables in the field of planning and land management to protect environment and eco systems. In [11] the purpose was to find a suitable geography to lay the vegetable water (waste water from the olive crushing). In [12] the objective is to measure the vulnerability of forest habitat interfaces, the authors used the AHP method (SAATY, 1980) to treat six decision criteria (layout, topography, vegetation structure, habitat structure, properties of buildings, socio-economic structures). They proceeded to map the vulnerability of each criterion by using ARCGIS. In [13] the goal is to alleviate the dissatisfaction of some group of citizens in Quebec when planning a linear park section of less than 15 km from the new port area of Quebec. Another work that is within the scope of energy diversification is to design a MCDA-GIS model to guide a project on wind energy in Canada [14].

### 3. THE INTEGRATION BETWEEN GIS AND MCDA

Spatial decision problems constitute a large part of decision-making problems. This type of problems is characterized by geographic data with spatial attributes (coordinates, shape...). Spatial problems whose complexity is related to the heterogeneity of data and concepts mobilized to model the geographical reality often have a multi-criteria aspect [2]. The complexity of these problems comes from (i) the multidimensional and interdisciplinary nature which is difficult to formalize, (ii) the involvement of several persons and institutions, generally with preferences and diverging objectives, (iii) the need to define multiple conflicting criteria whose importance is not the same [4]. The solution of such problems generates a spatial multi-criteria decision. Most of the involved geographical criteria must be mapped for decision makers. Alternative also, must be mapped and displayed. The spatial multi-criteria analysis for decision aid (SMADA) needs spatial and descriptive data both for imposed criteria and considered alternatives. All data are processed and aggregated in one hand by the MCDA using appropriate decision rules and by GIS for spatial analysis and mapping in the other hand, therefore the two tools are used interchangeably [3]. The conceptual idea on which the GIS-MCDA integration is based is to use the capabilities of GIS to prepare inputs necessary for the implementation of a multi-criteria method and exploit the potential of GIS for presentation to see the results of the analysis on maps [4]. Below some arguments in favor of the integration or coupling [4], [3], [2].

- To solve a spatial decision problem we need to consider both spatial and decision components problem.
- GIS is well suited for the representation of decision problems with spatial reference but it fails to take into account the multi-criteria decision dimension of the problem.

- MCDA allows formulation and modeling of spatial problems but it is limited to represent the spatial dimension of this problem.

GIS-MCDA approach is applied with a rate of 72.4% in the field of territorial and environmental management, transport, urban planning, waste management areas, hydrology and agriculture. 30% of treated decisional problems involve susceptibility or suitability analysis [3]. Among the questions related to the approach there is the lack of a related maintenance policy, absence of correlation between the problem, the aggregation rule and GIS. There is also an ambiguity related to the integration mode of the two tools. [4] Has proposed three integration mode (a: Indirect Integration, b: Built Integration, c: Complete integration).

In this paper a mixed integration is proposed, preparing geographic criteria to establish performance table is made by SIG independently (indirect integration) while visualizing function is integrated directly with the MCDA Module and is considered as a finality of the decision analysis (figure 1).

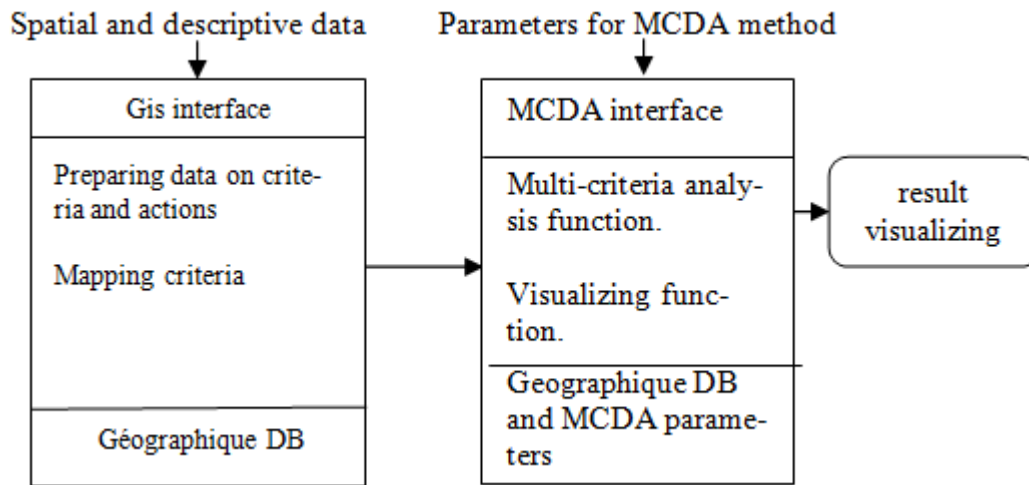


Figure 1. GIS-MCDA proposed integration.

#### 4. PROMETHEE

PROMETHEE I method for partial ranking and PROMETHEE II for total ranking were developed by JP Brans in 1982. Other extensions, PROMETHEE III and PROMETHEE IV and the GAIA interactive graphics module of this methodology were developed later. PROMOTHEE is successfully used in the fields of industrial location, investments, medicine and chemistry ... Its success returns to its user-friendly and there mathematical properties [1],[15]. It is an outranking method [18] based on dominance relationships among alternatives against criteria associated to a multicriteria problem (P: Preference I: indifference, R: In comparability):



$$\left\{ \begin{array}{l} \forall j : g_j(a) \geq g_i(b) \\ \exists k : g_k(a) > g_k(b) \end{array} \right\} \longleftrightarrow aPb \quad (1)$$

$$\forall j : g_j(a) = g_i(b) \longleftrightarrow aIb \quad (2)$$

$$\left\{ \begin{array}{l} \exists s : g_s(a) > g_s(b) \\ \exists r : g_r(a) < g_r(b) \end{array} \right\} \longleftrightarrow aRb \quad (3)$$

$g_j(a)$  denotes the evaluation of action  $a$  under criterion  $j$ . These relationships create a dominance graph. An appropriate multicriteria analysis method wish to enrich this graph, since the objective is to reduce the number of in-comparabilitys in order to optimize the decision. Consideration of in-comparabilitys is an asset because it reflects reality. Reduce the number of in-comparabilitys does not mean elimination of them. Other assets of PROMETHEE according to its inventor J.P. Brans [15] reside in the following requirements:

**R1:** The amplitude of the deviations between the alternatives within criterions is taken into account. It is denoted by:

$$d_j(a, b) = g_i(a) - g_j(b) \quad (4)$$

When  $d_j$  is negligible, dominance relation can be enriched by other means.

**R2:**  $g_i(a)$  is expressed with its own measurement unit, scale effects should be eliminated because it is not acceptable to obtain conclusions with scaling effects upon which the evaluations are expressed.

**R3:** The binary comparisons provide all possible information on the two compared alternatives ( $a$  is preferred then  $b$ ,  $a$  and  $b$  are indifferent,  $a$  and  $b$  are incomparable). The purpose is of course to reduce as much as possible the number of incomparability, but not when it is not realistic.

**R4:** The method must be understandable by the decision-makers. “Black box” procedures should be avoided.

**R5:** The technical parameters which have no significance for the decision maker must be rejected.

**R6:** An appropriate MCDA method must provide information about the confrontational nature of the criteria.

**R7:** The MCDA methods use the relative importance of the criteria through weights given by the decider tainted by his subjectivity and hesitation. A method that respects itself provides tools for the study of sensitivity.

In addition to the performance table necessary to the outranking methods, PROMETHEE requires two clear information, easy to find and assimilated by the decision maker for the best compromise solution :

#### 4.1 Information among the criteria

Described by the set  $\{w_j / j=1 \dots k\}$  of relative importance weights for different criteria. These are positive numbers independent of units of measurement with  $\sum w_j = 1$ . The decision maker is free to give these weights tainted with his hesitation and his subjectivity. If these notes are arbitrary they must be normalized by dividing each weight by the sum of all the others.

#### 4.2. Information on the criteria

PROMETHEE uses the difference between the evaluation of two alternatives a and b on any criterion j as in equation (4) to build a preference structure, this gap ( $d_j$ ) is reversed if the criterion is to minimize, cost criteria for example. The degree of preference is proportional to the degree of the gap, the preferences are measured by real numbers between 0 and 1. In the mind of the decision maker's preference between two alternatives a and b is a function of the gap  $d_j$  that is expressed mathematically by:

$$P_j(a, b) = F_j[d_j(a, b)] \quad (5) \quad \text{If the criteria is to maximize}$$

$$P_j(a, b) = F_j[-d_j(a, b)] \quad (6) \quad \text{If the criteria is to minimize}$$

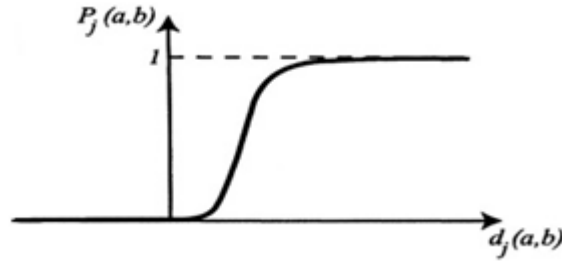


Figure 2 : Preference function [1]

The couple  $\{g_j(\cdot), P_j(a, b)\}$  is the generalized criterion associated to  $g_j(\cdot)$  criterion. It is the concept of generalization that characterizes PROMETHEE compared to other outranking methods. The generalized criterion function associated with preference, indifference and intermediate thresholds provides six types of preference functions. The behavior of each action overlooked to the others is enjoyed by three flows:

$$\text{The positive flow: } \varphi^+(a) = \sum_{x \in A} \pi(x, a) \quad (7)$$

$$\text{The negative flow: } \varphi^-(a) = \sum_{x \in A} \pi(a, x) \quad (8)$$

$$\text{The net flow: } \varphi(a) = \varphi^+(a) - \varphi^-(a) \quad (9)$$

The preference index for an alternative a compared to another b is denoted by:

$$\pi(a, b) = (1/m) \sum_{j=1}^k P_j(a, b) w_j \quad (10)$$

**m**: criteria number,  **$P_j(a, b)$**  : Preference of the action a over b,  **$w_j$** : Weight of criterion j.

The value of the net flow determines the rank of an alternative.

## 5. PROPOSED MODEL

The overall decisional system consists of three modules (Figure 3):

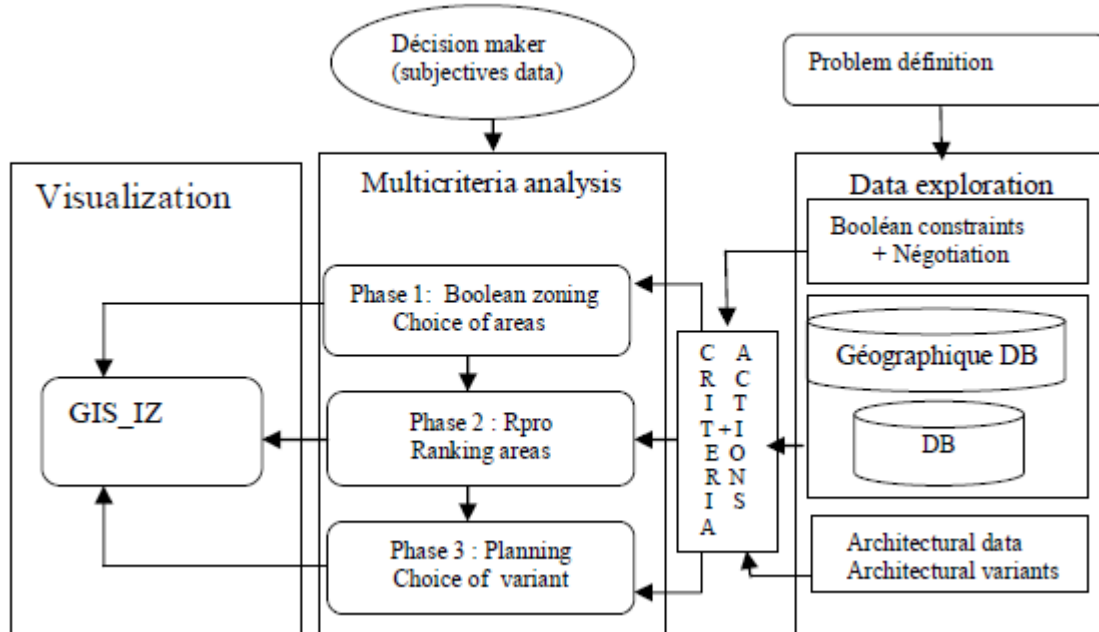


Figure 3: General System Architecture.

### 5.1. The visualization module:

Since actions in this study are all spatial, SIGZI ensures the display of these zones (actions) on the map of Algeria before and after each decision-making phase. To accomplish this task the vector mode is adopted, each industrial zone is a geographical entity of the abstract spatial type "POINT" it is implemented with geographical position (latitude and longitude).

### 5.2. Data Exploration module:

The main entities in our multi-criteria decision approach are the criteria and actions, data are collected from geographic and socioeconomic databases and from climate station as well as archives, criteria cards are built.

### 5.3. The Multi-criteria analysis module:

This is the main module for the solution of global decision problem, it operate in three phases (The aptitude study and geographical choice, ranking zones, choice of an architectural variant). Only Phase 2 (ranking zones) is explained in this paper.

### 5.3.1. First phase:

Conducted by ANIREF [16], is a zonal aptitude analysis. Such analysis is at the heart of the processes involved in planning and is a major component of spatial decision support [6]. The Boolean zoning used belongs to the class of non-compensatory aggregation methods [8] which operates according to rules such as: If (HT electric line > 10 m from the zone) then (suitable zone) else (unfit zone). The ability of a zone is calculated using the intersection of several indices.  $I_{APT,j} = C_{1,j} \cap C_{2,j} \cap \dots \cap C_{k,j}$  where  $C_{k,j}$  is the aptitude binary value of the criterion  $k$  for the zone  $j$ . The result is discussed, commented and complemented by negotiation.

### 5.3.2. Second phase:

It is the total ranking of zones using qualitative and quantitative values of the criteria. The total outranking method PROMETHEEII is used.

### 5.3.3. Third phase:

Consist to choose one of three available architectural variants, the selection criteria are the architecture, the management cost, the number of fragmented islands and the types of planned investments.

## 6. CASE STUDY

### 6.1. The set of actions:

Of the 39 industrial zones created through the entire national territory by ANIREF [16] our study has focused on the industrial areas of western Algerian. Each zone is an action (A1: Maghnia, Tlemcen. A2: Sidi Bel Abbès. A3: Ras Elma, Sidi Bel Abbès. A4: Sidi Ahmed, Saida. A5: Horchaia, Naama. A6: Tamazzoura, AinTémouchent. A7: Oggas Mascara. A8: El Haciane, Mostaganem. A9: Sidi Khettab, Relizane.).

### 6.2. The criteria:

The criteria used in this study were classified into three categories: natural constraints, the socio-economic and legal requirements and environmental constraints. According to these categories, 8 different evaluation criteria are defined. Figure 4 shows the hierarchy of criteria of judgment.

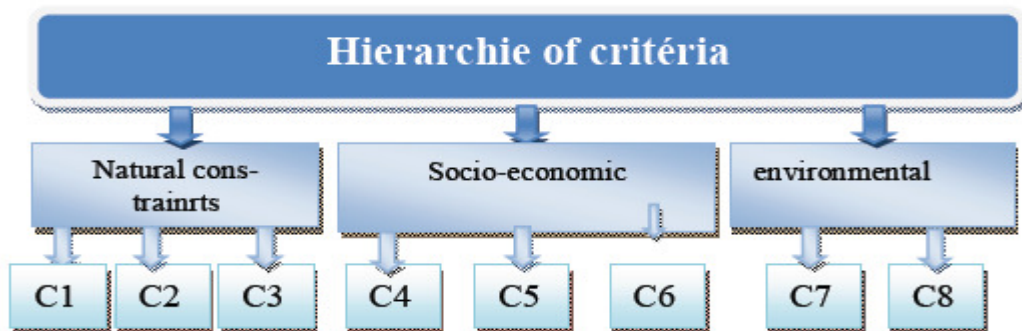


Figure 4: Hierarchy of judgment criteria.

- Criterion (C1): Seismicity.
- Criterion (C2): Climate Constraint: Rainfall.
- Criterion (C3): Climate Constraint: Temperature.
- Criterion (C4): Surface area.
- Criterion (C5): Management cost.
- Criterion (C6): Distance to transport networks.
- Criterion (C7): bioclimatic Constraint.
- Criterion (C8): Proximity to urban residential center.

To evaluate the different zones to rank on the basis of qualitative criteria, a rating scale (scale of 1 to 5) is associated with each qualitative criterion in order to make a measurable dimension. The principle used is to evaluate zones (actions) against the criterion on the basis of the mapping. The evaluation method is to analyze the geographical position of industrial zones on the corresponding thematic map of each geographical criterion (seismicity, humidity...).

### 6.2.1. Natural constraints:

**C1-Seismicity:** The seismic zoning of Algerian territory (Figure 5) reveals five seismic zones.

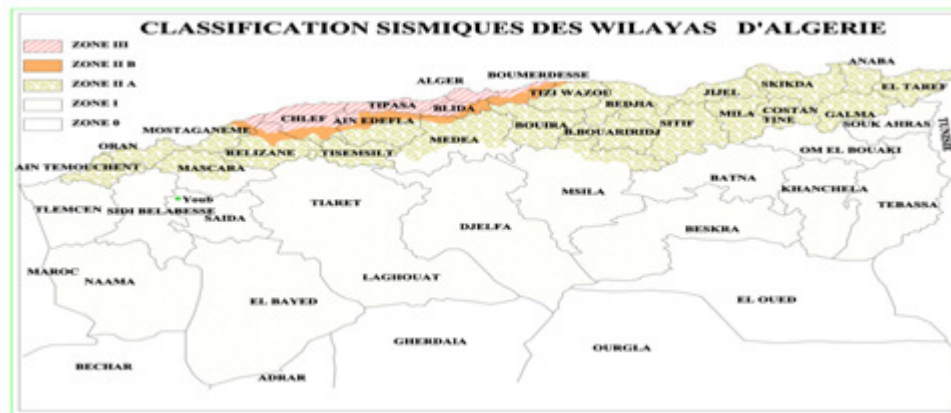


Figure 5: Seismic classification of Algeria [16].

Using the seismic map of Algeria and according to the proposed measure scale (Table 1) below contains actions values according to seismicity.

Table1: Evaluation of the actions according to the criterion of seismicity.

Actions	Seismicity	Numerical values
A1	Low to moderate	2
A2	Low to moderate	2
A3	Low to moderate	2
A4	Low to moderate	2
A5	Low	1
A6	Moderate	3
A7	Moderate	3
A8	Low to moderate	2
A9	Moderate	2

**C2, C3-** Climatic constraints: The average numerical values of these two criteria are taken from climate stations installed in the country.

Table2: Evaluation of the actions on criteria, rainfall and temperature

actions	Rainfall (mm)	temperature°C
A1	350	19
A2	310	24
A3	410	17
A4	380	19
A5	190	17
A6	400	18
A7	320	21
A8	350	20
A9	370	19

### 6.2.2 Socio-economic criteria:

**C4:** Area: This is quantitative information representing the area of each industrial zone.

**C5:** COST Management: This is quantitative information representing the management cost. Note that the location of the site (soil, slope, altitude ...) directly influences the lying and indirectly on the weight of this criterion.

**C6 -** Proximity to transport networks (roads, railway and airport): The evaluation of this criterion is done by comparing cartographically two thematic maps, the geographical situation of the zones in question with that of transport networks.

Table 3: Evaluation according to cost management, area and proximity to transport networks.

Actions	Cost planning (DA)	Area(Ha)	Proximity(m)
A1	900592576	104	2500
A2	867750000	100	4100
A3	523765223	60	5000
A4	867750000	100	6500
A5	1301625000	150	3500
A6	1778911797	205	3000
A7	851772119	98	8100
A8	1735585907	200	6500
A9	4338750000	500	3000

**C7-** environmental criteria: The map below shows the bioclimatic floors of Algeria

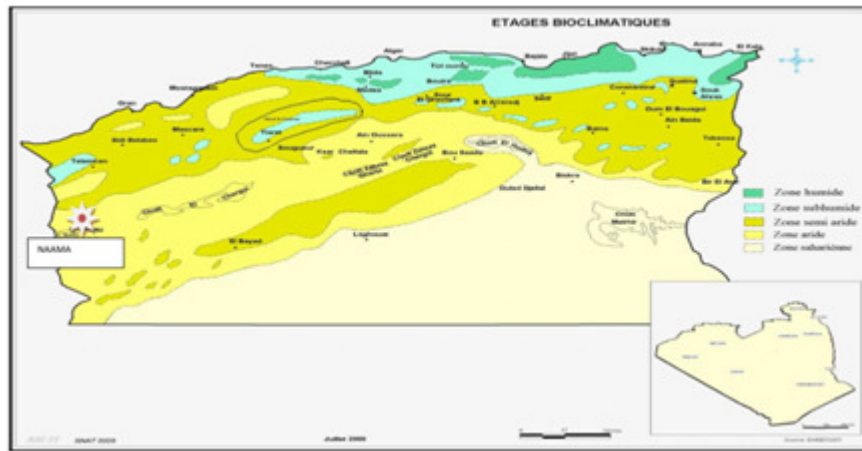


Figure 4: Bioclimatic Floors of Algeria [16].

**C8:** Proximity to the urban center of residence: This is the cause of noise pollution linked to increased traffic flows, propagation of harmful gases for respiratory health of citizens and liquid and solid industrial waste.

Table 4: Bioclimatic constraints and Proximity to urban centers.

Actions	Description	(Valeur Num) C7	Proximity(m) C8
A1	Semi aride zone	2	14000
A2	Semi aride zone	2	17000
A3	Semi arid zone	2	13500
A4	Semi aride zone	2	15000
A5	Arid zone	1	18000
A6	Semi arid zone	2	16500
A7	Semi arid zone	2	18300
A8	Semi aridezone	2	13000
A9	Arid Zone	1	17800



Weights for criteria are defined by the technical team of the ANIREF [16] direction following two steps:

- Classification of the eight criteria in descending order of importance according to a unanimous decision from a consultation among all members of the team (engineers, technicians and managers).
- The second step consists of distributing a set of 100 points between the various criteria. The values of the final weights are given in Table 5:

Table 5: Table of intra criteria weights

Criteria	Criteria description	weight(%)	Weight(point)
C1	Seismicity.	10%	10
C2	Climatic constraints: Rainfall	5%	5
C3	Climatic constraints: Temperature.	5%	5
C4	Acreage	20%	20
C5	COST	15%	15
C6	Proximity to transport networks	20%	20
C7	Bioclimatic constraints	5%	5
C8	Proximity to urban residential center	20%	20

Note that a criterion can be a factor to maximize or a constraint to minimize. The sense of each criterion was adopted in the opinion of the expert (table6).

After the evaluation of actions, weighting criterion and determination of the sense of each criterion we obtained the following performance table (table 6)

Table 6: Performance Table

Cr�terion/Action	C1	C2	C3	C4	C5	C6	C7	C8
A1	2	350	19	104	900592576	2500	3	14000
A2	2	310	24	100	867750000	4100	3	17000
A3	2	410	17	60	523765223	5000	3	13500
A4	2	380	19	100	867750000	6500	3	15000
A5	1	190	17	150	1301625000	3500	2	18000
A6	3	400	18	205	1778911797	3000	3	16500
A7	3	320	21	98	851772119	8100	3	18300
A8	4	350	20	200	1735585907	6500	3	13000
A9	3	370	19	500	4338750000	3000	2	17800
Criterion sense	Min	Min	Min	Max	Min	Min	Min	Max

### 6.3. Result

Although the significance of the result comes from the use of a validated method and a specific core GIS to data of the case study, remains a sensitivity analysis on preference and indifference thresholds to validate the stability of the solution. Here the indifference threshold is set at 5% of



the difference between the highest and lowest score while the preference threshold is set at 10% of the difference. Table 8 below shows zone rows obtained.

Table 8: Obtained ranking

Actions	Positive flow ( $\phi^+$ )	Negative flow ( $\phi^-$ )	Net flow ( $\phi$ )	Rank
A1	0.38531917	0.3-	0.08531916	4
A2	0.2971698	0.5250107	-0.2278409	8
A3	0.36762434	0.35000002	0.017624319	5
A4	0.23181818	0.4375	-0.20568182	7
A5	0.59375	0.2363376	0.3574124	1
A6	0.4321429	0.3094263	0.122716606	3
A7	0.3124035	0.37500003	-0.06259653	6
A8	0.22386363	0.58466977	-0.36080614	9
A9	0.5196429	0.24578992	0.27385297	2

## 6.4 Visualization:

### 6.4.1 Pre-visualization:

Pre visualization of industrial zones on the Algerian map before the multi-criteria decision analysis, the ranks are randomly assigned (Figure 5):

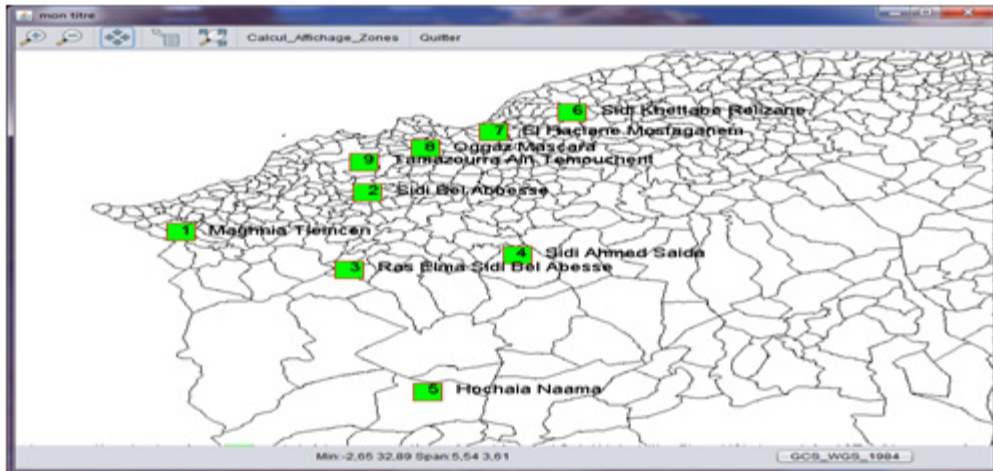


Figure 5: administrative Card with random rows.

### 6.4.2. Post Visualization:

Visualization of industrial zones on the map of Algeria after multi-criteria decision analysis with result ranks (Figure6):

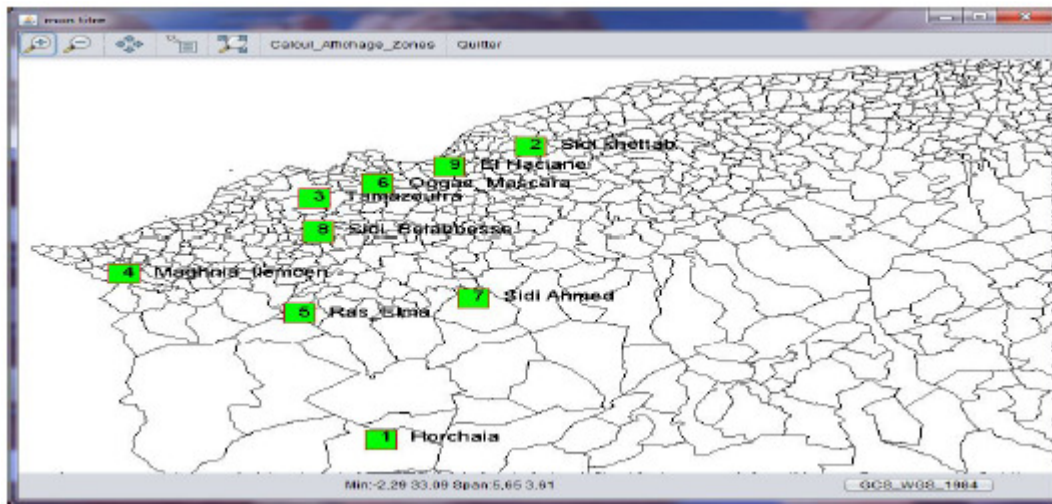


Figure 6: Viewing zone rows after the analysis.

## 7. CONCLUSION AND PERSPECTIVES

This work follows a zonal capability analysis based on Boolean non-compensatory multi-criteria methods. The constraints of selection of the zones were set by negotiation and by current legislation. The proposal in this paper is to start a second multi-criteria decision-analysis phase to consolidate those choices. The quantitative and qualitative information collected for each zone comforted the decision maker and has established a trust in him on the MCDA-GIS integration approach. This study allowed us to determine the usefulness of the approach for many sectors where the decision is important and dangerous, and intersects with the geography and even history. It is a contribution to make out the approach from the academic side to the field. The rank of an industrial area so obtained is an index that can:

- Call into question the choice of this zone.
- Alert the planners and builders of the area.
- Assign the area to adequate investment projects.

Our perspectives are to extend the study on all industrial zones at the national level which leads us to reconsider the choice of multi-criteria method. The choice of method is a crucial step, especially in our case study, four approaches can be used: Ad hoc, classification tree, multi-criteria method or expert systems [4].

To remedy the disadvantage of MCDA methods in modeling the preferences of decision makers with subjectivities and hesitance we decided to follow the approach proposed by [17] to test several data mining techniques.

## REFERENCES

- [1] J.P. Brans & Ph.Vincke, (1985). A preference Ranking Organisation Method:( The PROMETHEE Method for Multiple Criteria, Management Science, Vol. 31, No.6 (Jun., 1985), PP.647-656 Published by: INFORMS.
- [2] Raffaella Balzarini&Paule-Annick Davoine & Muriel Ney, (2012). Evolution et développement des méthodes d'Analyse spatiale multicritère pour des modèles d'aptitude : L'exemple des applications en Géosciences. Laboratoire d'Informatique de Grenoble (LIG) équipes Steamer et Metah.ESRI France, Département Education et Recherche.
- [3] J.Malczewski, (2006). GIS-based multi-criteria decision analysis: a survey of the literature, International Journal of Geographical Information, 20:7, 703-726.
- [4] Chakhar,Salem, (2006), Cartographie Décisionnelle Multicritère : Formalisation Et Implémentation Informatique, Thèse de doctorat , Université paris dauphine.
- [5] Marc.souris (2002),Les principes de systèmes d'information géographique : Principes algorithmes et architecture du système SAVANE, These de doctorat IRD ((Institut de recherche pour le développement) France.
- [6] Régis,Caloz & Claude collet, (2011). Analyse spatiale de l'information géographique, Presse polytechniques et universitaires romandes.
- [7] Fatima ZohraYounsi & Djamila Hamdadou & BouzianeBeldjilali, (2012). Proposition d'un Système Interactif d'Aide à la Décision Spatiale : Télédétection, SIG et Analyse Multicritère, Université d'Oran Es-Senia .
- [8] Randal Greene & Rodolphe Devillers & Joan E.Luther & BriaG.Eddy (2011), Gis-Based Multiple-Criteria Decision Analysis Department of geography, Memorial University of New-foundland, Canadian Forest service, Natural resources Canada. Geography compass 5/6.
- [9] Balali, Banafsheh Zahraie & Abbas Roozbahani, A Comparison of AHP and PROMETHEE Family Decision Making Methods for Selection of Building Structural System, American Jour-nal of Civil Engineering and Architecture, vol. 2, no. 5 (2014): 149-159.
- [10] Valentina Ferretti & Silvia Pomarico , integrating Multicriteria Analysis and Geographic In-formation Systems for studying ecological corridors in the Piedmont Region, 74th Meeting of the European Working Group "Multiple Criteria Decision Aiding".
- [11] S'habou R & ,Zairi M & Kallel A& Neji J. Ben Dhia H, (2011).Intégration du SIG et des méthodes d'analyse multicritère pour la gestion de la pollution: cas de stockage des margines Sfax Tunisie, Séminaire International, Innovation & Valorisation en Genie civil & Matériaux de construction, Rabat Maroc: INVACO2 N° : 50-309.
- [12] L. Pugnet & E.Maillé, (2013).Analyse multicritères pour l'évaluation de la vulnérabilité des interfaces habitat-foret, international conference on forest « fire risk modelling and mapping » Aix en Provence, France.
- [13] Carlo,Prévil & Marius,Thériault & Joël.Rouffignat, (2003).Analyse multicritère et SIG pour faciliter la concertation en aménagement du territoire : vers une amélioration du processus décisionnel. Les Cahiers de géographie du Québec, p. 35-61, N° 130 : Vol.. 47.

- [14] Vazquez Maria & L. WAAUB & Jean-Philippe CHAUMEL & Jean-Louis (2011), Analyse spatiale et approche d'aide multicritères et multi-acteurs à la négociation pour évaluer des scénarios d'implantation des parcs éoliens, publié dans "1ère Conférence Intercontinentale d'Intelligence Territoriale Interdisciplinarité dans l'aménagement et développement des territoires", Gatineau : Canada.
- [15] J-Pierre Brans & Bertrand Mareschal, PROMETHEE METHODS Centrum voor Statistiek en Operationeel Onderzoek Vrije Universiteit Brussel Pleinlaan 2, B-1050 Brussels & Service de Mathématiques de la Gestion Université Libre de Bruxelles Boulevard du Triomphe CP 210-01, B-1050 Brussels Belgium .
- [16] ANIREF (Agence Nationale d'Intermédiation et de Régulation Foncière) (2013). Etude préliminaire d'aménagement du parc industriel . Alger.
- [17] Kary Frakling (1996), Modélisation et apprentissage des préférences par réseaux de neurones pour l'aide à la décision multicritère. Thèse doctorat, Ecole nationale des mines de Saint-Etienne.
- [18] B.Roy, (1985), Méthodologie multicritère d'aide à la décision, Paris Economica.
- [19] B.taibi (2010) , L' analyse multicritère comme outil d'aide à la décision Application de la methode PROMETHEE etude de cas:l'entreprise SEROR, These de Magister ,Université de Tlemcen, Algerie.
- [20] Deliverable 4B, (2010). Multi-criteria analysis and ranking of alternative waste technologies/ management system, Rapport de recherche, Faculté des sciences Eljadida, National technical university of Athens, Municipality of the urban community of Azemmour, Maroc.
- [21] Egenhofer M.J A, (1989), Formal Definition of Binary Topological Relationships, Proceedings of the 3th International Conference on Foundations of data Organization and Algorithms. - Paris, France Lecture Notes in Computer Science 367, 1989. - pp. 457-472 .
- [22] H.Laurent, (2000). Systèmes d'évaluation et de classification multicritères pour l'aide à la décision. Thèse de doctorat Université Paris Dauphin .
- [23] Pugnet L. et E.Maillé, (2013). La modélisation et la cartographie des risques de forets, International conférence on Forest risk modelling and mapping.
- [24] S,BenMena, (2000). Introduction aux méthodes multicritères d'aide à la décision, B A S E Biotechnol. Agron. Soc. Environ.Unité de Mathématique. Faculté universitaire des Sciences agronomiques de Gembloux.
- [25] Michel grabisc (2004), Une approche constructive de la décision multicritère, manuscrit Université Paris I.
- [26] Gregory A. Kiker & Todd S. Bridges & Arun Varghese & Thomas P. Seager, & Igor Linkov (2005), Application of Multi-criteria Decision Analysis in Environmental Decision Making , Integrated Environmental Assessment and Management - Volume 1, Number 2 - pp. 95-108.

## AUTHORS

**Aissa Taibi** is a Ph.D. student and assistant professor at Oran University and affiliated researcher in Oran computer Lab. He is engineer in computer systems from Oran University in 1987; he received the Magister degree in system modelization from the same university in 2003. His interest field is Multi-criteria decision analysis and geographic information systems, expert systems and fuzzy reasoning.

**Baghdad Atmani** received his Ph.D. degree in computer science from the University of Oran (Algeria), in 2007. He is full professor in computer science. His interest field is knowledge discovery from data and his research are directed towards the Machine Learning problems. His research is based on inductive machine learning approaches, knowledge systems based and case-based reasoning. His research are guided and evaluated through various applications in the field of control systems, scheduling, production, maintenance, simulation, data integration and spatial data-mining.

*INTENTIONAL BLANK*

# A COMPARATIVE OVERVIEW TO DTC FOR PMSM DRIVEN PURE ELECTRIC VEHICLES

Fatih Korkmaz

Department of Electric-Electronic Engineering, Çankırı Karatekin University,  
Uluyazı Kampüsü, Çankırı, Turkey

## ABSTRACT

*For many industrial applications that including electric motor, selecting of the appropriate motor type is keypoint for the application success and efficiency. This paper presents a comparative performance evaluation between direct torque controlled Brushless AC and Brushless DC motors which have been considered as traction motor on electric vehicles. Moreover, a short review to brushless AC and brushless DC motor technology and vector driver systems are also included. The main structure of the direct torque control method is presented and small-scale pure electric vehicle model is integrated to the motor control model. To make a fair comparison, the brushless AC and brushless DC motors modeled as having the same electrical parameters and simulated with same control model parameters. The simulation studies have shown that direct torque controlled brushless AC and brushless DC motors have similar performance as vehicle traction motor, and both motor types can be considered as a good candidate for electric vehicles applications.*

## KEYWORDS

*PMSM motors, BLAC motor control, BLDC motor control, Electric vehicles, Direct Torque Control*

## 1. INTRODUCTION

In 1950s, the availability of high energy density permanent magnets (PMs) and developments in power switching technologies led the development of PM motors[1]. Two typical classifications of permanent magnet synchronous motors (PMSMs) have been presented in literature by scholars according to shape of the induced emf (also known as back emf). One is called as brushless AC (BLAC) which has a sinusoidal-wave back-emf and the other is brushless DC (BLDC) motor which has a trapezoid-wave/square-wave back-emf. A BLDC motor can be defined as a electronic commutated self synchronous rotary motor, where the rotor contains permanent magnet with rotor position sensors. [2] But, it's obvious that same definition easily can be used for the BLAC motors. So far, there has not been a unified standard about the classification or definition of the BLDC motors or BLAC motors.

The main question that should be asked about the PMSMs is “Why are the PM motors happening more and more popular with every past day?” A PMSM incorporates starting characteristic of serious excitation DC motors and speed-torque characteristic of shunt excitation DC motor. Furthermore, the PMSMs have structural advantages of conventional brushless AC motors[2]. The PMSMs have great advantages like high efficiency, simple structure, high power densities, high starting torque, wide speed ranges, linear torque and speed characteristics, low maintenance and works in any condition. Due to absence of mechanical brushes and commutator, PMSMs can be acceptable as best choice for high performance drive applications[3]. Hence, it has been preferred motor type in many field such as robotics, vehicle propulsion, aerospace, industrial control, machine tools, etc.

This paper presents an overview to PMSMs and vector drive technologies. Moreover, a comparative performance investigation has been occurred between DTC controlled BLAC and BLDC motors which have been considered as traction motor on electric vehicles. The main structure of the DTC method is presented and small-scale pure electric vehicle model is integrated to the motor control model. The BLAC and the BLDC motors modeled with Simulink as having the same electrical parameters and simulated with same control model parameters to make a fair comparison.

## 2. BRUSHLESS SYNCHRONOUS MOTORS

The AC drive systems can be grouped under two major heading – induction drive systems and synchronous drive systems. Not far, about 3 decades ago, the induction drive systems almost monopolized the whole market, but with the advent of high magnetic permanent magnets, synchronous drive systems, include the PM brushless motors, are becoming popular. [4]

A BLDC motor is designed to be supplied with a trapezoidal shape current while a BLAC with sinusoidal as can be seen in Figure 1. [5]

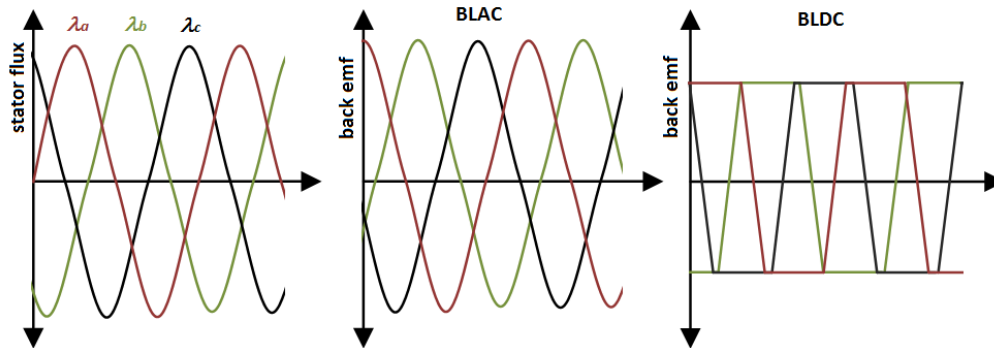


Figure 1. Back emf waveforms of BLAC and BLDC motors

The PM-BLAC motors fed by sinusoidal AC sources and produce an essentially constant torque, or so-called smooth torque. On the other hand, the PM-BLDC motors fed by rectangular or trapezoidal AC waves and have significant torque pulsations. However, due to interaction between a rectangular field and a rectangular current, the BLDC motors can produce higher torque than the BLAC motor [4].



Torque ripples issue is major problem for the many kind of systems, include EVs. The torque ripples causes torsional vibration and also acoustic noise. Due to vibration systems, mechanical parts of the system can be damage, even all system can be breakout [6]. For many years, reduction of the torque ripples in motor control systems have been major research subject and many kind of methods have been proposed by scholars and engineers to overcome this problem [7-9].

### 3. DIRECT TORQUE CONTROL

Direct torque control (DTC) is the name of reliable and energy efficiency vector control method. Although the DTC was originally proposed for asynchronous motors by Takahashi in 1986[10], over the years, it has been applied for various motor type such as linear motor, reluctance motor and also PM motors[11].

Another vector control method which has been preferred in high performance control of motors, is field oriented control (FOC). Almost half a century ago, the FOC method was also firstly proposed for asynchronous motors [12], but today, the FOC has been completely developed and it has been implemented for motor types such as PM motors. Conventional DTC method for AC motors with equations is given in Figure 2.

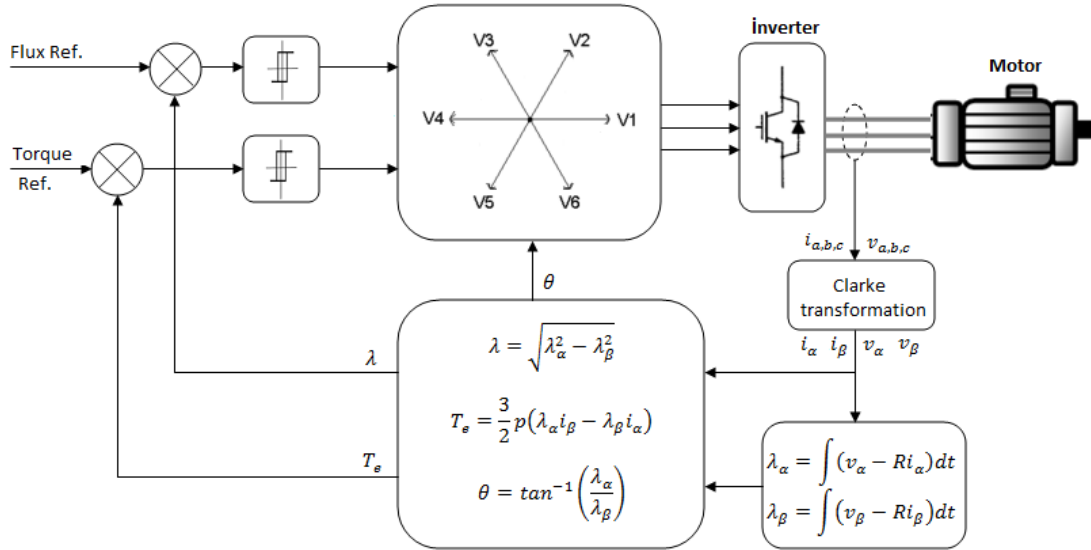


Figure 2. Conventional DTC motor scheme for AC motors

Over the last 40 years, many articles have been published about FOC and DTC of different motor types. However, few articles have focus on similarities or differences of between these methods. [13-15]

Some main structural advantages of the DTC can be listed as follow;

- Does not require current regulator
- Does not require PWM modulator

- Does not require rotor position measurement
- Does not require complex coordinate transformation
- Less parameter depended

To summarize, it can be pointed out that the DTC has simple structure when compared the FOC. On the other hand, the DTC method has some disadvantages and most important differences against to the FOC is high torque ripples[16].

In [13], the authors presented comparative analysis between two control methods for PMSMs. Advantages and disadvantages of the both methods were compared and discussed. Moreover, simulation based tests were performed for different working conditions. The authors claimed that the DTC had fast dynamic response while the FOC had better torque response and when selecting control method application requirements should be considered.

In another paper [17] the FOC and the DTC controlled PMSM was compared and discussed in detail. As a result, the authors claimed that there were no meaning differences in the main characteristics of the motor for both control method.

A comparisons between the FOC and the DTC controlled PMSMs had also been presented in [14] by numerical simulations. The paper supports almost same the claims with [13]. In other words, numerical simulations proved that the FOC was better on overall torque performance and the DTC had faster dynamic responses.

#### **4. MOTORS AND CONTROLS ON EVS**

Small-Scale Electric Vehicles (SSEV) are used in personal transportation for short distances. The SSEVs are powered by only battery packs and they only driven by electric motors-pure electric vehicle (PEV)- such as bicycles, scooters, skateboards, toy cars, rail cars, watercraft, forklifts, golf cars, and city cars, etc. The synchronous motors with PM-(BLAC/BLDC) are leading choice for vehicle propulsion in EVs due to advantages that mentioned earlier sections [18].

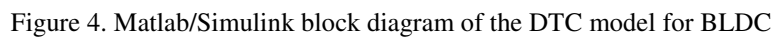
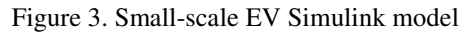
The DTC method is very useful and appropriate for EVs due to its great advantages such as; controlling torque directly, simple structure and easy applicable, low parameter sensitivity. [19]. To compare of the BLAC and BLDC traction motor performances on the DTC controlled system, Matlab/Simulink models of the controller has been created. Small-scale EV model has also included to model to obtain reference torque parameter for the DTC system. Figure 3 shows Small-scale EV Simulink model.

#### **5. SIMULATIONS AND RESULTS**

In simulations, BLAC and BLDC driven EV is accelerated to 0-2000 rpm motor speed, and the torque reference value is produced by EV model. Road angle is  $0^\circ$  for both motor. Matlab/Simulink block diagram of the DTC model for BLDC motor is given in Fig. 4.

The model and vehicle parameters which have been used in simulation as follows: Vehicle mass: 100kg, wheel radius: 0.4 m, Ratio of shaft: 1/10, BLAC/BLDC stator resistance: 0.4578 ohm,

The speed (in rad/s) and the torque responses of the motor are presented in Figure 5 – Figure 7.



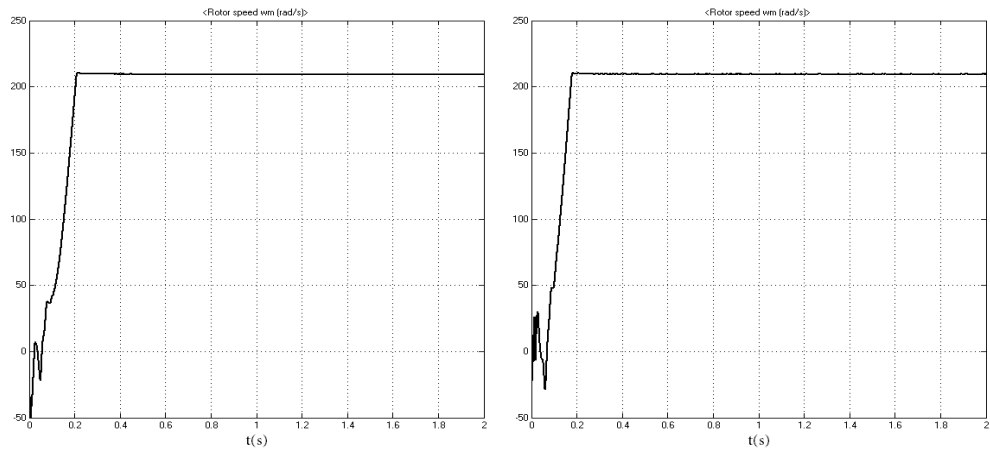


Figure 5. Overview of speed responses of BLAC (leftside) and BLDC (rightside) motors

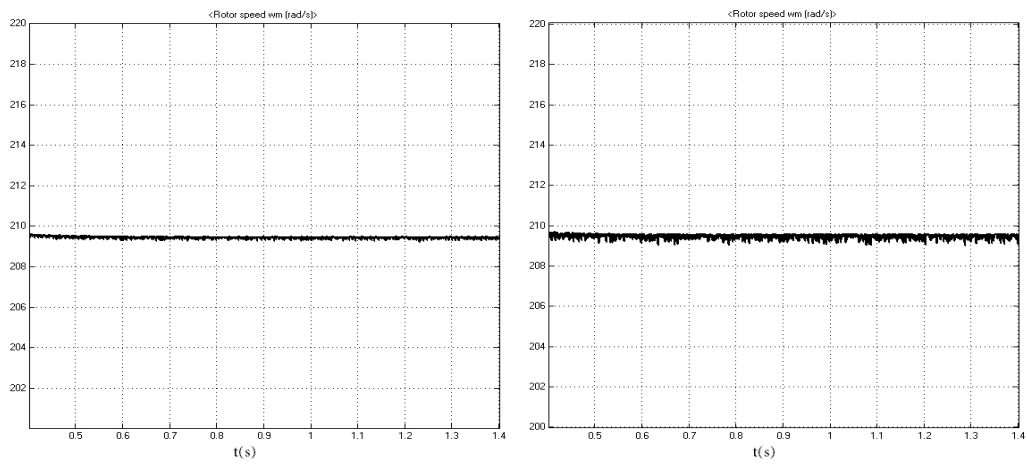


Figure 6. Zoomed view of speed responses of BLAC (leftside) and BLDC (rightside) motors

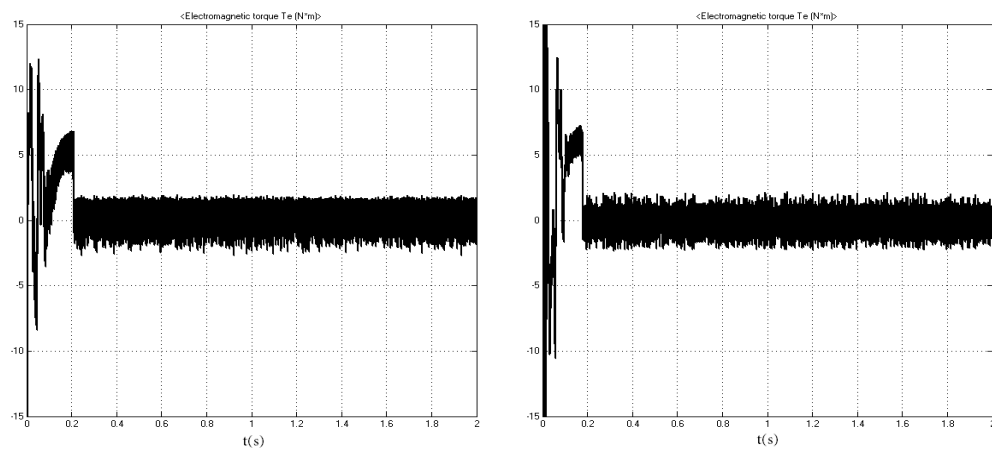


Figure 7. Overview of torque responses of BLAC (leftside) and BLDC (rightside) motors

When the speed and torque curves of the motor are compared, it can be clearly seen that motor performances are pretty close to each other for both motor types. However, the BLAC motor speed response is better than the BLDC motor due to less ripples (Figure 6.). The simulation studies proof that the DTC method is very appropriate for the BLAC and BLDC motors and vehicle designers/engineers can choose one of the two considering other physical or economical parameters.

## 6. CONCLUSIONS

Brushless synchronous motors are getting more and more popular every past day due to their well-known advantages like maintenance-free structures, height mass-power ratio, silent works, etc. BLAC and BLDC motors are indispensable motor types for electric vehicle traction systems. This paper aims to present a fair comparison between the BLAC and the BLDC motors for vehicle traction systems. For this purpose, direct torque controlled BLAC and BLDC motor modeled and simulated. The simulation studies show that both motors can be applicable for electric vehicle with DTC method. The speed and torque performances of the motors are almost same. So, it means, when brushless synchronous motor selection is necessary for electric vehicles or similar industrial system, some other physical or economical factors should be taken into consideration.

## REFERENCES

- [1] Krishnan, R. Permanent magnet synchronous and brushless DC motor drives, CRC Press, 2010
- [2] Chang-liang Xia, Permanent magnet brushless dc motor drives and controls, John Wiley & Sons Singapore Press, 2012
- [3] Purna Chandra Rao, Y. P. Obulesh and Ch. Sai Babu, "Mathematical Modeling of BLDC Motor With Closed Loop Speed Control Using PID Controller Under Various Loading Conditions" ARPN Journal of Engineering and Applied Sciences, vol. 7, no. 10, 2012.
- [4] K.T. Chau, Zheng Wang Chaos in electric drive systems : analysis, control and application, 2011
- [5] Torres, D., 2009. Comparing motor-control techniques. [Online] Available at: <http://www.ecnmag.com/articles/2009/10/comparing-motor-control-techniques> [Accessed 05 March 2016].
- [6] Z. Q. Zhu and J. H. Leong, "Analysis and Mitigation of Torsional Vibration of PM Brushless AC/DC Drives With Direct Torque Controller," in IEEE Transactions on Industry Applications, vol. 48, no. 4, pp. 1296-1306, July-Aug. 2012.
- [7] Fatih Korkmaz, Ismail Topaloglu, Hayati Mamur "Fuzzy logic based direct torque control of induction motor with space vector modulation", International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI), Vol.2, No. 5/6, December 2013
- [8] S. B. Ozturk, W. C. Alexander and H. A. Toliyat, "Direct torque control of four-switch brushless DC Motor with non-sinusoidal back-EMF," Power Electronics Specialists Conference, 2008. PESC 2008. IEEE, Rhodes, 2008, pp. 4730-4736.

- [9] M. Masmoudi, B. El Badi and A. Masmoudi, "Direct Torque Control of Brushless DC Motor Drives With Improved Reliability," in IEEE Transactions on Industry Applications, vol. 50, no. 6, pp. 3744-3753, Nov.-Dec. 2014.
- [10] Takahashi, I. & Noguchi, T. (1986) "A new quick-response and high efficiency control strategy of an induction motor," IEEE Transactions on Industrial Applications, vol.I A-22 ,No.5. , pp. 820–827.
- [11] Korkmaz, F., I. Topaloglu, and R. Gurbuz. "Simulink Model of Vector Controlled Linear Induction Motor with End Effect for Electromagnetic Launcher System." Elektronika Ielektrotehnika, 2014; 20(1)
- [12] Blaschke, F.,(1972) "The Principle of Field Orientation Applied to The New Transvector Closed-Loop Control System for Rotating Field Machines", Siemens-Rev., Vol. 39, 217–220.
- [13] X.T. Garcia, B. Zigmund, A. Terlizzi, R. Pavlanin, L. Salvatore "Comparison between FOC and DTC strategies for permanent magnet synchronous motors" Adv Electr Electron Eng, 5 (2012), pp. 76–81
- [14] Merzoug, M. S., and F. Naceri. "Comparison of field-oriented control and direct torque control for permanent magnet synchronous motor (pmsm)." Proceedings of world academy of science, engineering and technology. Vol. 35. 2008.
- [15] Korkmaz, F., Topaloglu, I., Cakir, M. F., & Gurbuz, R. (2013, May). Comparative performance evaluation of FOC and DTC controlled PMSM drives. In Power Engineering, Energy and Electrical Drives (POWERENG), 2013 Fourth International Conference on (pp. 705-708). IEEE.
- [16] Yılmaz Korkmaz, Fatih Korkmaz, İsmail Topaloğlu, Hayati Mamur"Adaptive bandwidth approach on DTC controlled induction motor"International Journal of Instrumentation and Control Systems,vol : 5 - no : 2, 2015
- [17] Navid Maleki, Mohammad Reza Alizadeh Pahlavani, Iman Soltani "A Detailed Comparison Between FOC and DTC Methods of a Permanent Magnet Synchronous Motor Drive"Journal of Electrical and Electronic Engineering,2015; 3(2-1): 92-100
- [18] James Goss, Mircea Popescu ""Malcolm Burwell – International Performance/cost comparison of induction-motor & permanent-magnet-motor in a hybrid electric car Copper Association,July 2013 – Tokyo Available at: <http://www.coppermotor.com/wp-content/uploads/2013/08/Techno-Frontier-2013-MBurwell-ICA-EV-Traction-Motor-Comparison-v1.8-Eng1.pdf> [Accessed 05 March 2016].
- [19] Devender Kumar, Ishan Thakur "A Review on Direct Torque Control for Induction Motor" International Journal of Science and Research (IJSR),Volume 3 Issue 7, July 2014

## AUTHORS

**Fatih Korkmaz** received the B.S., M.S., and Ph.D. degrees in electrical education from Gazi University, Ankara, Turkey, in 2000,2004 and 2011, respectively. His current major research interests include design, control and applications of motor drive systems. From 2013, he is an assistant professor in Çankırı Karatekin University, Turkey



# A NEW APPROACH TO DTC METHOD FOR BLDC MOTOR ADJUSTABLE SPEED DRIVES

Fatih Korkmaz

Department of Electric-Electronic Engineering, Çankırı Karatekin University,  
Uluyazı Kampüsü, Çankırı, Turkey

## ABSTRACT

*This paper proposes a new approach to direct torque control (DTC) method for brushless direct current (BLDC) motor drives. In conventional DTC method, two main reference parameters are used as: flux and torque. In this paper, the new approach has been proposed to improve the BLDC motor dynamic performance. A main difference from the conventional method of it was that only one reference parameter (speed) was used to control the BLDC motor and the second control parameter (flux) was obtained from speed error through the proposed control algorithm. Thus, the DTC performance has been especially improved on systems which need variable speed and torque during operation, like electric vehicles. The dynamic models of the BLDC and the DTC method have been created on Matlab/Simulink. The proposed method has been confirmed and verified by the dynamic simulations on different working conditions.*

## KEY WORDS

*Brushless machines, Direct torque control, Vector control, Torque control*

## 1. INTRODUCTION

In recent years, the BLDC motors applications become more popular in industrial and daily utilization, like CNC machines, servo systems and electric vehicles. The main reasons of the increasing on popularity of BLDC motors, in other words, main advantages of the BLDC motors, can be listed as follows[1-4]:

- High efficiencies
- High power densities
- High starting torque
- Wide speed ranges
- Linear torque and speed characteristics
- Low maintenance and works in any condition

The BLDC motors not only have advantages of conventional DC motors but also have advantages of AC motors as can be seen from the list. Because, they have the AC motor mechanical structures –no brushes or collectors–while they have the DC motors electrical structures.

The DTC method was developed by Takahashi in the middle of the 1980s, for three phase induction motors. The DTC method converts three phase parameters of the motor (three phase voltage and currents) to two phase independent vector components with Clarke transformation, thus, it proposes

control of every components separately, like conventional DC motors. Although, it was developed for the induction motors, it has been applied many other motor types like permanent magnet AC motors, BLDC motors, switched reluctance motors, linear motors[5-6].

In literature, many kind of BLDC motor drive methods can be found. In [7], artificial neural networks based method was used in modeling of BLDC to get the maximum power consumption. Very simple and effective three-level neutral point clamped inverter was proposed to drive axial flux BLDC motors, in [8]. Field Programmable Gate Array (FPGA) based BLDC motor driver with using digital pulse-width modulation (PWM) is presented in [9]. In addition, several different methods, which based of the DTC, were studied in BLDC drivers. Reference [10] proposes the DTC method for matrix converter fed BLDC motor. The DTC of BLDC motor method using with four-switch inverter in constant torque region was proposed in [11].

In this paper, a new approach to the DTC method has been proposed for small sized electric vehicles that work in variable speed and torque conditions, naturally. In the proposed method, optimum stator flux reference value was obtained by PI controller with usage of the speed error. The dynamic model of the proposed method was developed with Matlab/Simulink. The dynamic simulations were performed and results were presented to illustrate the validity of the proposed method.

## 2. DIRECT TORQUE CONTROL OF BLDC

The BLDC motor has three phase stator windings with permanent magnet rotor and electrical model of the motor that connected with PWM inverter, is given in Fig. 1[12].

Voltage equations of the motor can be obtained by the following equation;

$$\begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} = \begin{bmatrix} R & 0 & 0 \\ 0 & R & 0 \\ 0 & 0 & R \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} + \begin{bmatrix} L & 0 & 0 \\ 0 & L & 0 \\ 0 & 0 & L \end{bmatrix} \frac{d}{dt} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} + \begin{bmatrix} e_a \\ e_b \\ e_c \end{bmatrix} \quad (1)$$

Where  $V_a, V_b, V_c$  are phase voltages,  $R$  is resistance,  $L$  is phase inductance,  $i_a, i_b, i_c$  are phase currents and  $e_a, e_b, e_c$  are back EMFs.

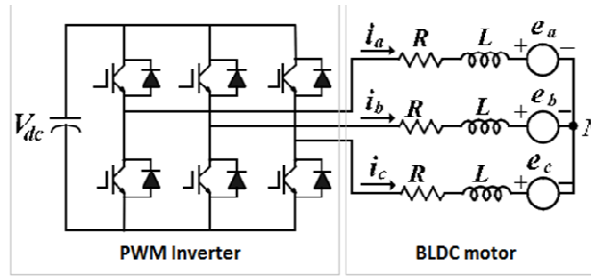


Figure 1. Electrical model of the motor that connected with PWM inverter

The mechanical moment equation of the motor given by the following equation;

$$T_e = T_L + B\omega_m + j \frac{d\omega_m}{dt} \quad (2)$$

$T_e$  and  $T_L$  describes generated electromagnetic torque and load torque, respectively.  $B$  is the friction



coefficient,  $j$  is the inertia and  $\omega_m$  is the angular velocity of rotor[13].

The DTC method needs to transformation of the three phase motor parameters to two phase. In this transformation, electrical parameters of the motor (voltages, currents, back emf) should to be transformed to stationary reference frame and it can also be named as  $\alpha - \beta$  transformation or Clarke transformation in many sources. The Clarke transformation matrix is given (3).

$$\begin{bmatrix} f_\alpha \\ f_\beta \end{bmatrix} = \frac{2}{3} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} f_a \\ f_b \\ f_c \end{bmatrix} \quad (3)$$

Where,  $f_\alpha, f_\beta$  are  $\alpha - \beta$  components of motor parameters, and  $f_a, f_b, f_c$  are the abc frame components [14].

With the transforming of the three phase parameters of the motor,  $\alpha - \beta$  components of the phase voltages can be written as:

$$v_{s\alpha} = R_s i_{s\alpha} + L_s \frac{di_{s\alpha}}{dt} + e_\alpha \quad (4)$$

$$v_{s\beta} = R_s i_{s\beta} + L_s \frac{di_{s\beta}}{dt} + e_\beta \quad (5)$$

Where  $v_{s\alpha}, v_{s\beta}$  are the stator voltages,  $i_{s\alpha}, i_{s\beta}$  are the stator currents and  $e_\alpha, e_\beta$  are back emf in the  $\alpha - \beta$  reference frame.

In the DTC scheme, stator flux components are obtained from  $\alpha - \beta$  components of the measured stator voltages and currents as given below[15]:

$$\lambda_{s\alpha} = \int (v_{s\alpha} - R_s i_{s\alpha}) dt \quad (6)$$

$$\lambda_{s\beta} = \int (v_{s\beta} - R_s i_{s\beta}) dt \quad (7)$$

The magnitude of the flux can be calculated with;

$$\lambda = \sqrt{\lambda_{s\alpha}^2 + \lambda_{s\beta}^2} \quad (8)$$

and position of the stator flux vector can be calculated with;

$$\theta = \arctan \frac{\lambda_{s\beta}}{\lambda_{s\alpha}} \quad (9)$$

In the conventional DTC (C-DTC) method, control algorithms works with two separate reference values as torque (or speed) and flux references. Because, in idea of the DTC, stator flux vector has two components ( $\alpha - \beta$  components) and they can be controlled independently from each other. One of them controls flux, while the other one controls torque. In generally, flux reference is kept constant and the speed control of the motor can be achieved by setting up the torque reference value. This approach is very appropriate for constant torque-variable speed applications. The conventional DTC block diagram is given in Fig. 2.



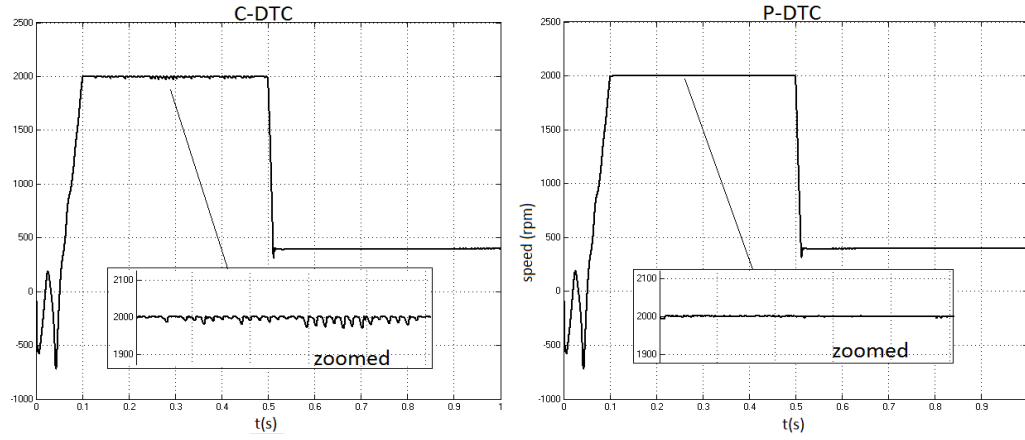


Figure 4. Speed responses of the BLDC in variable speed-constant load

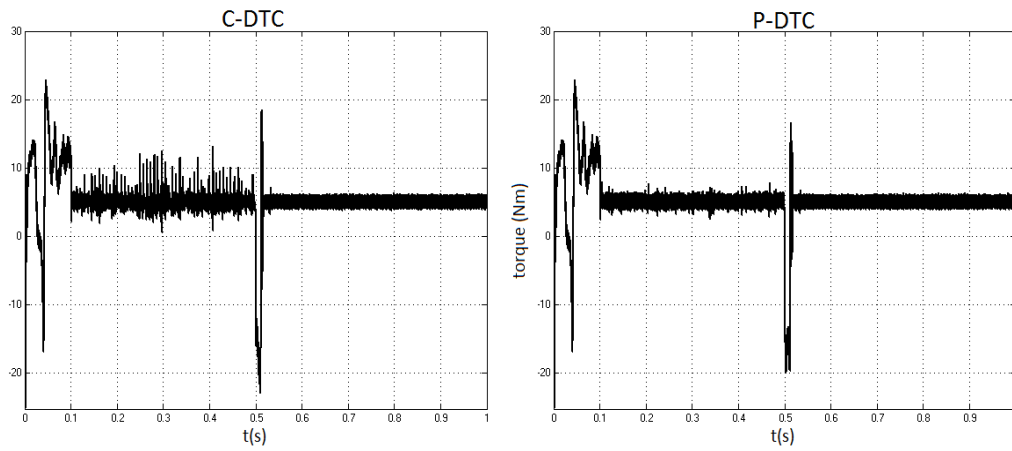


Figure 5. Torque responses of the BLDC in variable speed-constant load

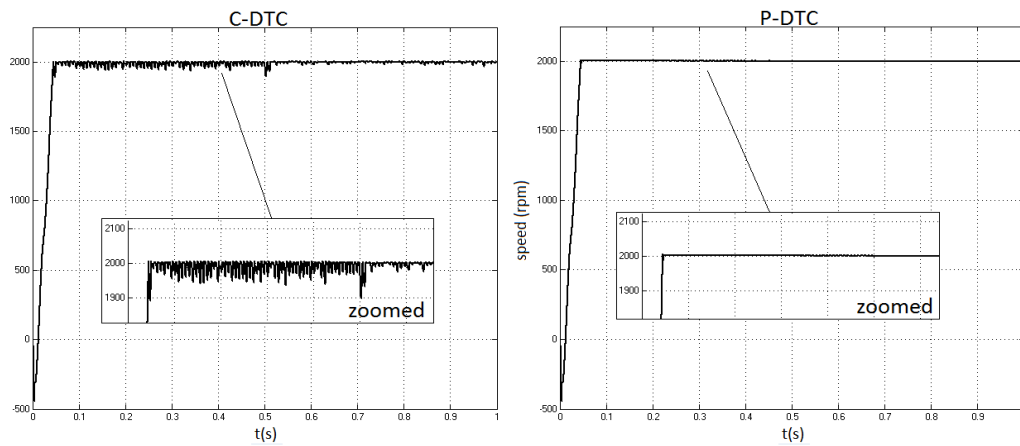


Figure 6. Speed responses of the BLDC in variable torque-constant speed

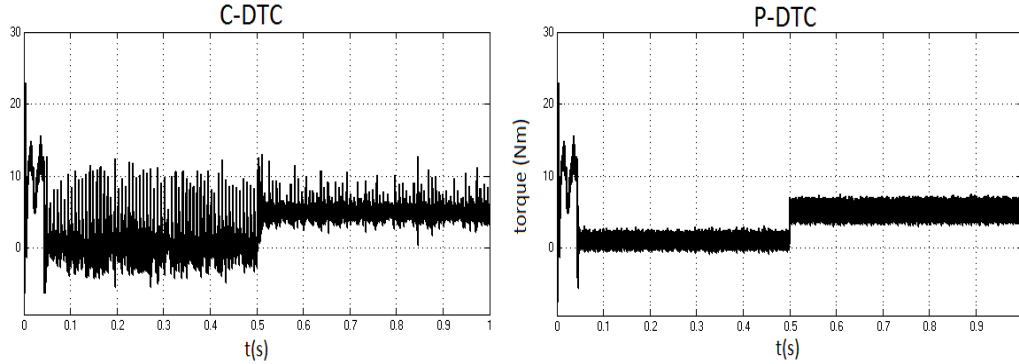


Figure 7. Torque responses of the BLDC in variable torque-constant speed

In the second test, the motor was simulated at constant speed reference (2000 rpm) and the torque reference was changed to 1 Nm to 5 Nm at 0.5. sec. The speed and the torque responses of the motor were given in Fig. 6 and Fig. 7, respectively.

It can be seen in from figures (Fig. 4–Fig. 7) that, the BLDC motor speed and torque responses were improved with the P-DTC method. The torque and speed ripples reduced remarkably with the P-DTC method especially in high speed reference values. In 400 rpm speed reference, the motor dynamic behaviors were almost same with both methods.

#### 4. CONCLUSIONS

Over the last decades, the BLDC motors have been used many industrial applications and it has gain great popularity between electric motors. On the other hand, the DTC is well-known high performance control method not only for induction motors but also many other motor types. This paper presents a new perspective to the conventional DTC method. The proposed DTC method applied to the BLDC motors which considered drives small size electric vehicles. These vehicles works in variable speed and torque conditions with single reference input: speed. The conventional DTC method was modified with the adding of second PI controller that produces optimum flux reference. In order to test the validity and applicability of the proposed method, dynamic simulations were been performed under different working conditions and results were presented. The simulations results showed that the proposed method reduced remarkably the speed and the torque ripples when compared conventional DTC method. The proposed method had also very simple structure to apply the conventional DTC and its extra computational load to the controller was almost zero.

#### APPENDIX

The parameters of the blocks and the motor that used in dynamic simulations are, DC bus voltage 400V, stator phase resistance,  $R_s$  (ohm) = 0.45, stator phase inductance  $L_s$  (H) =  $8.5e-3$ , flux hysteresis band limits =  $\pm 0.005$ , torque hysteresis band limits =  $\pm 0.02$ , stator flux reference = 0.32 weber, sampling time 10  $\mu$ sec.

#### REFERENCES

- [1] C. Xia, Z. Li, and T. Shi, “A Control Strategy for Four-Switch Three-Phase Brushless DC Motor Using Single Current Sensor”, *Industrial Electronics, IEEE Transactions on*, vol.56, no.6, pp. 2058–2066, June 2009.

- [2] Z. Xiaofeng, L. Zhengyu, "A new BLDC motor drives method based on BUCK converter for torque ripple reduction," Power Electronics and Motion Control Conference, 2006. IPEMC 2006. CES/IEEE 5th International , vol.2, no., pp.1–4, 14-16 Aug. 2006
- [3] I. Topaloglu, F. Korkmaz, H. Mamur, R. Gurbuz, "Closed-Loop speed control of PM-BLDC motor fed by six step inverter and effects of inertia changes for desktop CNC machine" Elektronika IR Elektrotehnika, vol. 19, no 1, pp. 7–10, 2013.
- [4] I. Tarimer, A. Akpunar, R. Gurbuz, "Design of a direct sliding gearless electrical motor for an ergonomic electrical wheelchair", Elektronika IR Elektrotehnika, no 3, pp. 75–80, 2008.
- [5] I. Takahashi and T. Noguchi , "A new quick-response and high efficiency control strategy of an induction motor" IEEE Transactions on Industrial Applications, vol.I A-22 , no.5, pp. 820–827, 1986.
- [6] F. Korkmaz, I. Topaloglu, R. Gurbuz, "Simulink model of vector controlled linear induction motor with end effect for electromagnetic launcher system", Elektronika IR Elektrotehnika, vol. 20, no 1, pp. 29–32, 2014.
- [7] M. Nizam, A. Mujiyanto, H. Triwaloyo, Inayati, "Modelling on BLDC motor performance using artificial neural network (ANN)", Rural Information & Communication Technology and Electric-Vehicle Technology (rICT & ICeV-T), 2013 Joint International Conference on , vol., no., pp.1–4, 26-28 Nov. 2013.
- [8] S. De, M. Rajne, S. Poosapati, C. Patel, K. Gopakumar, "Low-inductance axial flux BLDC motor drive for more electric aircraft", Power Electronics, IET , vol.5, no.1, pp.124–133, January 2012.
- [9] A. Tashakori, M. Hassanudeen, M. Ektesabi, "FPGA based controller drive of BLDC motor using digital PWM technique", Power Electronics and Drive Systems (PEDS), 2015 IEEE 11th International Conference on , vol., no., pp.658–662, 9-12 June 2015.
- [10] R. Muthu, M.S. Kumaran, L.A. Rajaraman, P. Ganesh, P. Reddy, "Direct Torque Control of matrix converter fed BLDC motor", Power Electronics (IICPE), 2014 IEEE 6th India International Conference on , vol., no., pp.1–6, 8-10 Dec. 2014.
- [11] S.B. Ozturk, W.C. Alexander, H.A. Toliyat, "Direct torque control of four-switch brushless DC Motor with non-sinusoidal back-EMF", Power Electronics Specialists Conference, 2008. PESC 2008. IEEE , vol., no., pp.4730–4736, 15-19 June 2008.
- [12] W.S. Im, W. Liu, J.M. Kim, "Sensorless control of 3-phase BLDC motors using DC current model", Energy Conversion Congress and Exposition (ECCE), 2014 IEEE , vol., no., pp.4484–4490, 14-18 Sept. 2014.
- [13] P.K. Girija, A. Prince, "Robustness evaluation of SMO in sensorless control of BLDC motor under DTC scheme", Power Signals Control and Computations (EPSCICON), 2014 International Conference on , vol., no., pp.1-6, 6-11 Jan. 2014.
- [14] F. Korkmaz, I. Topaloglu, M.F. Cakir, R. Gurbuz, "Comparative performance evaluation of FOC and DTC controlled PMSM drives", Power Engineering, Energy and Electrical Drives (POWERENG), 2013 Fourth International Conference on , vol., no., pp.705–708, 13-17 May 2013.
- [15] F. Korkmaz, M.F. Cakir, Y. Korkmaz, I. Topaloglu, "Fuzzy based stator flux optimizer design for direct torque control", International Journal of Instrumentation and Control Systems (IJICS), vol.2, no.4, pp. 41–49, October 2012.

**AUTHORS**

**Fatih Korkmaz** received the B.S., M.S., and Ph.D. degrees in electrical education from Gazi University, Ankara, Turkey, in 2000, 2004 and 2011, respectively. His current major research interests include design, control and applications of motor drive systems. From 2013, he is an assistant professor in Çankırı Karatekin University, Turkey.



# DETECTION OF AUTOMATIC THE VOT VALUE FOR VOICED STOP SOUNDS IN MODERN STANDARD ARABIC (MSA)

Sulaiman S. AlDahri

King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia  
saldahri@kacst.edu.sa

## ABSTRACT

*Signal processing in current days is under studying. One of these studies focuses on speech processing. Speech signal have many important features. One of them is Voice Onset Time (VOT). This feature only appears in stop sounds. The human auditory system can utilize the VOT to differentiate between voiced and unvoiced stops like /p/ and /b/ in the English language. By VOT feature we can classify and detect languages and dialects. The main reason behind choosing this subject is that the researches in analyzing Arabic language in this field are not enough and automatic detection of VOT value in Modern Standard Arabic (MSA) is a new idea. In this paper, we will focus on designing an algorithm that will be used to detect the VOT value in MSA language automatically depending on the power signal. We apply this algorithm only on the voiced stop sounds /b/, /d/ and /d<sup>2</sup>/, and compare that VOT values automatically generated by the algorithm with the manual values calculated by reading the spectrogram. We created the corpus, and used CV-CV-CV format for each word, the target stop consonant is in the middle of word. The algorithm resulted in a high accuracy, and the error rate was 0.80%, 26.62% and 11.71% for the three stop voiced sounds /d/, /d<sup>2</sup>/ and /b/ respectively . The standard deviation was low in /d/ sound because it is easy to pronounce, and high in /d<sup>2</sup>/ sound because it is unique and difficult to pronounce.*

## KEYWORDS

Arabic, VOT, MSA, POA, TEO

## 1. INTRODUCTION ABOUT VOICE ONSET TIME

VOT (Voice Onset Time) feature appears important in distinguishing between voiced and unvoiced stops in various languages. Phonation onset or VOT is defined as the interval (period) between the release burst of the stop and the onset of glottal vibration [1]. This period is measured in msec. This makes the VOT detection automatically is difficult. VOT as we have just described is relevant only for stop consonants [2]. Stop consonants are produced with a closure of the vocal tract at a specific place which is known as the place of articulation (POA) [3]. So, VOT is affected by the stop consonant's POA, which is different from one language to another. Also, VOT is affected by the speaker linguistic knowledge and vowel duration [3][4].

VOT value can be divided to three types, zero VOT, positive VOT and negative VOT. Zero VOT means where the onset of vocal fold vibration coincides very close to the release of the stop closure. Positive VOT means that there is a delay in the onset of vocal-fold vibration after the release of the stop closure; in this case, the voicing lag. In Negative VOT, the onset of vocal fold vibration precedes the release of the stop closure; in this case, the voicing lead [1][5]. Categories range from two-four depending on the certain language. For example, English and Spanish have two voicing categories, whereas Eastern Armenian and Thai have three voicing categories [6].

The VOT is an important characteristic of stop consonants that play a great role in perceptual discrimination of phonemes of the same POA [7]. Also, VOT is an important feature in stress related phenomena, word segmentation, and dialectal and accented variations in speech patterns [2]. Moreover, the previous researches, found that VOT values are not affected by the change of the age in both male and female [8]. It is well known that VOT varies to some extent with place of articulation. The principal findings are that: (1) the further back the closure, the longer the VOT (Fischer Jorgensen, 1954; Peterson & Lehiste, 1960); (2) the more extended the contact area, the longer the VOT (Stevens, Keyser & Kawasaki, 1986); and (3) the faster the movement of the articulator, the shorter the VOT (Hardcastle, 1973). These patterns have been known for many years [3].

VOT values are generally unobserved in fixed length frame based speech investigation. On the other hand, it is known that automatic speech recognition and performances can be understood by the help of VOT. Among the various applications of the use of VOT is the difficulty of accent detection. Non-native language can affect both the length and the quality of the VOT of English stops [9]. Depending on a research effort [9], VOT values can be used to discriminate Mandarin, Turkish, German, Spanish and English accents.

In languages which process two categories of voicing. Voiced and voiceless, the voicing onset usually starts before the release of the stop closure for a voiced stop. However, languages differ in the manifestation of these two categories. Depend on VOT, Lisker et al. [1] have a distinct number of voicing categories that are used contrastively, thereby forming different phonological entities of the respective system. Such categories range from two-to-four depending on the particular language. For example, English, Spanish, Tamil, Hungarian, and Dutch have two voicing categories, whereas Eastern Armenian and Thai have three voicing categories [1][6].

## **2. LITERATURE REVIEW**

### **2.1. Automatic detection the VOT**

There have been a number of previous researches proposing algorithms for automatic VOT measurement. Previous studies have used automatic measurements for speech recognition tasks (Niyogi and Ramesh, 1998, 2003; Ali, 1999; Stouten and van Hamme, 2009), phonetic measurement (Fowler et al., 2008; Tauberer, 2010), and accented speech detection (Kazemzadeh et al., 2006; Hansen et al., 2010). Some studies, focus on the problem of VOT measurement itself, and evaluate the proposed algorithm by comparing automatic and manual measurements (Stouten and van Hamme, 2009; Yao, 2009; Hansen et al., 2010; Lin and Wang, 2011). Moreover, one study used machine learning technique to design an algorithm to automatically measure VOT (Morgan and Joseph, 2012) [4].



On the other side, English was studied by Lisker et al. [1] using American and Britain dialects among more than nine other languages and dialects under different environments. In general, languages like English, Japanese, and German, were investigated for more than forty years for VOT stops [1]. Several studies that have been conducted in English, showed similar results to those of Lisker et al [1]. Peterson et al. [10] present their VOT results of /p/, /t/, /k/ as 58 msec, 69 msec and 75 msec respectively. Flege et al. [11] found the VOT of /p/ is 46 msec, /t/ is 56 msec and /k/ is 67 msec. In addition, close relations were found between the 18 languages investigated by Cho T. and Ladefoged P. [3] which represented 12 different language families.

In Spanish and French languages that treat long lead or pre-voiced stops (i.e., long, negative VOTs) as voiced, while short-lag stops are classified as voiceless (Caramazza & Yeni-Komshian, 1974; Deuchar & Clark, 1996; Zampini & Green, 2001). In other words, in English (and Swedish) voiced stops are characterized by an articulator timing similar to Spanish voiceless stops [12]. In French language, the VOT is sufficient phonological cue for the distinction of the homorganic stop consonant pairs in French [13].

However, to replace manual measurement, we believe that an automatic VOT measurement algorithm should meet three criteria. Both the burst and voicing onsets are often highly transient, and because the effects of interest (e.g., VOT difference between two conditions) in studies using VOT measurements are often very small, the algorithm should have high accuracy by the chosen measure of performance. The cues to the burst and voicing onset locations vary depending on many factors like speaking style, speaker's native language, and different labs have slightly different VOT measurement criteria [4].

In addition, there was a general age effect on the L2 learners' categorical-perception behavior mirrored by negative correlations and also the overall differences between listener groups were significant for all three voicing continua [12].

Das and et al. was Automatic detected of VOT for unvoiced stops (/t/, /k/ and /p/) used the Teager Energy Operator (TEO). This algorithm is applied to accent classification using English, Chinese, and Indian accented speakers. Using the 546 tokens and consisting of 3 words from 12 speakers. The VOT is detected with less than 10% error when compared to the manual detected VOT. Also, pairwise English accent classification are 87% for Chinese accent, 80% for English accent, and 47% for Indian accent [9]. The TEO is a nonlinear energy tracking signal operator which has been used in speech and signal processing. It has been shown that TEO can be useful for detecting voiced/unvoiced speech, speech under stress, speech under vocal fold pathology and an automated sub-band frequency analysis is performed to detect VOT value [9].

In another effort, Okalidou A. et al. found that the developmental patterns Standard-Greek (SG) and Cypriot-Greek (CG) were different due to the number of contrasting voicing categories in each language/dialect. They found the VOT value three-way voicing contrast in CG is acquired later than the two-way voicing contrast in SG [6].

In several methods of VOT detection are considered, with the most accurate method based on tracking the laryngographic signal. This method is not possible unless a laryngograph is used while recording speech from the speaker. The other methods are based on tracking formant frequencies (F1, F2 or F3), performing spectrographic analysis, or tracking the onset of speech (f0) periodicity in the acoustic waveform. Manual involvement is required in all of these methods

to calculate the VOT value [9]. In study Parkash C. et al., they found approach for detection of VOT based on Bessel expansion and amplitude modulation component of the TEO [14].

## 2.2. VOT in Arabic language

The researches and resources in the speech processing field for Arabic language is not enough, especially in the subject related to VOT. Studies based on Arabic language varies in results, Al-Ani [15] and Mitleb [16] suggested that Arabic is a member of Group A, but Yeni-Komshian et al. [17], who based his research on Lisker et al. [1], found it to be from Group B. On the other hand, Flege [11] considered it neither belongs to Group A nor B.

In a study by Alghamdi [2], he analyzed the role of VOT in speaker identification and the effect of acquiring a second language in Ghamdi dialect for Saudi speaker. This research showed that a phonetic diversity between the first and second language is maximized when the speakers are more fluent in the second language. Moreover, Alghamdi [2] investigated the Saudi dialect of Arabic language, and the results of the average VOT for /t/, /k/ and /tʔ/ were 39 msec, 42 msec and 21 msec, respectively.

Another effort, Mitleb [16] analyzed VOT of Jordanian Arabic stops. One of his results is that the VOT value depends on the vowel length in long vowel environment more than its dependency on the length in short vowel environment. Also, he realized that VOT distinguishes Arabic unvoiced and voiced stops as the case in English.

In Aldahri research, he concluded that VOT values of these stops are positive regardless of the voicing where /d/ is a voiced sound, but /t/ is not. This is not the case for the same sounds in English language, where voiced stops have negative VOT values, but it is positive for unvoiced (e.g., /t/) [19]. Another research, Aldahri and Alotiabi, they found the emphaticness property decreases VOT values if compared with VOTs of nonemphatic [20]. In another effort, MSA Arabic language is found to have both long and short VOT for unvoiced and voiced sounds respectively [18]. However, the researches to detect the VOT in Arabic language are not applied until now. So, that appears the important this research in this field.

The aim of this study is to design an algorithm to automatically detect the VOT for voiced stop sounds in MSA Arabic. The rest of this paper is organized as following: Section 3 describes the used corpus and the methodology. Section 4 gives the results of the research in addition to some discussions. Before the final section, Section 5 summaries the results of the research. Finally, Section 6 is to list our references.

## 3. EXPERIMENTAL FRAMEWORKS

The set of stop phonemes in MSA language consists of eight phonemes that can be classified into: emphatic and non-emphatic or voiced and unvoiced [21]. These sets are shown in Table 1 with a description of their place of articulation, voicing, and emphaticness property [19].

Table 1. Stop phonemes in MSA Arabic language.

			Bilabial	Alveo-dental	Velar	Uvular	Glottal
Stop	Voiced	Emphatic		/d <sup>ʔ</sup> /			
		Non-emphatic	/b/	/d/			
	Unvoiced	Emphatic		/t <sup>ʔ</sup> /			
		Non-emphatic		/t/	/k/	/q/	/ʔ/

### 3.1. Data set

The used corpus is based on the previous work explained in [18][19], with each word containing one of the seven targeted sounds. The recorder speakers were chosen carefully to insure achieving the utterance quality required for the work of this search.

Those speakers include native and non-native Arabic speakers. The speakers are between thirteen and forty years old. The words are chosen to make sure that the targeted sounds are in the middle of the word while the preceding and the succeeding phonemes with respect to the targeted phonemes are always the same (/a/). The word structure is CV-CV-CV. The speaker repeats this set of words for 5 trials.

The total number of the recorded utterances is 2800 (80 speaker's × 7 words × 5 trials) recorded words. As we know, the voiced stop sounds in MSA Arabic are /d/, /b/ and /d<sup>ʔ</sup>/. We managed to record three words (each one containing one of the voiced stop sounds) for 1200 recorded words (80 speaker's × 3 words × 5 trials). The sampling rate was set at 16000 sample/seconds (16 kHz) and resolution at 16 bit using one channel (mono).

Each record file is named according to the following naming pattern: SxxCyEzTw.wav. In this string S, C, E and T stand for speaker, consonant, emphatic, and trial, respectively. The xx (two digits number) displays the speaker number. The one digit y, 1 refers to /d/ or /d<sup>ʔ</sup>/, 3 refers to /t/ or /t<sup>ʔ</sup>/, 5 refers to /k/, 6 refers to /q/ and finally 7 refers to /b/. The fourth digit z, is a binary flag set to 0 for non emphatic and 1 for emphatic. One digit z is the emphatic/non-emphatic sound identifier as following: 1 refers to the pair /d<sup>ʔ</sup>/ or /d/, 3 refers to the pair /t<sup>ʔ</sup>/ and /t/. The last digit, w, is a one-digit number representing the trial number.

For the goal of this paper, we used carrier words that only carry the stop voiced consonants /d/, /d<sup>ʔ</sup>/ and /b/. The Table 2 is lists the part of the corpus that is targeted in this research work.

Table 2. Stop sounds and carrier words which used in this research

Arabic Alphabet Carrier	IPA Symbol	Carrier words (CV-CV-CV)	Transcription	Code
daal د	/d/	نَدَرَ	/nadara/	C1e0
Dhaad ض	/d <sup>ʔ</sup> /	نَضَرَ	/nad <sup>ʔ</sup> ara/	C1e1
Baa ب	/b/	نَبَرَ	/nabara/	C7e0

### 3.2. VOT Detection Algorithm

The methodology used to extract the VOT value is written using Java. The Power values of the signal file are used as input to the program in form of an array, each line or each value represents 1 msec in time. Then, the array is scanned line by line to create queues of increasing values. Each queue ends when the next value in the array is less than the current value (each queue represents a part of the signal from a local min to the next local max). However, In order to handle noise, the following condition is performed:

The queue contains a series of increasing values, but to ignore slight decreases that happened due to signal noise, we check if the decrease from the current power value to the next power value is less than or equal to a given value (0.1 dB by default), ignore it and continue building the queue, but if it's larger than that, stop and consider the current value the end of this queue.

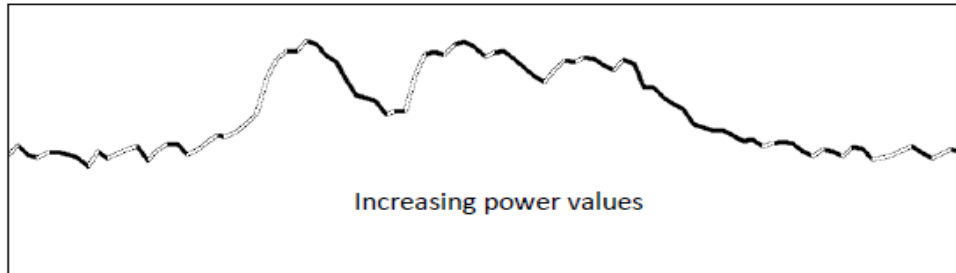


Figure 1. All created queues after scanning

To remove useless queues (that can't be a VOT), each queue is then checked for the following:

1. Power Delta: which is the difference between the highest and the lowest power values in the queue (the start and end). If it is less than a given value then ignore the queue. The value depending on our experience is between 2 to 4 dB.
2. Signal Length (Queue Length): If the length in msec is less than a given value then ignore the queue. Depending on our experience the VOT value is always more than 5 msec in MSA language.

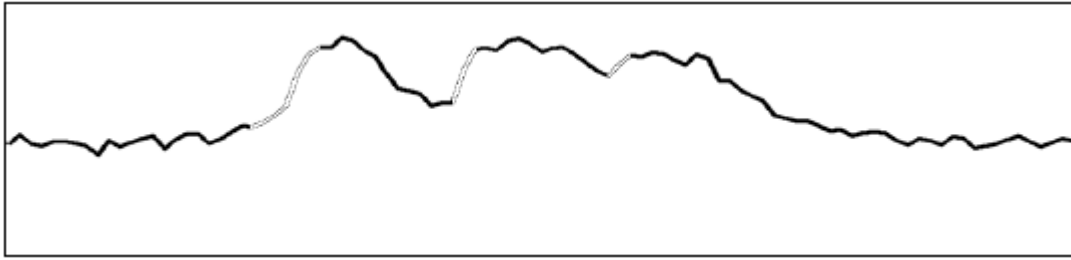


Figure 2. Remaining queues after the cleaning step

We're always looking for the VOT in the second letter (the middle letter) of three letter words. The algorithm benefits from the Pitch Contour data (if available) to skip the parts of the signal with a Pitch Contour value of 0 (unvoiced sound, which is usually the beginning and end of the file) as show in Figure 3. It also has an option to only read a window of a given length (usually between 200 to 400 msec) that represents the center of the signal, and ignores the rest of the file. The equation to calculate the center is:

$$X1 = \frac{\text{Pitch counter}}{2} - \text{Given length}$$

$$X2 = \frac{\text{Pitch counter}}{2} + \text{Given length}$$

The VOT value should be between X1 and X2 as show in Figure 3.

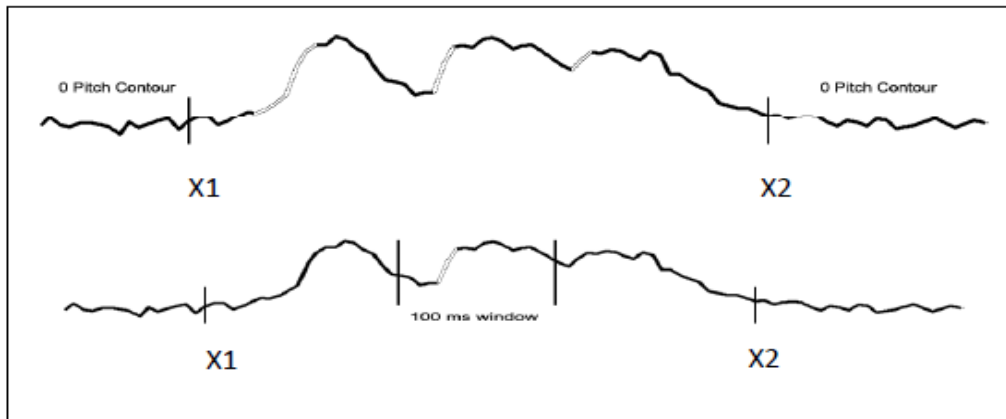


Figure 3. Remaining VOT candidates

#### 4. DISCUSSION AND RESULTS

We apply and investigation the algorithm on voiced stop sounds in MSA Arabic which are /d/, /dʒ/ and /b/. We examined many audio files in our corpus, but the tables lists partial subset of VOT values, specifically sixteen audio files for each of the three investigated stops. Twenty audio files are for each sound as shown in Tables 1, 2 and 3. VOT values were obtained by measuring the distance between the onset of energy in the formant frequency range representing the release of air pressure and the first vertical striations representing glottal pulsation from wide-

band spectrograms of recorded words [13]. Our algorithm depends on the power signal to determine VOT value. So, the start release of vocal tract is represented by the point where the signal start increasing and the start vibration vocal cord is represented by the point where the signal stops increasing as shown in Figure 3.

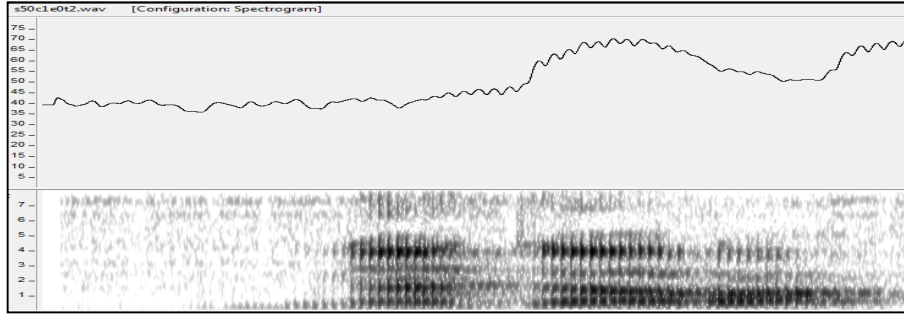


Figure 4. Spectrogram and power signal for stop voiced sound. Down arrow represent start release and up arrow represent start vibration

In any language, stop sounds are divided to voiced sounds and voiceless sounds. Lisker and et al.[1] divided languages based on VOT value. From our previous researches [18] we found that the VOT value in Arabic language is positive. Also, we found that the VOT value in voiced sounds is short VOT, and long VOT in unvoiced sounds.

Each Table contains three spectrogram readings for VOT values, the average for it, VOT value detected by our algorithm, standard division between algorithm VOT value and the average VOT value by reading the spectrogram, and the error rate. The error rate is calculated by the following equation:

$$\text{Error Rate} = \frac{|x-y|}{x} * 100$$

where  $x$  represents VOT value by reading the spectrogram and  $y$  represents VOT value by the algorithm.

Table 3. Algorithm statistics for /d/ sound.

Name File	By Read Spectrogram 1	By Read Spectrogram 2	By Read Spectrogram 3	Mean	By algorithm	Standard Deviation	Error Rate
s50c1e0t2	16	12	12	13.33333333	15	1.178511302	12.5
s51c1e0t2	13	11	10	11.33333333	7	3.064129385	38.24
s52c1e0t2	19	12	13	14.66666667	14	0.471404521	4.55
s53c1e0t2	12	11	12	11.66666667	12	0.23570226	2.86
s54c1e0t1	18	15	18	17	19	1.414213562	11.76
s55c1e0t2	9	9	10	9.333333333	10	0.471404521	7.14
s56c1e0t2	13	8	14	11.66666667	11	0.471404521	5.71
s59c1e0t4	10	10	12	10.66666667	14	2.357022604	31.25
s60c1e0t2	11	8	9	9.333333333	7	1.649915823	25.00
s61c1e0t2	19	18	11	16	15	0.707106781	6.25
Average	14	11.4	12.1	12.5	12.4	0.070710678	0.80

In Table 3, we analyze the algorithm for sound /d/. The average VOT value by the algorithm is 12.4 msec which are inside the VOT value range of three spectrograms readings. Eight files have VOT value inside the range of the three spectrogram readings and two files were outside this range. The pronunciation of this sound is easy, because it is found in most languages. So, the error rate is 0.80%. Also, the standard division for VOT value between the algorithm and manual spectrogram reading is 0.07 msec.

Table 4. Algorithm statistics for /d?/ sound.

Name File	By Read Spectrogram 1	By Read Spectrogram 2	By Read Spectrogram 3	Mean	By algorithm	Standard Deviation	Error Rate
s51c1e1t2	11	11	12	11.3333333	15	2.592724864	32.35
s53c1e1t2	10	8	13	10.3333333	18	5.421151989	74.19
s54c1e1t1	15	15	10	13.3333333	19	4.006938427	42.50
s55c1e1t3	12	8	13	11	15	2.828427125	36.36
s57c1e1t2	7	8	8	7.66666667	8	0.23570226	4.35
s58c1e1t2	10	8	12	10	10	0	0.00
s59c1e1t2	12	12	6	10	14	2.828427125	40.00
s60c1e1t3	11	10	9	10	10	0	0.00
s61c1e1t2	7	11	8	8.66666667	6	1.885618083	30.77
s62c1e1t2	10	11	10	10.3333333	15	3.299831646	45.16
Average	10.5	10.2	10.1	10.26666667	13	2.309882152	26.62

The outcomes of VOT value by algorithm and spectrogram readings for /d?/ sound is represented in Table 4, which contains 10 speakers. This sound is unique to MSA Arabic language and difficult to pronunciation. So, it is difficult to measure VOT value by reading the spectrogram hence it is difficult to detect it automatically. By algorithm, the VOT values in 6 files are outside the range of three spectrogram readings. Also, the maximum VOT value outside the spectrogram range by 5 msec and the minimum is outside by 2 msec. In addition, the average VOT value detected by algorithm is 13 msec, which is outside the three spectrogram readings. The error rate is 26.62% for this sound.

Table5. Algorithm statistics for /b/ sounds.

Name File	By Read Spectrogram 1	By Read Spectrogram 2	By Read Spectrogram 3	Mean	By algorithm	Standard Deviation	Error Rate
s50c7e0t2	8	7	9	8	9	0.7071068	12.50
s51c7e0t2	13	13	12	12.6666667	15	1.6499158	18.42
s52c7e0t2	10	9	8	9	7	1.4142136	22.22
s53c7e0t2	10	13	9	10.6666667	11	0.2357023	3.13
s54c7e0t3	11	11	12	11.3333333	7	3.0641294	38.24
s55c7e0t2	10	12	13	11.6666667	14	1.6499158	20.00
s58c7e0t2	8	8	7	7.66666667	5	1.8856181	34.78
s59c7e0t2	12	14	8	11.3333333	5	4.4783429	55.88
s60c7e0t2	8	12	7	9	6	2.1213203	33.33
s61c7e0t2	15	15	12	14	14	0	0.00
Average	10.5	11.4	9.7	10.53333333	9.3	1.7206265	11.71

In Table 5, we investigate and analyze 10 speakers for /b/ sounds. We found by algorithm four files with a VOT value far from the spectrogram range by 3 msec or less, and another two files

far by 5 msec. We ignore the different between the algorithm and spectrogram reading when it equals 1 msec. The accuracy or error rate for this sound is 11.71%. In addition, the average VOT value by algorithm is inside the range of spectrogram readings for this sound.

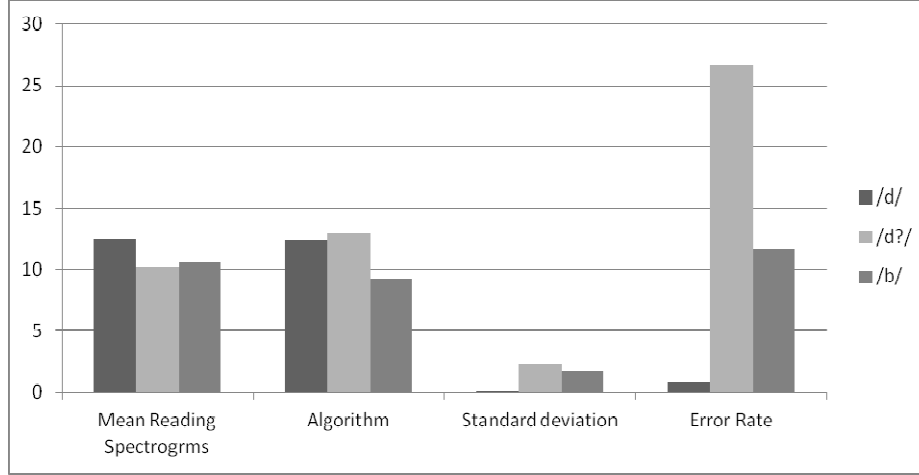


Figure 5. Statistic of VOT value for three voiced stop sounds.

In another side, Figure 5 shows the statistic to compare the VOT value between the automatic VOT estimated by the algorithm and the manual VOT by spectrogram reading for three stop voiced sounds. We found that the VOT values are positive and short VOT. In addition, the smallest standard deviation for these three phonemes is 0.07 msec in sound /d/. So, the error rate of the algorithm for this sound is low, which represents 0.80%. Also, sound /dʔ/ standard deviation is high with 2.3 msec, because this sound is unique, and difficult to pronounce and detect manually by spectrogram. Therefore, this sound is difficult to detect automatically, and the error rate is 26.62% which is high comparing to other voiced stop sounds.

## 5. CONCLUSION AND FUTURE WORK

In this paper, we detected the VOT value automatically for voiced stop sounds in MSA Arabic namely /d/, /dʔ/ and /b/ and compared the accuracy algorithm with spectrogram reading. Automatically detecting VOT in speech signal is a challenging problem because it combines temporal and frequency structure over short duration. Our algorithm depends on reading the power signal. We create our database and the structure CV-CV-CV. We ended to a conclusion that the error rates were 0.80%, 26.62% and 11.71% for /d/, /dʔ/ and /b/ respectively. Moreover, we found that the standard deviation between the algorithm and spectrogram reading is low in /d/ sound because this sound is easy to pronounce, and high in the unique /dʔ/ sound, the main reason is that it is difficult to pronounce.

As a future and continuing research of this effort, we will continue to study automatic VOT for unvoiced stop sounds in MSA Arabic and compare our algorithm with manual spectrogram reading. Also, we will extend our database. On the other hand, we will study the environmental effect such as gender, age, effect of neighboring phonemes and place of autocorrelation to the VOT value.



## ACKNOWLEDGEMENTS

Grateful acknowledgment to King Abdulaziz City for Science and Technology for its support of this work. Also, acknowledgment to Eng. Hazzam Alhakami.

## REFERENCES

- [1] L. Lisker and A. S. Abramson, "A Crosslanguage Study of Voicing in Initial Stops: Acoustical Measurements", *Word*, Volume 20, No. 3, , pp. 384-442, December 1964.
- [2] M. Alghamdi, "Voice Print": Voice Onset Time as a Model. *Arab Journal for Security Studies and Training*. 21, pp. 42: 89-118, 2006 (in Arabic).
- [3] Cho, T. and Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, 27:207–229.
- [4] Sonderegger, M. and Keshet, K., "Automatic measurement of voice onset time using discriminative structured prediction", *Journal Acoustical Society of America, Soc. Am.*, 132, Dec 2012.
- [5] A. S. Abramsont and L. Lisker, "A Cross-language Experiments in identification and Discrimination", 21-24 May 1968.
- [6] Okalidou, A., Petinou, K., Theodorou, E. & Karasimou, E. (2010). Development of voice onset time in Standard Greek and Cypriot Greek speaking preschoolers. *Clinical Linguistics and Phonetics*, 24(7), 503–519.
- [7] J. Jiang, M. Chen, and A. Alwan,"On the perception of voicing in syllable-initial plosives in noise", *Journal of the Acoustical Society of America*, vol. 119, no. 2, Febuary 2006.
- [8] P. Sweeting, and R. Baken, "Voice onset time in normal-aged population", *Journal of Speech and Hearing Research*. 25, 129-134, 1982.
- [9] S. Das, and J. Hansen, "Detection of voice onset time (VOT) for unvoiced stops (/p/, /t/, /k/) using the Teager energy operator (TEO) for automatic detection of accented English", *Proceedings of the 6th NORDIC Signal Processing Symposium (NORSIG'4)*, pp. 344 – 347, June 9-11 2004.
- [10] G. E. Peterson and I. Lehiste, "Duration of syllable nuclei in English", *Journal of the Acoustical Society of America* 32, PP. 693-703, 1960.
- [11] J. E. Flege and R. Port, "Cross-language phonetic interference: Arabic to English", *Language and Speech* 24, PP. 125-146, 1981.
- [12] Stölten, Katrin; Abrahamsson, Niclas; Hyltenstam, Kenneth, Effects of Age of Learning on Voice Onset Time: Categorical Perception of Swedish Stops by Near-native L2 Speakers, December 2014, *Language & Speech*;Dec2014, Vol. 57 Issue 4, p425.
- [13] Caramazza A. and Yeni-Komshian G. H,(1974). Voice onset time in two French dialects, *Journal of Phonetics*, 2, 239–245.
- [14] C. Prakash, N. Dhananjaya, S. Gangashetty, Bessel features for detection of voice onset timeusing AM-FM signal, in *Proc. of Int. Conf. on the Systems, Signals and Image Processing (IWSSIP)*, (IEEE, Sarajevo, Bosnia and Herzegovina, 2011), pp. 1–4.

- [15] S. Al-Ani, "Arabic Phonology", The Hague, 1970.
- [16] F. Mitleb, "Voice onset time of Jordanian Arabic stops", The 3rd International Conference on Arabic Language Processing (CITALA'09), Rabat, Morocco, pp. 133-135, May 4-5 2009.
- [17] G. H. Yeni-Komshian, A. Caramaza and M. S. Preston, "A study of voicing in Lebanese Arabic", Journal of Phonetics 5:1, PP. 35-48, 1977.
- [18] Aldahri S., Studying VOT of unique Arabic stop and designing system to identify stop sounds in Modern Standard Arabic (MSA), Intelligent Signal Processing and Communication Systems (ISPACS), November 12-15, 2013, Naha, Okinawa, Japan.
- [19] Alotiabi Y. and Aldahri S., Phobnnetic Investigation of MSA Arabic Stops (/t, d/)", Yantai, China, 2010.
- [20] Aldahri S., The Effect of Arabic Emphaticness on Voice Time Onset (VOT), This conference (the 3rd International Conference on Audio, Language and Image Processing (ICALIP 2012), China Shanghai, 2012.
- [21] J. Deller, J. Proakis and J. H. Hansen, "Discrete-Time Processing of Speech Signal", MacmillAnn, 1993, Autor Engineering in computer science, researcher in field speech processing.

# APPROACHES IN USING EXPECTATION-MAXIMIZATION ALGORITHM FOR MAXIMUM LIKELIHOOD ESTIMATION OF THE PARAMETERS OF A CONSTRAINED STATE SPACE MODEL WITH AN EXTERNAL INPUT SERIES

Chengliang Huang<sup>1</sup>, Xiao-Ping Zhang<sup>1</sup> and Fang Wang<sup>2</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, Ryerson University  
Toronto, On, Canada

c28huang@ryerson.ca, xzhang@ee.ryerson.ca

<sup>2</sup>School of Business and Economics, Wilfrid Laurier University  
Waterloo, On, Canada

fwang@wlu.ca

## ABSTRACT

*EM algorithm is popular in maximum likelihood estimation of parameters for state-space models. However, extant approaches for the realization of EM algorithm are still not able to fulfill the task of identification systems, which have external inputs and constrained parameters. In this paper, we propose new approaches for both initial guessing and MLE of the parameters of a constrained state-space model with an external input. Using weighted least square for the initial guess and the partial differentiation of the joint log-likelihood function for the EM algorithm, we estimate the parameters and compare the estimated values with the “actual” values, which are set to generate simulation data. Moreover, asymptotic variances of the estimated parameters are calculated when the sample size is large, while statistics of the estimated parameters are obtained through bootstrapping when the sample size is small. The results demonstrate that the estimated values are close to the “actual” values. Consequently, our approaches are promising and can be applied in future research.*

## KEYWORDS

*State-Space Model, Maximum Likelihood Estimation, Expectation Maximization Algorithm, Kalman filtering and smoothing, Asymptotic variances, Bootstrapping*

## 1. INTRODUCTION

One of the advantages of the state-space models for linear dynamic systems is their ability to fit more parsimonious structures with fewer parameters to describe a multivariate time series. As a result, the application of state-space models is not limited to engineering practice.

For example, in certain scenarios, the dynamics of the brand equity of a firm can be represented by a state-space model in the form of Kalman filter [1],

$$\begin{cases} x_t = \alpha x_{t-1} + \gamma u_t + e_t \\ y_t = h x_t + w_t \end{cases}, \quad (1)$$

where  $x_t$  is the invisible brand equity,  $u_t$  is the investment, the external input,  $y_t$  is the brand performance, the output, at a certain step  $t$ . The transition coefficient,  $\alpha$  ( $|\alpha| < 1$ ), and the input coefficient,  $\gamma$ , are parameters of interest. The observation coefficient,  $h$ , is constrained to be a constant, i.e.,  $h=1$ . The observation noise,  $w_t$ , the process noise,  $e_t$ , and the initial state,  $x_0$ , at each step,  $\{x_0, e_1, \dots, e_t, w_1 \dots w_t\}$ , are all assumed to be mutually independent, where  $w_t \sim N(0, \sigma_w^2)$ ,  $e_t \sim N(0, \sigma_e^2)$ , and  $x_0 \sim N(\mu_0, \sigma_0^2)$ .

As exemplified in Equation (1) where  $h=1$ , a constrained state-space model means that some parameters, some elements in the parameter matrix of a Kalman filter, are fixed or shared thus not all of them have to be estimated. Conversely, if the model is unconstrained, any parameters or any elements of each parameter matrix will be estimated. Moreover, the model has an time-variant external input series,  $u_t$ . By the way, parameter estimation of state space models with external inputs can be seen as supervised problems while that of state space models without external inputs can be treated as unsupervised problems.

Maximum likelihood estimation (MLE) [2] is used to obtain the time-invariant parameters of the Kalman filter from input,  $u_t$ , and the output,  $y_t$ . Because of the existence of hidden variables,  $x_t$ , in the formulated likelihood function, expectation-maximization (EM) algorithm [3], an iterative method, is used to complete MLE of parameters in state-space models.

The approaches for the implementation of EM algorithm in the MLE of unconstrained state-space model are delivered in [4]-[7]. The approaches for certain constrained state-space models are covered in [7]-[10]. Since those constraints are quite specific, the approaches proposed are lack of generality. On the other hand, approaches for state-space models with external inputs are usually not provided except in [7]-[9]. However, the state space model in [7]-[9] is not typical because the inputs are constants or do not affect the hidden variable. Especially in [5], the external input is actually a constant, which is a parameter to be estimated. Therefore, as we know from extant literature, no innovative methods are explored to estimate the unknown parameters of such constrained state-space models in the recent years.

Our research is to find appropriate methods for the supervised problem: using EM algorithm for the MLE of those constrained state-space models who have external inputs. In this paper, we use the model represented by Equation (1) as an example. Our target is to estimate the parameters,  $\alpha$ ,  $\gamma$ ,  $\sigma_e$ ,  $\sigma_w$ ,  $\mu_0$  and  $\sigma_0$ . Our work is carried out mainly in two phases: (i) initial guessing using our innovative approach, and (ii) implementation of MLE using EM algorithm with approaches different from those in extant literature. In addition, if the sample size is large, we use the asymptotic variances of the estimated parameters to check the accuracy of the estimation. If the sample size is small, we introduced bootstrapping to examine the distribution of the estimated parameters.

## 2. INITIAL GUESSING

The initial guess is performed through two steps. Firstly, the system parameter  $\alpha$  and the variances  $\sigma_e$ ,  $\sigma_w$  and  $\sigma_0$  are guessed using autocovariance of the observations. Then the initial state mean  $\mu_0$  and input parameter  $\gamma$  are guessed using weighted linear square.

### 2.1. Guessing System Parameter and the Variances

Denote  $\eta_t = \sum_{i=0}^{\infty} \alpha^i e_{t-i}$ , from (1) it can be proved that

$$x_t = \frac{\gamma u_t}{1-\alpha} + \eta_t. \quad (2)$$

For  $h = 0, 1, 2 \dots$ , we can obtain the covariance and autocovariance of  $\eta_t$  as

$$\gamma_\eta(h) = \frac{\sigma_e^2 \alpha^h}{1-\alpha^2}, \quad (3)$$

while it is obvious that

$$\sigma_0^2 = \frac{\sigma_e^2}{1-\alpha^2}. \quad (4)$$

which means that we don't have to estimate  $\sigma_e$  and  $\sigma_0$  separately, as was performed in extant literature.

Moreover, we have the variance of  $y_t$

$$\gamma_y(0) = v_0 + v_w, \quad (5)$$

and the covariance of  $y_t$  when  $h = 1, 2, 3 \dots$ ,

$$\gamma_y(h) = \gamma_\eta(h). \quad (6)$$

Hence we can obtain the guessed initial values for system parameter,  $\alpha$ , and standard deviations,  $\sigma_e$  and  $\sigma_w$ , from

$$\begin{cases} \alpha = \frac{\gamma_y(2)}{\gamma_y(1)} \\ \sigma_e^2 = \frac{(1-\alpha^2)\gamma_y(1)}{\alpha} \\ \sigma_w^2 = \gamma_y(0) - \frac{\sigma_e^2}{1-\alpha^2} \end{cases}. \quad (7)$$

## 2.2. Guessing Initial State Mean and Input Parameter

Denoting

$$z_t = \sum_{i=1}^t \alpha^{t-i} u_i, \quad (8)$$

and

$$\zeta_t = \sum_{i=1}^t \alpha^{t-i} e_i + w_t, \quad (9)$$

we have,

$$y_t = \alpha^t \mu_0 + \gamma z_t + \zeta_t, \quad (10)$$

where  $\zeta_t \sim N(0, \sigma_\zeta^2)$  and

$$\sigma_\zeta^2 = \frac{1-\alpha^{2t}}{1-\alpha^2} \sigma_e^2 + \sigma_w^2 \quad (11)$$

In order to estimate  $\mu_0$  and  $\gamma$ , we perform linear regression between  $y_t$ , as dependent variable, and  $\alpha^t$  and  $z_t$ , as independent variables, using  $T$  samples of  $u_t$  and  $y_t$ . Since  $\zeta_t$  is heteroscedastical, we apply weighted least square (WLS). WLS finds its optimum when the weighted sum,  $S$ , of squared residuals is minimized where

$$S = \sum_{t=1}^T \frac{(y_t - \alpha^t \mu_0 - \gamma z_t)^2}{\frac{1-\alpha^{2t}}{1-\alpha^2} \sigma_e^2 + \sigma_w^2}. \quad (12)$$

Denote that  $v_t = (1 - \alpha^{2t})\sigma_e^2 + (1 - \alpha^2)\sigma_w^2$ , we solve the gradient equation (regarding  $\mu_0$  and  $\gamma$  respectively) for the sum of squares

$$\begin{cases} \sum_{t=1}^T \frac{\alpha^t(y_t - \alpha^t \mu_0 - \gamma z_t)(1 - \alpha^2)}{v_t} = 0 \\ \sum_{t=1}^T \frac{z_t(y_t - \alpha^t \mu_0 - \gamma z_t)(1 - \alpha^2)}{v_t} = 0 \end{cases}. \quad (13)$$

Therefore, we will have initial guess about as below:

$$\begin{cases} \gamma = \frac{\sum_{t=1}^T \frac{\alpha^{2t}}{v_t} \sum_{t=1}^T \frac{z_t y_t}{v_t} - \sum_{t=1}^T \frac{\alpha^t y_t}{v_t} \sum_{t=1}^T \frac{\alpha^t z_t}{v_t}}{\sum_{t=1}^T \frac{\alpha^{2t}}{v_t} \sum_{t=1}^T \frac{z_t^2}{v_t} - \sum_{t=1}^T \frac{\alpha^t z_t}{v_t} \sum_{t=1}^T \frac{\alpha^t z_t}{v_t}} \\ \mu_0 = \frac{\sum_{t=1}^T \frac{\alpha^t y_t}{v_t} \sum_{t=1}^T \frac{z_t^2}{v_t} - \sum_{t=1}^T \frac{\alpha^t z_t}{v_t} \sum_{t=1}^T \frac{y_t z_t}{v_t}}{\sum_{t=1}^T \frac{\alpha^{2t}}{v_t} \sum_{t=1}^T \frac{z_t^2}{v_t} - \sum_{t=1}^T \frac{\alpha^t z_t}{v_t} \sum_{t=1}^T \frac{\alpha^t z_t}{v_t}} \end{cases}. \quad (14)$$

### 3. ESTIMATION USING EM ALGORITHM

The conditional density for the states and outputs are,

$$P(x_t | x_{t-1}) = \frac{1}{\sigma_e \sqrt{2\pi}} \exp \left[ -\frac{(x_t - \alpha x_{t-1} - \gamma u_t)^2}{2\sigma_e^2} \right], \quad (15)$$

$$P(y_t | x_t) = \frac{1}{\sigma_w \sqrt{2\pi}} \exp \left[ -\frac{(y_t - x_t)^2}{2\sigma_w^2} \right]. \quad (16)$$

Assuming a Gaussian initial state density

$$P(x_0) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left[ -\frac{(x_0 - \mu_0)^2}{2\sigma_0^2} \right], \quad (17)$$

By the Markov property implicit in this model, we calculate the joint probability, not the partial probability used by Shumway (2011), regarding all T samples of  $x_t$  and  $y_t$ , denoted as  $\{x\}$  and  $\{y\}$  respectively:

$$P(\{x\}, \{y\}) = P(x_0) \prod_{t=1}^T P(x_t | x_{t-1}) \prod_{t=1}^T P(y_t | x_{t-1}) \quad (18)$$

We denote the joint log probability as

$$\Omega = \log P(\{x\}, \{y\}). \quad (19)$$

According to Equation (4), we only need to estimate the parameter set,  $\Psi = \{\alpha, \gamma, \sigma_e, \sigma_w, \mu_0\}$ , through maximizing the objective function:

$$\begin{aligned} \Omega(\alpha, \gamma, \sigma_e, \sigma_w, \mu_0) = & -\frac{(1-\alpha^2)(x_0 - \mu_0)^2}{2\sigma_e^2} - \frac{1}{2} \log \frac{\sigma_e^2}{1-\alpha^2} - \sum_{t=1}^T \frac{(x_t - \alpha x_{t-1} - \gamma u_t)^2}{2\sigma_e^2} - \frac{1}{2} T \log \sigma_e^2 - \\ & \sum_{t=1}^T \frac{(y_t - x_t)^2}{2\sigma_w^2} - \frac{1}{2} T \log \sigma_w^2 - \frac{2T+1}{2} \log(2\pi). \end{aligned} \quad (20)$$

#### 3.1. EM Algorithm

Since the objective function expressed by Equation(20) depends on the unobserved data series,  $x_t (t = 1, 2, \dots, T)$ , we consider applying the EM algorithm conditionally with respect to the observed output series  $y_1, y_2, \dots, y_T$ . The objective function above has an input series. Accordingly, the input coefficient has to be estimated. Consequently, our approaches are unlike the approaches[6] used in the implementation of EM algorithm for linear dynamic systems.

The EM algorithm mainly has two steps: the E-STEP and the M-Step. During the E-Step, the parameters are assumed known, the hidden states and their variance are estimated over all the samples, and then the likelihood function constructed from joint probability are calculated.

During the M step, we have to find the parameter set  $\hat{P}(k) = \{\alpha(k), \gamma(k), \sigma_e(k), \sigma_w(k), \mu_0(k)\}$  for the  $k$ th counts of the recursions by maximizing the conditional expectation, or the above objective function.

The overall procedure for the implementation of EM algorithm is as below:

- (i) Initialize the procedure by selecting the guessed values as starting values for the parameters.  
On iteration  $k$ , ( $k=1,2,\dots$ )
- (ii) Compute the log-likelihood (optional),
- (iii) Use the parameters to obtain the smoothed values of the hidden states and their correlations, for  $t= 1,2, \dots, T$ .
- (iv) Use the smoothed values to calculate the updated parameters.
- (v) Repeat Steps (ii) – (iv) to convergence.

We mainly perform two sub-steps in the E-step of EM algorithm: Kalman filtering and Kalman smoothing.

### 3.2. Kalman Filtering and Smoothing

Assuming that we already know the parameter set  $\{\alpha, \gamma, \sigma_e, \sigma_w, \mu_0, \sigma_0\} (x_0 \sim N(\mu_0, \sigma_0^2))$ , and the observations  $y_t$  and  $u_t$ , we have the estimation of the hidden state, as well as the variances estimated based on the observations for the period 1 to  $t$ .

$$x_{t|t-1} = \alpha x_{t-1|t-1} + \gamma u_t, \quad (21a)$$

$$v_{t|t-1} = \alpha^2 v_{t-1|t-1} + \sigma_e^2, \quad (21b)$$

$$\tilde{y}_t = y_t - x_{t|t-1}, \quad (21c)$$

$$s_t = v_{t|t-1} + v_w, \quad (21d)$$

$$k_t = v_{t|t-1} s_t^{-1}, \quad (21e)$$

$$x_t^t = x_{t|t-1} + k_t \tilde{y}_t, \quad (21f)$$

$$v_t^t = (1 - k_t) v_{t|t-1}, \quad (21g)$$

where  $x_{0|0} = \mu_0$  and  $v_{0|0} = \sigma_0^2$ .

According to [6], to compute  $E[x_t | \{y, u\}] \equiv x_t^T$  and the correlation matrices  $p_t \equiv v_t^T + x_t^T (x_t^T)'$  one performs a set of backward recursion using

$$j_{t-1} = \alpha \frac{v_{t-1|t-1}}{v_{t|t-1}}, \quad (22)$$

$$x_{t-1}^T = x_{t-1|t-1} + j_{t-1}(x_t^T - \alpha x_{t-1|t-1} - \gamma u_t), \quad (23)$$

$$v_{t-1}^T = v_{t-1|t-1} + j_{t-1}(v_t^T - v_{t|t-1})j_{t-1}', \quad (24)$$

where  $x_T^T = x_{T|T}$  and  $v_T^T = v_{T|T}$ . We also have  $p_{t,t-1} \equiv V_{t,t-1}^T + x_t^T(x_{t-1}^T)'$ , where  $v_{t,t-1}^T$  can be obtained through the backward recursions

$$v_{t-1,t-2}^T = v_{t-1|t-1}j_{t-2}' + j_{t-1}(v_{t,t-1}^T - \alpha v_{t-1|t-1})j_{t-2}', \quad (25)$$

which is initialized using  $v_{T,T-1}^T = \alpha(1 - k_T)v_{T-1}^{T-1}$ .

Note that the state estimate,  $x_t^T$ , differs from the one computed in a Kalman filter in that it is the smoothed estimator of  $x_t$  based on all of the observed data (and input data), i.e. it depends on past and future observations; the Kalman filter estimates  $E[x_t|\{y\}_1^t]$  is the usual Kalman filter estimator based on the observed data up to the current time instant  $t$ .

### 3.3. Expected Log-Likelihood Formulation

After we have got the expected values for  $x_0$  and  $x_t$  as  $x_0^T \equiv E[x_0|\{y, u\}]$  and  $x_t^T \equiv E[x_t|\{y, u\}]$  respectively, we can calculate the expectation of the log-likelihood

$$E(\Omega) = E[\log P(\{x\}, \{y\})]. \quad (26)$$

Denote

$$\begin{cases} P_{t-1}^T = \sum_{t=1}^T E[x_{t-1}x_{t-1}'|\{y\}] \\ P_t^T = \sum_{t=1}^T E[x_t x_t'|\{y\}] \\ P_{t,t-1}^T = \sum_{t=1}^T E[x_t x_{t-1}'|\{y\}] \end{cases}, \begin{cases} M_{t-1}^T = \sum_{t=1}^T E(x_{t-1}u_t) \\ M_t^T = \sum_{t=1}^T E(x_t u_t) \\ U_t^T = \sum_{t=1}^T u_t^2 \end{cases}, \text{ and } \begin{cases} W_{t-1}^T = \sum_{t=1}^T E(x_{t-1}y_t) \\ W_t^T = \sum_{t=1}^T E(x_t y_t) \\ Y_t^T = \sum_{t=1}^T y_t^2 \end{cases},$$

we have

$$E(\Omega) = -\frac{1}{2}E(\Omega_0) - \frac{1}{2}E(\Omega_1) - \frac{1}{2}E(\Omega_2) - \frac{2T+1}{2}\log(2\pi) \quad (27)$$

where

$$E(\Omega_0) = \frac{1-\alpha^2}{\sigma_e^2} [v_0^T + (x_0^T - \mu_0)^2] + \log \frac{\sigma_e^2}{1-\alpha^2} \quad (27a)$$

$$E(\Omega_1) = \sigma_e^{-2} (P_t^T + \alpha^2 P_{t-1}^T + \gamma^2 U_t^T - 2\alpha P_{t,t-1}^T - 2\gamma M_t^T + 2\alpha\gamma M_{t-1}^T) + T \log \sigma_e^2 \quad (27b)$$

$$E(\Omega_2) = \sigma_w^{-2} (Y_t^T + P_t^T - 2W_t^T) + T \log \sigma_w^2 \quad (27c)$$

### 3.4. Estimation of the Parameters

We use the first order condition on partial derivatives of  $E(\Omega)$  to individual parameters to obtain the gradient and then the values of individual parameters. This method is not the multivariate regression approach [6]. The parameters are chosen when the objective function is maximized, i.e., the gradients are all zero. The estimates of  $\alpha$ ,  $\gamma$ ,  $\sigma_e^2$ ,  $\sigma_w^2$  and  $\mu_0$  are from below five equations:

$$\alpha(1 - \alpha^2)v_0^T - \alpha\sigma_e^2 - \alpha(1 - \alpha^2)P_{t-1}^T + (1 - \alpha^2)P_{t,t-1}^T - \gamma(1 - \alpha^2)M_{t-1}^T = 0 \quad (28a)$$

$$\gamma U_t^T - M_t^T + \alpha M_{t-1}^T = 0 \quad (28b)$$

$$(1 - \alpha^2)v_0^T - (1 + T)\sigma_e^2 + P_t^T + \alpha^2 P_{t-1}^T + \gamma^2 U_t^T - 2\alpha P_{t,t-1}^T - 2\gamma M_t^T + 2\alpha\gamma M_{t-1}^T = 0 \quad (28c)$$



$$\sigma_w^2 = \frac{1}{T}(Y_t^T + P_t^T - 2W_t^T) \quad (28d)$$

$$\mu_0 = x_0^T \quad (28e)$$

Moreover, we use the second orders of the derivatives of  $E(\Omega)$  to calculate the second derivatives and then the information matrix. Most of the second order derivatives of  $E(\Omega)$  are zero except those listed below:

$$\frac{\partial^2 E(\Omega)}{\partial \alpha^2} = \frac{v_0^T + (x_0^T - \mu_0)^2 - P_{t-1}^T}{\sigma_e^2} - \frac{1 + \alpha^2}{(1 - \alpha^2)^2}, \quad (29)$$

$$\frac{\partial^2 E(\Omega)}{\partial \gamma \partial \alpha} = \frac{\partial^2 E(\Omega)}{\partial \alpha \partial \gamma} = -\frac{M_{t-1}^T}{\sigma_e^2}, \quad (30)$$

$$\frac{\partial^2 E(\Omega)}{\partial \sigma_e \partial \alpha} = \frac{\partial^2 E(\Omega)}{\partial \alpha \partial \sigma_e} = \frac{2(\alpha P_{t-1}^T - P_{t,t-1}^T + \gamma M_{t-1}^T) - 2\alpha[v_0^T + (x_0^T - \mu_0)^2]}{\sigma_e^3}, \quad (31)$$

$$\frac{\partial^2 E(\Omega)}{\partial \gamma^2} = -\frac{U_t^T}{\sigma_e^2} \quad (32)$$

$$\frac{\partial^2 E(\Omega)}{\partial \sigma_e \partial \gamma} = \frac{\partial^2 E(\Omega)}{\partial \gamma \partial \sigma_e} = \frac{2(\gamma U_t^T - M_t^T + \alpha M_{t-1}^T)}{\sigma_e^3} \quad (33)$$

$$\frac{\partial^2 E(\Omega)}{\partial \sigma_e^2} = \frac{T+1}{\sigma_e^2} - \frac{3(1-\alpha^2)[v_0^T + (x_0^T - \mu_0)^2]}{\sigma_e^4} - \frac{3(P_t^T + \alpha^2 P_{t-1}^T + \gamma^2 U_t^T - 2\alpha P_{t,t-1}^T - 2\gamma M_t^T + 2\alpha\gamma M_{t-1}^T)}{\sigma_e^4} \quad (34)$$

$$\frac{\partial^2 E(\Omega)}{\partial \sigma_w^2} = \frac{T}{\sigma_w^2} - \frac{3(Y_t^T + P_t^T - 2W_t^T)}{\sigma_w^4} \quad (35)$$

$$\frac{\partial^2 E(\Omega)}{\partial \mu_0^2} = -\frac{1 - \alpha^2}{\sigma_e^2} \quad (36)$$

According to Cramer-Rao Theorem, the MLE is an efficient estimate. When the sample size is large enough, the asymptotic variances of the estimates can be considered as the metric of the accuracy of the estimation. The asymptotic variances are calculated using the inverse of the information matrix, which is the inverse of the negative of the expected value of the Hessian matrix. The vector of the asymptotic variances of the estimates is

$$\begin{bmatrix} \frac{(T+1)(\alpha^2 - 1)^2 U \sigma_e^2}{(T+1)(1-\alpha^2)^2 (P_0 U - U \sigma_0^2 - M_0^2) + (1+T-\alpha^2+T\alpha^2) U \sigma_e^2} \\ \frac{[(1+T)(1-\alpha^2)^2 (P_0 - \sigma_0^2) + (1+T-\alpha^2+T\alpha^2) \sigma_e^2] \sigma_e^2}{(T+1)(1-\alpha^2)^2 (P_0 U - U \sigma_0^2 - M_0^2) + (1+T-\alpha^2+T\alpha^2) U \sigma_e^2} \\ \frac{[(1-\alpha^2)^2 (P_0 U - M_0^2 - U \sigma_0^2) + (1+\alpha^2) U \sigma_e^2] \sigma_e^2}{(T+1)(1-\alpha^2)^2 (P_0 U - U \sigma_0^2 - M_0^2) + (1+T-\alpha^2+T\alpha^2) U \sigma_e^2} \\ \frac{\sigma_w^2}{2T} \\ \sigma_0^2 \end{bmatrix}$$

If the sample size is small, we introduce boot-strapping procedure where the estimates are obtained from likelihood constructed from re-sampled standardized innovation,  $\tilde{y}_t$ , in Equation (21c). Moreover, the mean squared errors (MSE) of the state variables which is estimated from Equations (21a-g) using estimated parameters are also estimated.

## 4. SIMULATION AND RESULTS

The output data is generated through presetting the input series and the values of the parameters. We implement the initial guessing, EM iteration, and finally obtain the parameter estimates. This makes it easier to evaluate our work by comparing the actual values with the estimated values, or by checking the standard deviation of the estimates.

### 4.1. Data Generation

We generate data from the state-space model described as Equation (1). We assume  $\alpha = 0.8$ ,  $\gamma = 1.5$ , and  $\mu_0 = 0$ . Moreover, the process noises,  $e_t$ , and observation noise,  $w_t$ , are generated independently where  $e_t \sim N(0, 1.1^2)$  and  $w_t \sim N(0, 0.9^2)$ . We assume that the input,  $u_t$ , is a slow changing periodical square wave signal whose period is 10 time unit. The standard deviation of initial state,  $\sigma_0$ , is not needed during the data generation but can be calculated according to Equation (2). The expected log-likelihood can be calculated using Equation (27). Both are treated as “actual” values to be compared with guessed values and estimated values.

We performed our simulation using two different sample sizes: the large size of 1000 and the small size of 50. When the sample size is small, we applied bootstrapping method to estimate the accuracy of the estimate.

### 4.2. Results

We provide the results of the simulation with small sample size of 50 in Table 1, and the results of the simulation with large sample size of 1000 in Table 2.

Table 1. The parameters estimate with small sample size

Parameters	Actual	Guessed	Estimated	Std. Dev.
$\alpha$	0.8	0.879	0.801	0.023
$\gamma$	1.5	1.075	1.438	0.104
$\sigma_e$	1.1	1.400	0.708	0.158
$\sigma_w$	0.9	0.922	0.862	0.193
$\mu_0$	0	0.992	2.143	1.231
$\sigma_0$	1.83	3.811	1.183	0.252

In Table 1 and Table 2, we displayed the actual values, the guessed values, the estimated values and the standard deviations of the estimated values for transition coefficient,  $\alpha$ , input coefficient,  $\gamma$ , standard deviation of process errors,  $\sigma_e$ , standard deviation of observation errors,  $\sigma_w$ , mean of initial state,  $\mu_0$ , and standard deviation of initial state,  $\sigma_0$ . In general, the guessed value is near the actual value while the estimated value is much more close to the actual value than the guessed ones for the parameters of the most interest:  $\alpha$  and  $\gamma$ . The standard deviations in Table 1 are from bootstrapped distribution while the standard deviations in Table 2 are from the asymptotic variances.

Table 2: The parameters estimated with large sample size

Parameters	Actual	Guessed	Estimated	Std. Dev.
$\alpha$	0.8	0.879	0.800	0.0001
$\gamma$	1.5	1.276	1.424	0.0012
$\sigma_e$	1.1	2.024	1.033	0.0011
$\sigma_w$	0.9	0.533	0.971	0.0005
$\mu_0$	0	2.459	1.667	2.9614
$\sigma_0$	1.83	4.250	1.721	NA

In general, the guessed value is near the ‘actual’ value while the estimated value is much closer to the “actual” value than the guessed ones, especially for the parameters of interest:  $\alpha$  and  $\gamma$ . It is worth noting that the deviations of the estimated values are larger than the asymptotic ones due to the imperfectly generated data.

## 5. CONCLUSIONS

The research is to validate our innovative approaches in the application of EM algorithm in the MLE of a constrained dynamic linear system with external input. There are two main contributions in this research. Firstly, we realized that  $\sigma_0$  and  $\sigma_e$  has the relationship expressed by Equation (4) thus we don’t have to estimate both of them during the implementation of EM algorithm. Accordingly, the likelihood function in Equation (27) is not similar with those researchers who ignored the relationship. Secondly, in initial guessing of the value of input coefficient and the mean of initial state, we introduce weighted least square for the guessing of input coefficient,  $\gamma$ , and the mean of the initial state,  $\mu_0$ .

It is obvious that more techniques have to be discovered for the initial guessing, and the estimating based on the guessed initial guessing of, the parameter values, especially for the implementation of the M-step of the EM algorithm. The approaches we proposed can be a new start point for the future research on the estimation of dynamic systems with higher dimensions of external inputs, hidden states and observations.

## REFERENCES

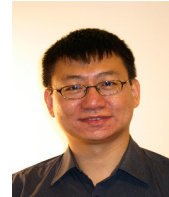
- [1] Kalman, R. E., (1960) "A new approach to linear filtering and prediction problems", *Journal of Basic Engineering*, 82 (1), pp. 35–45.
- [2] James D. Hamilton, *Time Series Analysis*, Princeton University Press, Princeton, New Jersey, 1996.
- [3] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977) "Maximum likelihood from incomplete data via the EM algorithm", *Journal of the Royal Statistical Society, Series B*, Vol. 39, No. 1, pp. 1–38.
- [4] Ghahramani, Z. and Hinton, G. E. (1996) "Parameter estimation for linear dynamic systems", Technical report CRC-TR-96-2, Department of Computer Science, University of Toronto.
- [5] Deng, L. and Shen, X. (1997) "Maximum likelihood in statistical estimation of dynamic systems: decomposition algorithm and simulation results", *Signal Processing*, Vol. 57, No. 1.
- [6] Shumway, Robert H. and Stoffer, David S. (2011) *Time Series Analysis and Its Applications: With R Examples*, New York, NY, Springer, New York.
- [7] Holmes, Elizabeth Eli. (2013) "Derivation of an EM algorithm for constrained and unconstrained multivariate autoregressive state-space (MARSS) models", Technical Report, Northwest Fisheries Center, NOAA Fisheries.
- [8] Wu, Lilian Shiao-Yen, Pai, Jeffrey S. and Hosking, J. R. M. (1996) "An algorithm for estimating parameters of state-space models", *Statistics & Probability Letters*, Vol. 28, pp. 99–106.
- [9] Zuur, A. F., Fryer, R. J., Jolliffe, I. T., Dekker, R. and Beukema, J. J. (2003) "Estimating common trends in multivariate time series using dynamic factor analysis", *Environmetrics*, Vol. 14, No. 7, pp. 665–685.
- [10] Dewar, M. Scerri, K. and Kadirkamanathan, V. (2009) "Data-driven spatio-temporal modeling using the integro-difference equation", *IEEE Transactions on Signal Processing*, Vol. 57, No. 1, pp. 83–91.

## AUTHORS

**Chengliang Huang** is currently a Ph. D. candidate at Ryerson University. He had been a senior engineer and project manager at Huawei and Intel for four years respectively. He graduated as MASc in Electrical and Computer Engineering, Ryerson University in 2009, MBA in Economic Management, Beijing Jiaotong University and B. Sci. in Electronic Engineering, Fudan University in 1990. He has published papers on peer-reviewed academic journals, including IEEE Sensor Journal and Wireless Personal Communications. His research interest includes wireless communications, random signal processing, and system identification, time series analysis, marketing data mining/analysis, and marketing productivity.



**Xiao-Ping Zhang** (M'97, SM'02) received the B.S. and Ph.D. degrees from Tsinghua University, in 1992 and 1996, respectively, all in electronic engineering. He holds an MBA in Finance, Economics and Entrepreneurship with Honors from Booth School of Business, the University of Chicago. Since Fall 2000, he has been with the Department of Electrical and Computer Engineering, Ryerson University, where he is now Professor, Director of Communication and Signal Processing Applications Laboratory (CASPAL) and Program Director of Graduate Studies. His research interests include multimedia communications and signal processing, multimedia retrieval and video content analysis, sensor networks and electronic systems, computational intelligence, and applications in bioinformatics, finance, and marketing. He is a frequent consultant for biotech companies and investment firms. He is cofounder and CEO for EidoSearch, Inc., which offers a content-based search and analysis engine for financial data. Dr. Zhang is a registered Professional Engineer in Ontario, Canada, a Senior Member of IEEE and a member of Beta Gamma Sigma Honor Society. He is the publicity co-chair for ICME'06 and program co-chair for ICIC'05. He is currently an Associate Editor for IEEE Transactions on Signal Processing, IEEE Signal Processing Letters and Journal of Multimedia.



**Dr. Wang's** research focuses on brand equity assessment, marketing data mining/analysis, long-term marketing productivity, and marketing strategy. She also works on e-commerce and information systems related topics. Her research on brand equity assessment is funded by the Social Sciences and Humanities Research Council of Canada (SSHRC). She is a key team member for a NSERC (Natural Sciences and Engineering Research Council of Canada) supported project on financial market information analysis. Her work has appeared in many academic journals, including Journal of Marketing, Journal of the Academy of Marketing Science, Information & Management, Journal of Consumer Marketing, among others.



# MLTDD : USE OF MACHINE LEARNING TECHNIQUES FOR DIAGNOSIS OF THYROID GLAND DISORDER

Izdiyar Al-muwaffaq and Zeki Bozkus

Department of Computer Engineering, Kadir Has University  
izdiyar.Mofek@stu.khas.edu.tr, zeki.bozkus@khas.edu.tr

## ABSTRACT

*Machine learning algorithms are used to diagnosis for many diseases after very important improvements of classification algorithms as well as having large data sets and high performing computational units. All of these increased the accuracy of these methods. The diagnosis of thyroid gland disorders is one of the application for important classification problem. This study majorly focuses on thyroid gland medical diseases caused by underactive or overactive thyroid glands. The dataset used for the study was taken from UCI repository. Classification of this thyroid disease dataset was a considerable task using decision tree algorithm. The overall prediction accuracy is 100% for training and in range between 98.7% and 99.8% for testing. In this study, we developed the Machine Learning tool for Thyroid Disease Diagnosis (MLTDD), an Intelligent thyroid gland disease prediction tool in Python, which can effectively help to make the right decision, has been designed using PyDev, which is python IDE for Eclipse.*

## KEYWORDS

*Machine Learning, Thyroid diseases, CRT decision tree algorithm, PyDev, Python IDE.*

## 1. INTRODUCTION

Classification algorithms are very important category of supervised machine learning algorithms. These algorithms require a very large training sets. These training data sets are consisting of many features or attributes which describe the individual sample. Since we are doing supervised learning algorithm. All of the training set are labelled correctly. The classification algorithms such as decision trees and support vector machines (SVM), develop model with these data with many different parameters. When we have a new unlabeled sample, we can use the model to predict the label of the new sample. These techniques are used for disease diagnosis to help doctor to effectively label the new case.

The thyroid releases two principal hormones. The first is called thyroxine (T4) and the other one is triiodothyronine (T3) into the blood stream. The main functions of the thyroid hormones are to regulate the growth rate of metabolism. There are two common problems of thyroid disorder: Hyperthyroidism and Hypothyroidism. The first one releases too much thyroid hormone into the blood stream and the second one releases too low thyroid hormones to the blood stream. These

means that thyroids are very active in Hyperthyroidism. In contrast, thyroids are not active in Hypothyroidism [1] [2].

A decision tree is one of the very effective machine learning classifier where the algorithm makes a tree structure, where every non-leaf node denotes a test on an attribute, each branch performs an outcome of the test and each leaf node holds a class label. Decision tree is very simple and effective classifier so this algorithm can be used in several application areas such as medicine, financial analysis, astronomy and molecular biology for classification [3] [4]. We are going to use decision tree to identify and predict thyroid diseases whether new patient data (symptom or features) show thyroid disorder or normal.

This study developed a tool to be used for diagnosing thyroid diseases. We named our tool as a MLTDD (machine learning tool for thyroid disease diagnosis). MLTDD could predict 99.81% accuracy thyroid diseases with sample dataset.

The paper is organized as follows: The Section 2 describes the thyroid dataset and introduces decision tree techniques. The obtained experimental results in application are given in Section 3. Discussions and comparisons with previous work will be found in Section 4. Finally, Section 5 presents the conclusions.

## 2. METHODOLOGY USED

This section explains about the algorithm, language and software used for this work. The data set used for experimental purpose is downloaded from university of California of Irvin (UCI) repository site (web source <http://www.archive.ics.uci.edu/ml/datasets.html>). Details of data set is given in Section 2.1.

An implementation process of thyroid gland diagnosis process is depicted in Figure 1.

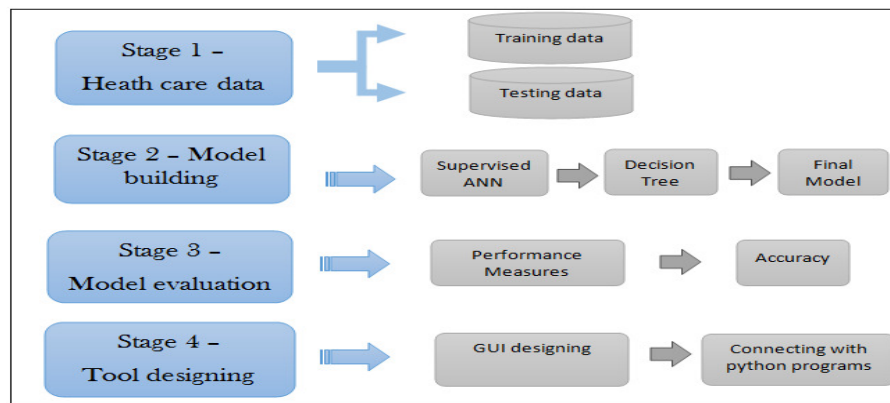


Figure 1: Implementation of MLTDD Process

### 2.1 Description of data-set

Thyroid dataset for this work had been collected from UCI machine learning repository. It consists of 7200 instances and 3 classes, 3772 are training instances, 3428 testing are instances

and 21 attributes as shown in Table 1 [5]. The task is to detect if a given patient has a normal condition (1) or suffers from hyperthyroidism (2) or hypothyroidism (3).

This section describes the main characteristics of the thyroid data set and its attributes: Each measurement vector consists of 21 values – 15 binary and 6 are continuous. Three classes are assigned to each of the measurement vectors, which correspond with hyper-rthyroidism, hypothyroidism and normal function of the thyroid gland (Table 2).

Table 1. General information

Thyroid Disease (ann) data set			
Type	Classification	Origin	Real world
Features	21	(continuous / categorical)	(6 / 15 )
Instances	7200	Classes	3
Missing values?			No

Table 2. Attribute description.

Attribute	Domain	Attribute	Domain	Attribute	Domain
Age	[0.01, 0.97]	Thyroid surgery	[0, 1]	Hypopituitary	[0, 1]
Sex	[0, 1]	Il31_treatment	[0, 1]	Psych	[0, 1]
On_thyroxine	[0, 1]	Query_hypothyroid	[0, 1]	TSH	[0.0, 0.53]
Query_on_thyroxine	[0, 1]	Query_hyperthyroid	[0, 1]	T3	[0.0005, 0.18]
On_antithyroid_medication	[0, 1]	Lithium	[0, 1]	TT4	[0.002, 0.6]
Sick	[0, 1]	Goiter	[0, 1]	T4U	[0.017, 0.233]
Pregnant	[0, 1]	Tumor	[0, 1]	FTI	[0.002, 0.642]
Class	{1,2,3}				

## 2.2 Decision Tree classification method

We used decision tree classification for our thyroid diseases diagnosis tool. Decision tree takes training datasets and construct an internal tree structure. The non-leaf node will have a question or a condition to check based on feature of the data. The leaf nodes of the tree have the labels.

The algorithm of constructing decision tree is a divide-and-conquer algorithm. The root of the tree asks one of the feature questions to splits datasets into multiple branches based on the answer of the parent question. If all the subsets of data in the branch have the same label, that branch will stop. It will not grow from that branch any further. Otherwise, the constructor algorithm can split the data further with new conditions from other features of the data sets. The algorithm will recursively try to create new branches of the tree based on the features (or attributes) of datasets. Basically algorithm try to understand which features effect the decision for labeling.

When we have a record of a new patient, we use the decision tree as if it is a flow-charts. We try to go the labeled leaf node by asking the same questions of the tree until we reach the leaf node. We then classify the new patient the same as the leaf node label which we reach from the root of the decision tree. There are huge number of research how to select which attributes is best to select for the root and non-leaf nodes. Basically, the constructor will choose the attributes to have maximum information gains [6].

## 2.3 MLTDD: Intelligent Thyroid Gland Disease Prediction System

We designed this application to be used by doctors, who are not expert computer users. So our design goal was to develop the user interface of the thyroid diseases application as user friendly as possible. The application has been implemented in Eclipse Python (PyDev plugged in eclipse JUNO Version: 4.2.2) coded in python 2.7. The GUI has been designed in Qt designer version 5.5.1 which is graphical user interface for Qt application. MLTDD runs on windows environment (Fig. 2).

Thyroid Gland predictor

Menu

**Intelligent Thyroid gland Disease Prediction System**

Enter Name:

Enter age:

Select Gender: Enter (M) for Male // (F) for Female

Enter TSH:

Enter T3:

Enter TT4:

Enter T4U:

Enter FTI:

The following are True or False ,please enter the letter ( T ) or ( F )

On-Thyroxine:

Pregnant:

Query\_hypothyroid:

psychology sympt:

**Predict**

The diagnosis of patient 1 is normal

**Clear All** **Quit**

[Click here for more information](#) Prediction accuracy is 99.7 % >> Decision Tree algorithm CRT was used>>

Fig. 2. Screen of prediction test of MLTDD app.



Ann dataset has been used to train the decision tree algorithm and create a model then predict a diagnosis for a new patient whose data will be entered by the user. Taking the data from GUI and pass them to a python program which uses the model created to classify the patient according to its data as a hypothyroidism or hyperthyroidism or a normal diagnosis, then by passing the diagnosis to the program which dealing with the GUI so that it can be printed out to the user.

### 3. RESULTS

For preprocessing, we applied “PCA” for feature selection but we noticed that the accuracy decreased from 99.644 to 97.533, we also applied another preprocessing method “ROUNDUM subset” but we got the same result. For the purpose of decreasing the number of features to be able to design an acceptable GUI, since the dataset has 21 features “information gain attribute evaluation” has been applied as a feature selection method. A good ranking has been obtained which helps to eliminate the ten least important attributes and keep the other 11 attributes. With 11 features we got a 99.70 % accuracy and it doesn't decrease so much, since the best accuracy we got with the all 21 features is 99.82% with 23% testing dataset as shown at Table 3. So we decided to eliminate those 10 feature to make my "GUI" more acceptable and easier to use. In Table 3, there are different accuracy values with different splitting of dataset comparing before and after applying the feature selection method and eliminating the less important features, in Figure 3 the graph shows this comparison as well.

Table 3: Specifies the values of Accuracy before and after eliminating 10 features.

Percentage split for testing dataset	Accuracy (%) With 21 features	Accuracy (%) With 11 features
10 %	99.58 %	99.17 %
15 %	99.72 %	99.35 %
20 %	99.79 %	99.51 %
23 %	99.82 %	99.58 %
30 %	99.68 %	99.68 %
32 %	99.65 %	99.70 %
33 %	99.62 %	99.66 %

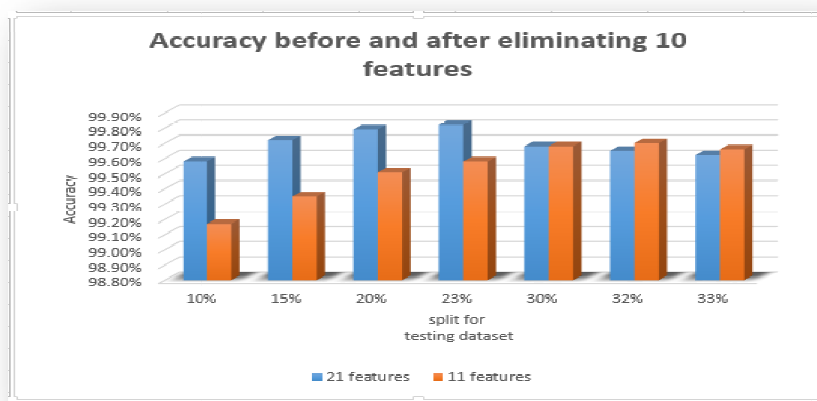


Figure 3: Accuracy comparison graph for the splitting percentage

## 4. RELATED WORK

Many methods and algorithms have been utilized in the past in medical disease classification. A skimpy abstract of them as follows:

Anupam Skukla et al. [8] suggested the diagnosis of thyroid disorders with Artificial Neural Networks (ANNs). Three ANN algorithms were; the Radial Basis Function (RBFN), the Back propagation algorithm (BPA), and the Learning Vector Quantization (LVQ) Networks have been utilized for the diagnosis.

Lale Ozyilmaz et al. [9] concentrated on suitable interpretation of the thyroid data. Several neural network methods such as fast back-propagation (FBP), Radial Basis Function (RBF), adaptive Conic Section Function Neural Network (CSFNN), and Multi-Layer Perceptron (MLP) with back-propagation (BP) have been utilized and compared for the diagnosis of thyroid disease.

Fatemeh Saiti et al. [10] proposed two algorithms, which are Support Vector Machines and Probabilistic Neural Network considering separating and classification the Hypothyroidism and hyperthyroidism diseases, which plays a vital role for thyroid diagnosis. These methods depend on robust classification algorithms, in order to deal with irrelevant and redundant features.

Carlos Ordonez et al. [11] compared decision tree rules with association rules. Association rules research for hidden patterns turning them out to be suitable for discovering predictive rules including subsets of the medical data set.

A.S.Varde et al. [12] has developed the clinical laboratory expert system in 1991 for the diagnosis of thyroid disorder. The system had considered clinical findings and the results of applicable laboratory tests along with the patient's medical history. The system had been execute using VP-Expert, version 2.02 that is commercially available software.

Palanichamy Jaganathan et al. [13] developed F-score feature selection method that used to nominate the most relevant features for classification of thyroid dataset. The result shows that their new feature selection method applied to this dataset has generated better classification accuracy than GDA-WSVM combination, with an amelioration of 1.63%; of accuracy for improved F-score-MLP combination (93.49%).

## 5. CONCLUSIONS

In recent years, machine learning is becoming very important computer science fields. It is slowly changing many industry including health industry. We collect many data and we have very fast computational power. Together with very effective algorithms, we can develop tools to help human to do their job better. In this study we develop a tool to help doctor for the disease diagnosis. We do not expect our tool to be used by a doctor as it is, but these kind of tool can be used by medical student to be a learning tool.

Thyroid disease can be tricky to diagnose because symptoms are easily confused with other illness condition. When thyroid disease is caught early, treatment can be given patients very effectively. In this study, diagnosing thyroid disease is aimed with a machine learning tool called

as MLTDD (machine learning tool for thyroid disease diagnosis). MLTDD could form a foresight diagnosis with 99.7% accuracy for thyroid diseases in the datasets we used.

In our opinion, we can develop some more diseases diagnosis tool to help medical students to learn the disease characteristics for testing purposes. We do not believe these tools will replace the doctor yet. But they can be a helper to doctor if we have further improvement at technology. However, students who is studying endocrinology for thyroid diseases can use this tool for testing their knowledge by comparing their predictions with MLTDD. We have to collect more data to training these algorithms. The more data means we need to have high performing computational units. We believe that when we put big data, high speed computational power and advance machine learning algorithm, we can have very effective diagnostic tools.

## REFERENCES

- [1] Dr. Sahni BS, Thyroid Disorders .Online Available:  
<http://www.homoeopathyclinic.com/articles/diseases/tyroid.pdf>
- [2] Thyroid: <https://en.wikipedia.org/wiki/Thyroid>
- [3] Jiawei Han , Micheline Kamber, Data Mining Concepts and Techniques. Published by Elsevier 2006.
- [4] V. Podgorelec, P. Kokol, B. Stiglic, and I. Rozman, Decision trees: An overview and their use in medicine. In Proceedings of Journal Medical System 2002.
- [5] <http://archive.ics.uci.edu/ml/machine-learning-databases/thyroid-disease/>
- [6] H.S.Hota, Diagnosis of Breast Cancer Using Intelligent Techniques, International Journal of Emerging Science and Engineering (IJESE), January 2013 .
- [7] Jiawei Han , Micheline Kamber, Data Mining Concepts and Techniques. Published by Elsevier 2006.
- [8] Anupam Shukla, Prabhdeep Kaur, Ritu Tiwari and R.R. Janghel, Diagnosis of Thyroid disease using Artificial Neural Network. In Proceedings of IEEE IACC 2009.
- [9] Lale Ozyilmaz , Tulay Yildirim, Diagnosis of Thyroid disease using Artificial Neural Network Methods. In Proceedings of ICONIP 2002.
- [10] Fatemeh Saiti and Mahdi Aliyari, Thyroid Disease Diagnosis based on Genetic algorithms using PNN and SVM. In Proceedings of IEEE 2009.
- [11] Carlos Ordonez University of Houston, Houston, TX, “Comparing association rules and decision trees for disease prediction”. In Proceedings of Int. Conf. Inf. Knowl. Manage., 2006.
- [12] A.S.Varde,K.L.Massey and H.C.Wood, “A Clinical Laboratory Expert System for the Diagnosis of Thyroid Disfunction”, Proceeding of Annual International Conference of the IEEE Engineering in Medicine and Biology Society.
- [13] Palanichamy Jaganathan and Nallamuthu Rajkumar, ” An expert system for optimizing thyroid disease diagnosis”. In Proceedings of International Journal of Computational Science and Engineering,2012.

*INTENTIONAL BLANK*

# WEB SERVICE COMPOSITION IN DYNAMIC ENVIRONMENT: A COMPARATIVE STUDY

Aram AlSedrani<sup>1</sup> and Ameer Touir<sup>2</sup>

<sup>1</sup>Department of Computer Science, King Saud University, Riyadh, KSA  
Aram\_sedrani@ccis.imamu.edu.sa

<sup>2</sup>Department of Computer Science, King Saud University, Riyadh, KSA  
tuir@ksu.edu.sa

## ABSTRACT

*Web service composition development is a complex and dynamic process. It is one of the challenges in distributed dynamic environments. Although, SOA (Service Oriented Architecture) facilitates service composition process through standard protocols in searching and binding with web services. Yet composition in SOA paradigm faces many challenges. One of the main challenges is the environment in which composition services are developed. Nowadays the environment becomes more dynamic due to the increase in the number of web services that are frequently changing. Therefore, the need for self-adapted composition methods that acts according to environment changes is advocated. In this paper, we will study the existed researches that address the web service composition in a dynamic environment to state the art in this area and assist future research.*

## KEYWORDS

*Service composition, web services, dynamic environment.*

## 1. INTRODUCTION

The fast growth in the utilization of SOA led to the growth of the number of web services published on the net. Web services offer huge potentials through a large number of simple services developed by several service providers in various servers [1]. However, today's world has become not only more complex but also more dynamic. The single service offers simple and basic functionality that became inadequate to satisfy the needs for future requirement. Therefore combining multiple services to provide complex composite service is of a high demand nowadays.

Web service composition under dynamic environment gained researchers attention recently since around 2004 compared to the emergence of SOA and web services since the late 1990s [2]. This recent interest in this issue arises because of the significant number of available web services and the diversity of web service providers. Due to those reasons, the environment of web services becomes more dynamic with a great probability of web service added, deleted or even moved.

Therefore, service composition that depends on dynamic entities must adapt to any changes occurred during its phases.

In this paper, we aim at investigating the existed researches that deal with the problem of web service composition in a dynamic environment in order to identify the research gap of this field and improve future research. The remainder of the paper is organized as follow: in section two, an introduction of web service composition is provided. Section three will highlight the main challenges that web service composition faces in a dynamic environment. We conduct our study in section four, and the dissection of study results is provided in section five. In the last section, we present the conclusions of our study.

## 2. WEB SERVICE COMPOSITION

The services composition process consists of several phases. The first is the *composition planning* where specifying service request and decompose it into an abstract set of tasks. During the next phase, *service discovery*, a search for services that match the functionality and non-functionality requirements for each task in the composition is performed. Next, from the multiple services discovered in the previous phase, *service selection* is about selecting the most appropriate service for each task in the composition to satisfy user requirement. The last phase is the *service execution* where the individual task in the composition is invoked and executed to come up with the final service.

## 3. SERVICE COMPOSITION CHALLENGES IN DYNAMIC ENVIRONMENT

Due to the success of SOA, the development of web services gains much attention from the software developer. This success leads to the growth in the number of web services available over the internet and thus increases the dynamic nature of web services environment. Therefore, service composition that depends on the dynamic environment of web services demands dynamic approaches to ensure the reliability of the composition process.

To illustrate the dynamism of composition environment we need to understand the possible changes that could occur in such environment. The dynamism is due to two types of changes; one is caused by the web service, and the other is caused by the environment itself. For example, some of the changes in web services that could affect the composition are as follow: First, one or more of the participant services fail to accomplish its task, or they become unavailable. Second, participant services could not provide its expected QoS or updated their values. Third, new candidate services advertise for better QoS than participant service. On the other hand, some major faults that could affect the execution of web service composition caused by the environment like, network or connection failure during service invocation or unexpected server crash.

Therefore, service composition under dynamic environment needs new requirements. According to [3] we can set the new composition requirement as follow: self-configuring, self-optimizing, self-healing and self-adapting. Self-configuring composition indicates that the composition able to discover and select services automatically. Self-optimizing means that the services within the composition are chosen according to QoS constraint. The self-healing composition is the composition that is capable of detecting constraints violation and react accordingly. The last

requirement is the self-adapting composition that is capable of altering its process to adapt to changes occurs with minimum human intervention.

## 4. THE COMPARATIVE STUDY

Dynamic web service composition had researchers attention in recent years. Many solutions are proposed to handle the dynamic environment changes during composition process. In this study, we will investigate some recent researches in this field and conduct a comparative study about them. Later we will review our observations and recommendations.

### 4.1. Overview of Major Researches:

**1.. (Chang et.al 2004)** [4] is a middleware platform for service composition. The main goal of this work is to provide selection algorithms to maximize the user satisfaction through QoS. The platform considers the changing in QoS attribute values during Execution and proposed a re-planning technique for contingencies. The plan, in the case of QoS changes, is partitioned into three regions based on the criteria if the task has executed, currently executing, or will be executed. Then, based on a particular set of constraints the planner will identify the task (T) affected and perform a re-planning starting of the region that contains (T).

**2.. (Canfora et. al 2005)** [5] is a triggering algorithm to control the workflow of service composition in the case of QoS values changes during execution. The algorithm in this work basically will estimate the global QoS value of the workflow whenever an update occurs. However, the planning in this algorithm is static designed by the user.

**3.. (Nariai et. al 2005)** [6] the proposal in this research uses the situation calculus (SC)theory, in the form of first-order logic language, to enable instant reaction of user preferences changes during composition and planning. The authors incorporate the SC theory in an intelligent infrastructure to analyze the pattern of exceptions experienced by a particular machine to manage the user requirements. In fact, they discussed five exception patterns such as unavailability exception and un-satisfy user constraints exception. However, the work considers mainly the user preferences as the dynamic change and doesn't take into account the QoS changes during planning and composition. Moreover, the SC theory does not support the constructing of complex composed service.

**4.. (Sapena et. al 2008)** [7] this research proposes an on-line planning algorithm based on overlapping the planning and execution phases. The algorithm mainly developed for the systems under time constraints. Therefore, a temporary initial plan is constructed first and then whenever there is available time an improvement of this plan is performed. The algorithm follows a state space search method, specifically, a depth-first search to improve the initial plan. However, a conflict repairing mechanism is used in case if some action fails to complete the plan. This mechanism based on repairing the preconditions of affected tasks to override the failed action and find alternative tasks.

**5.. (Yi-an et. al 2009)** [8] this research proposes the NVDSCM (iNcomplete Visible and Dynamic Service Composition Method) algorithm to compose services in the partial-known and vibrant environment. The main concept of NVDSCM algorithm is to monitor the cost changes between neighbor services, and if changes happen to this cost, an adjustment to the following

costs will be performed until the end of the composition path. The algorithm utilizes the D\* search for finding composition path.

**6.. (Ying et. al 2009)** [9] the research utilizes the ECA rules in multimedia conference systems to manages web service composition in the case of updating user requirement. The proper event will trigger when the business process request is changed and allow service rescheduling. However, the proposed composition is semi-automated and focuses on the changes made by the user for updating requirement. Moreover, the work does not consider the changes in QoS of participant services.

**7.. (Dai et.al 2009)** [10] the authors in this research present a self-healing service composition framework based on performance prediction. The framework contains a QoS monitor that collects the quality values for participating services. At runtime, when a service is predicted to be failed according to its quality values, a reselection algorithm will be triggered to replace the failed service before it invoked. The aim of the proposed framework is to shorten the delay time from service failure recovery by early predicting the failures.

**8.. (Wiesner et.al 2009)** [11] in this research a recovery mechanisms are proposed for OWL-S semantic services in case of service failure. These mechanisms are based on OWL-S file annotation. By extracting the semantic information of services, the research proposes a set of operations to find alternative services in order to compensate the failed ones.

**9.. (Friedrich et.al 2010)** [12] The aim of this research is to provide a model-based approach and tools to analyze the reparability of service-based processes. The approach uses the forward reparability by creating repair plans to reach a correct execution state in case of fault or failure processes. In fact, the approach is divided into two steps. First, at design time, the heuristic-based analysis will check whether the plan is applicable for repairable executions of its processes. Second, at runtime, the approach will generate the repair plan based on the information from the previous step and predefined actions.

**10.. (Bartalos et.al 2011)** [13] the algorithm proposed in this research starts first by pruning the search space of web services by defining them as useable and Unusable. Then, the algorithm searches for the composition path using Fast Forward search. After adding a single web service to the composition plan, the algorithm entered a waiting mode to check if any changes occurred to web services. If so, it will react to the change according to three situations; if service deleted, the algorithm will delete the service, and its connections from the repository; if services added, the algorithm will classify it as useable or Unusable. If QoS values are updated, the algorithm will change accordingly. However, this research focuses on preparing and classifying the web services before composition to fasten the search process.

**11.. (Kuzu et.al 2012)** [14] the research proposes an algorithm for anytime planning that is based on overlapping between planning and execution. The tasks in anytime planning are executed ones it entered the plan. The execution outcomes of a single task are used to enhance the remaining planning process and provide information about services that could improve planning in non-deterministic situations. This approach is motivated by the fact that a web service repository at planning time could not be the same at execution time, so an overlapping between planning an execution of tasks are proposed. However in the case of contingency situation at some point during planning and effect previous executed task the re-planning will be costly since the



approach needs to plan and execute from the start point. Another problem in anytime planning is the risk of planning and executing of a single task one by one without considering a situation where the plan could be unachievable.

**12.. (Barakat et.al 2012)** [15] the work proposes an algorithm that reacts to changes occurred in the participating web services during the selection phase. The algorithm modeled the problem as a graph, and each node will store the optimal path from started node. Then when a change occurred the algorithm will re-track the stored information in each node in the path and reselect a new service. The proposed algorithm will adapt to a dynamic environment. However, with the increase in the number of nodes, the storage capacity will increase.

**13.. (Sabooohi et.al 2013)** [16] the SRPFR algorithm proposed in this work is an automatic sub-digraph renovation plan for failure recovery. It consists of two steps: an offline and an online process. First, the offline process begins when composite services are stored in a repository; the algorithm calculates all possible sub-digraph of a composite web service digraph. Then, it finds an appropriate replacement sub-digraph based on ranking mechanism. On the other hand, the online process starts when a service failure occurred in the execution time. In this case, the executor exchanges the sub-digraph that contains the failed service with a pre-calculated sub-digraph from the previous step.

**14.. (Markou et.al 2014)** [17] the MAPPPA algorithm proposed in this research is an alternative plan generation and merging algorithm. It aims basically at generating multiple plans for the same problem and merges these plans in a single decision tree. The algorithm is designed mainly for non-deterministic problems and could be applied to unexpected situations. The algorithm used A\* search over the decision tree and guided by probability values for successful execution of specific branches.

**15.. (Gupta et.al 2014)** [18] in this research the authors propose a mechanism for fault recovery during service composition execution. The mechanism is performed by a broker to generate all possible subsets of the composition sets if some service in the set is failed. Then, it is the responsibility of the broker to rank the subsets according to QoS values and replace the fault service.

**16.. (Wang et.al 2016)** [19] this research presents a framework for service composition based on multi-agent re-enforcement learning. The composition goal, in this framework, is decomposed among the agents to reduce the computational cost. Then, each agent will conduct a learning process to search for a solution to its goal. Furthermore, the research introduces a sharing algorithm between agents in order to reduce the time of convergence. From the experiences of agents, the framework able to adapt to changes occurred in a dynamic environment. However, the framework does not consider the quality of web services as a dynamic factor in the environment.

#### **4.2. Study Evaluation Factors:**

The next table illustrates the analytical study performed on these selected researches. In this study, we will overview the main technique used by the research to adapt to environment changes in order to evaluate the similarity of those approaches. Moreover, the main factor of this study is the composition phase that the research considers in detecting and reacting to dynamic changes. Moreover, with the importance of QoS in determining the quality of composite service, we will

analyze the main quality attributes that affect the composition and whether the researches consider them as a primary factor of composition or not. Then the main research outcomes will be reviewed to highlight the benefits of those researches. The last factor in our analysis is whether the research presents an experiment to proof the correctness of its technique or not.

Table 1: Comparative Table

Number	Research	Technique used	Dynamic Considered phase	QoS Considered attributes*	Research outcomes	Correctness proof
1	[4] Chang et.al 2004	Constraints set	Execution	EC, ET, REL, AV, REP	AgFlow platform	NO
2	[5] Canfora et. al 2005	Triggering Algorithm	Execution	ET, AV	Simulation	YES
3	[6] Nariai et. al 2005	Situation calculus	Execution	NON	Infrastructure	NO
4	[7] Sapena et..al 2008	Conflict-repairing	Planning Execution	NON	Algorithm	YES
5	[8] Yi-an et. al 2009	D* Search	Planning	EC	Algorithm	YES
6	[9] Ying et. al 2009	ECA Rules	Execution	NON	Algorithm	NO
7	[10] Dai et.al 2009	Performance prediction	Execution	EC, ET	Framework	YES
8	[11] Wiesner et.al 2009	Set of operations	Execution	NON	Algorithm	NO
9	[12] Friedrich et.al 2010	Model-based	Execution	NON	WS-Diamond Architecture	YES
10	[13] Bartalos et.al 2011	FF Search	Planning	NM	Algorithm	YES
11	[14] Kuzu et.al 2012	Triggering Signals	Execution	NON	Simplanner Framework	NO
12	[15] Barakat et.al 2012	Graph-based	Selection	EC, ET	Algorithm	YES
13	[16] Saboohi et.al 2013	Subdigraph replacement	Execution	EC, ET, REL	Algorithm	YES
14	[17] Markou et.al 2014	Alternative Plan	Planning	NON	Algorithm	YES
15	[18] Gupta et.al 2014	Subset replacement	Execution	NM	Broker	YES
16	[19] Wang et.al 2016	Agent Re-enforcement Learning	Execution	NON	Simulation	YES

\*EC: Execution Cost, ET: Execution Time, REL: Reliability, AV: Availability, REP: Reputation, NM: Not-Mentioned.

## 5. DISCUSSION

We can see from the number of researches on web service under a dynamic environment that this area is a new research direction, and is not mature enough to satisfy composition process requirements. The need for new techniques is highly recommended; especially with the fast grow of the web, which is the main environment where web service composition is operated. By conducting our study we identify several issues in this area.

As we can see from the next chart, which demonstrates the composition phase considered by subjected researches, that most existed service composition methods (75%) handles the dynamism of the environment in the execution phase and treats the environment as static in early phases. In such situations, any change will be detected and handled after their occurrence. This mechanism causes undesirable problems at execution time that leads to decrease the composition performance with re-planning or re-selecting overhead. However, some researches consider planning mechanisms to handle changes in web services. These approaches detect changes at early phases which build a more reliable composition plan. Thus, prevent re-planning overhead in case of changes during later phases. Most of the planning mechanisms proposed in subjected researches are focused on searching for alternative paths to overcome the changes of web services in original plan path.

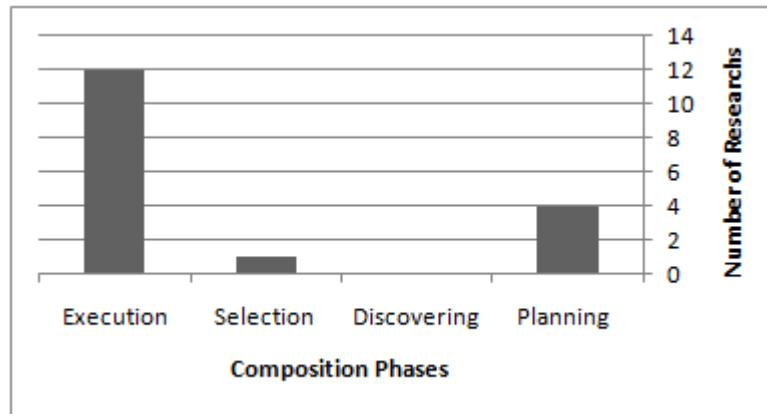


Figure 1: Composition Phases Chart

We can see also from the chart in figure 1 that almost none of the researches except [15] consider proposing a solution to dynamic changes in discovery and selection phases. Those two phases, in fact, deal directly with available web services while planning is considered as ordering tasks to achieve a particular goal. In the case of the discovery phase, the composition process will search for all possible web services to do a specific task in the plan. However, with dynamic web environment, the discovery phase outcomes must be a set of active, and reliable services that afterward the composition process can select from. Therefore, discovery process must depend on algorithms that can react to changes occurred on services, either when deleting inactive discovered services incidence or in the case of some newly exist equivalent service that must be included in the discovered services set.

The same concept is applied to the selection phase. First of all, the set of candidate web services to be select from must be reliable and active. Then, during selection, the changes on web services

must be considered in order to improve the selection. As an example, if new service advertises for QoS values that could improve the quality of overall composite service, then the selection algorithm must consider this service and add it to the candidate services set.

In another hand, 50% of the subjected researches, as shown in figure 2, consider the QoS attributes as the primary factor in the dynamic environment. QoS attributes are an effective factor in the dynamic environment. Increasing or decreasing the quality values of individual web service in the composition will affect the overall quality of composite service. We can observe from Table 1, that the most used QoS attributes in these researches are the execution cost (EC) and the execution time (ET). We can refer to the reason for these choices, is that most of the researches aim at enhancing the performance of the composition by reducing the time and cost. Since the composition process is a long process and dealing with multiple providers, the execution time and cost could be high in some cases.

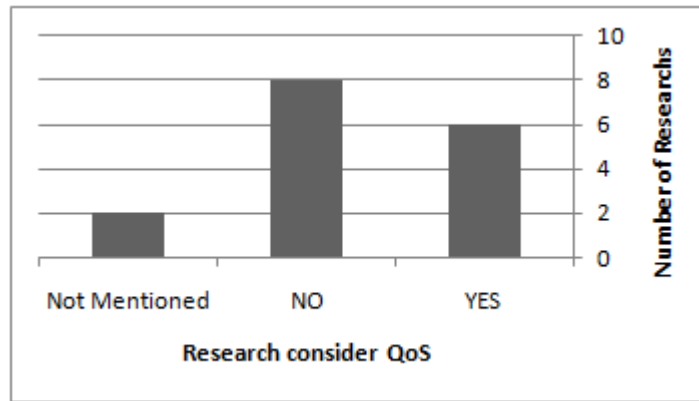


Figure 2: QoS chart

Another observation from the study is that there are two main methods to handle the changes in the composition. First, a pre-processing method in which all possible situations are predefined before composition process. As an example, Saboohi et al. proposed a method to identify all the potential sub-graphs and their alternative to replace a service failure before the execution of the composite web service. However, identifying all the sub-graphs and their alternatives are time-consuming, even if failure does not occur. In another hand, a re-planning and re-selection process are performed in the case of service failure during execution as in [9]. However, this method is inefficient due to the re-processing overhead. To be in the middle, we can recommend for methods that prepare for any changes could be occurred while selecting and planning at the first time, so in the case of service failure, the re-selection and re-planning will be in linear time. As an example, when selecting a service, we can identify some possible alternative services that equal in functional and non-functional properties. Therefore, in the case of service failed during selection or execution prepared solutions are replaced directly, according to its similarity ranking, into the composition.

The last comment on our study findings is that some researches conduct an experiment on their algorithms to proof the enhancement on execution time instead of proofing that the composition service is impeccable, and the dependencies between services after altering the composition are resolved.

## 6. CONCLUSION

In this paper, a comparative study is conducted that targets the web service composition methods in a dynamic environment. To the best of our knowledge, the selected researches are the only ones that have been found in this area. From our study we can summarize our observations and recommendations in the field of web service composition under dynamic environment as follow:

1. The need for composition methods that is capable of monitoring and reacting to any environment changes during all the composition phases.
2. An enhancement of QoS-aware web service composition methods is recommended due to the importance of quality values changes in the dynamic environment.
3. During each composition phase, preparation mechanisms that define alternative solutions to replace any service failure in linear time are important.
4. Any proposed solution must guarantee the correctness of composition after altering it.

## ACKNOWLEDGMENTS

This research has been supported by the Computer Science Dept. of the College of Computer and Information Sciences at King Saud University.

## REFERENCES

- [1] A. Jula, E. Sundararajan, and Z. Othman, "Cloud computing service composition: A systematic literature review," *Expert Syst. Appl.*, vol. 41, no. 8, pp. 3809–3824, Jun. 2014.
- [2] T. Erl, *Service-Oriented Architecture: Concepts, Technology, and Design*, 7th Ed. Indiana: Prentice Hall, 2008.
- [3] Q. Z. Sheng, X. Qiao, A. V Vasilakos, C. Szabo, S. Bourne, and X. Xu, "Web services composition : A decade ' s overview," *Inf. Sci.*, vol. 280, pp. 218–238, 2014.
- [4] L. Zeng, B. Benatallah, A. H. H. Ngu, M. Dumas, J. Kalagnanam, and H. Chang, "QoS-aware middleware for Web services composition," *IEEE Trans. Softw. Eng.*, vol. 30, no. 5, pp. 311–327, May 2004.
- [5] G. Canfora, M. Di Penta, R. Esposito, and M. L. Villani, "QoS-Aware Replanning of Composite Web Services," in *IEEE International Conference on Web Services (ICWS'05)*, 2005, pp. 121–129.
- [6] K. Nariai, I. Paik, and M. Shinozawa, "Planning and composition of Web services with dynamic constraints using situation calculus," in *Proceedings of The Fifth International Conference on Computer and Information Technology (CIT'05)*, 2005, pp. 1009–1013.
- [7] O. Sapena and E. Onaindía, "Planning in highly dynamic environments: An anytime approach for planning under time constraints," *Appl. Intell.*, vol. 29, no. 1, pp. 90–109, 2008.

- [8] Z. Yi-an and L. Wan-bao, "The Services Composition Algorithm in Incomplete Visible and Dynamic Environment," in WCSE '09. Second International Workshop on Computer Science and Engineering, 2009, vol. 1, pp. 169–171.
- [9] Z. Ying, C. Junliang, C. Bo, and Z. Yang, "Using ECA rules to manage web service composition for multimedia conference system," in 2nd IEEE International Conference on Broadband Network & Multimedia Technology, 2009, pp. 545–549.
- [10] Y. Dai, L. Yang, and B. Zhang, "QoS-Driven Self-Healing Web Service Composition Based on  $\pi$ ," J. Comput. Sci. Technol., vol. 24, no. 2, pp. 250–261, 2009.
- [11] K. Wiesner, R. Vacul, M. Kollingbaum, and K. Sycara, "Recovery Mechanisms for Semantic Web Services," in International Conference on Distributed applications and interoperable systems (DAIS), 2009, pp. 100–105.
- [12] G. Friedrich, M. Fugini, E. Mussi, B. Pernici, and G. Tagni, "Exception Handling for Repair in Service-Based Processes," in IEEE Transactions On Software Engineering, 2010, vol. 36, no. 2, pp. 198–216.
- [13] P. Bartalos and M. Bielíková, "Effective QoS-Aware Web Service Composition in Dynamic Environment," in Information Systems Development, 1st ed., W. W. Song, S. Xu, C. Wan, Y. Zhong, W. Wojtkowski, G. Wojtkowski, and H. Linger, Eds. Springer New York, 2011, pp. 101–113.
- [14] M. Kuzu and N. K. Cicekli, "Dynamic planning approach to automated web service composition," Appl. Intell., vol. 36, no. 1, pp. 1–28, 2012.
- [15] L. Barakat, S. Miles, and M. Luck, "Reactive Service Selection in Dynamic Service Environments," in First European Conference in Service-Oriented and Cloud Computing, 2012, pp. 17–31.
- [16] H. Saboohi and S. Abdul Kareem, "An automatic subdigraph renovation plan for failure recovery of composite semantic Web services," Front. Comput. Sci., vol. 7, no. 6, pp. 894–913, 2013.
- [17] G. Markou and I. Refanidis, "Anytime Planning for Web Service Composition via Alternative Plan Merging," in IEEE 26th International Conference on Tools with Artificial Intelligence (ICTAI), 2014, pp. 91 – 98.
- [18] S. Gupta and P. Bhanodia, "A Flexible and Dynamic Failure Recovery Mechanism for Composite Web Services Using Subset Replacement," Int. J. Sci. Res., vol. 3, no. 12, pp. 1886–1890, 2014.
- [19] H. Wang, X. Wang, X. Zhang, Q. Yu, and X. Hu, "Effective service composition using multi-agent reinforcement learning," Knowledge-Based Syst., vol. 92, pp. 151–168, 2016.

# FIRST-ORDER MATHEMATICAL FUZZY LOGIC WITH HEDGES

Van-Hung Le

Faculty of Information Technology,  
Hanoi University of Mining and Geology, Vietnam  
levanhung@humg.edu.vn

## ABSTRACT

*In this paper, we consider first-order mathematical fuzzy logic expanded by many hedges. This is based on the fact that, in the real world, many hedges can be used simultaneously, and some hedge modifies truth (or meaning of sentences) more than another hedge. Moreover, each hedge may or may not have a dual one. We expand two axiomatizations for propositional mathematical fuzzy logic with many hedges to the first-order level and prove a number of completeness results for the resulting logics. We also consider logics with many hedges based on  $\Delta$ -core fuzzy logics.*

## KEYWORDS

*Mathematical Fuzzy Logic, First-Order Logic, Hedge, Strong Completeness, Standard Completeness*

## 1. INTRODUCTION

Extending logical systems of mathematical fuzzy logic (MFL) with hedges is axiomatized by Hájek [1], Vychodil [2], Esteva *et al.* [3], among others. Hedges are called *truth-stressing* or *truth-depressing* if they, respectively, strengthen or weaken the meaning of the applied proposition. Intuitively, on a chain of truth values, the truth function of a truth-depressing (resp., truth-stressing) hedge (connective) is a superdiagonal (resp., subdiagonal) non-decreasing function preserving 0 and 1. In [1, 2, 3], logical systems of MFL are extended by a truth-stressing hedge and/or a truth-depressing one.

However, in the real world, we often use many hedges, e.g., *very*, *highly*, *rather*, and *slightly*, at the same time to express different levels of emphasis. Furthermore, a hedge may or may not have a dual one, e.g., *slightly* (resp., *rather*) can be seen as a dual hedge of *very* (resp., *highly*). Therefore, in [4, 5], Le *et al.* propose two axiomatizations for propositional logical systems of MFL with many hedges. In the axiomatization in [5], each hedge does not have any dual one whereas in the axiomatization in [4], each hedge can have its own dual one. In [5, 6], logical systems with many hedges for representing and reasoning with linguistically-expressed human knowledge are also proposed.



Moreover, since first-order logic is more powerful than propositional logic in terms of knowledge representation and reasoning. In this paper, we expand the axiomatizations in [4, 5] to the first-order level and prove a number of completeness results for the resulting logics w.r.t. the underlying logic. More concretely, we propose first-order fuzzy logics with many hedges based on a first-order core fuzzy logic for two cases: (i) each hedge does not have a dual one, and (ii) each hedge can have its own dual one. Then, we prove that the new first-order logics are a conservative expansion of the underlying first-order core fuzzy logic and gives a characterization of their strong completeness, especially, the strong standard completeness. We also discuss logics with many hedges based on  $\Delta$ -core fuzzy logics, which are extensions of core fuzzy logics by  $\Delta$  connective.

The remainder of the paper is organized as follows. Section 2 gives an overview of notions and results of MFL used in this paper and the two axiomatizations for propositional MFL with many hedges proposed in [4, 5]. Section 3 presents first-order MFL with many hedges and proves a number of completeness results. Section 4 discusses logics with many hedges based on  $\Delta$ -core fuzzy logics. Section 5 concludes the paper.

## 2. PRELIMINARIES

### 2.1. Preliminaries on Mathematical Fuzzy Logic

Let  $L$  be a propositional logic in a language  $\mathcal{L}$ , a set of connectives with finite arity. A truth constant  $\bar{r}$  is a special formula whose truth value under every evaluation is  $r$ . Formulae are built from variables and truth constants using connectives in  $\mathcal{L}$ . Each evaluation  $e$  of variables by truth values uniquely extends to an evaluation  $e(\varphi)$  of all formulae  $\varphi$  using truth functions of connectives. A formula  $\varphi$  is called a *1-tautology* if  $e(\varphi) = 1$  for all evaluations  $e$ . Several 1-tautology formulae are taken as *axioms* of the logic. A *theory* is a set of formulae. An evaluation  $e$  is called a *model* of a theory  $T$  if  $e(\varphi) = 1, \forall \varphi \in T$ . A *proof* in  $T$  is a sequence  $\varphi_1, \dots, \varphi_n$  of formulae whose each member is either an axiom of the logic or a member of  $T$  or follows from some preceding members of the sequence using the deduction rule(s) of the logic. If  $\varphi$  is the last member of a proof in  $T$ ,  $\varphi$  is called a *provable* formula, denoted  $T \vdash_L \varphi$ . If  $T = \emptyset$ ,  $\varphi$  is said to be *provable* in the logic [7, 8].

Most logical systems called *core fuzzy logics* are a finitary Rasiowa-implicative logic [8]. Every finitary Rasiowa-implicative logic  $L$  is algebraizable. Its equivalent algebraic semantics, a class of  $L$ -algebras, is a quasivariety. Each  $L$ -algebra  $\mathbf{A}$  is endowed with a relation  $\leq$  (called *preorder*) by setting,  $\forall a, b \in A, a \leq b$  iff  $a \Rightarrow b = 1$ , where  $\Rightarrow$  is the truth function of the connective  $\rightarrow$ . If  $\leq$  is a total order,  $\mathbf{A}$  is called an *L-chain*.  $L$  is called a *semilinear* logic iff it is strongly complete w.r.t. the class of  $L$ -chains [7, 8]. Core fuzzy logics are semilinear. They belong to a large class of systems which are axiomatic expansions of MTL (monoidal t-norm based logic). Other well-known examples of core fuzzy logics are BL (basic logic), G (Gödel logic), L (Łukasiewicz logic),  $\Pi$  (product logic) [7], SBL, NM, WNM, IMTL, and SMTL [8].



The  $\Delta$  connective is the one whose truth function is given by: if  $x = 1$ ,  $\Delta(x) = 1$ ; otherwise,  $\Delta(x) = 0$ . It is used to express the notion of full truth of a proposition which is not expressible in either MTL or other logics of left-continuous t-norms. A core fuzzy logic  $L$  expanded by this connective is called “ $L$  with  $\Delta$ ” and denoted by  $L_\Delta$ . It can be axiomatized by adding the following axioms to the axiomatic system for the logic  $L$  [8]:

- ( $\Delta 1$ )  $\Delta\varphi \vee \neg\Delta\varphi$
- ( $\Delta 2$ )  $\Delta(\varphi \vee \psi) \rightarrow \Delta\varphi \vee \Delta\psi$
- ( $\Delta 3$ )  $\Delta\varphi \rightarrow \varphi$
- ( $\Delta 4$ )  $\Delta\varphi \rightarrow \Delta\Delta\varphi$
- ( $\Delta 5$ )  $\Delta(\varphi \rightarrow \psi) \rightarrow (\Delta\varphi \rightarrow \Delta\psi)$

and the rule of  $\Delta$ -necessitation: from  $\varphi$  infer  $\Delta\varphi$ .

Besides the notion of full truth, the ordering of truth values is also internalized in  $L_\Delta$ , as  $\Delta(x \rightarrow y) = 1$  iff  $x \leq y$ . Moreover, the connective  $\Delta$  makes it possible to interpret classical logic in  $L_\Delta$  by prefixing each propositional variable by  $\Delta$ . Extensions of all core fuzzy logics by  $\Delta$  are called  $\Delta$ -core fuzzy logics, e.g.,  $MTL_\Delta$ .

**Definition 1** (**KC**, **FSKC**, **SKC**) [8] *Let  $L$  be a core fuzzy logic and  $\mathbb{K}$  a class of  $L$ -chains. It is said that  $L$  has the (finitely) strong  $\mathbb{K}$ -completeness, (F)SKC for short, if for every (finite) set of formulae  $T$  and every formula  $\varphi$ , it holds that  $T \vdash_L \varphi$  iff  $e(\varphi) = 1$  for every  $L$ -algebra  $\mathbf{A} \in \mathbb{K}$  and each  $\mathbf{A}$ -model  $e$  of  $T$ . It is said that  $L$  has the  $\mathbb{K}$ -completeness, KC for short, when the equivalence is true for  $T = \emptyset$ .*

Clearly, the SKC implies the FSKC, and the FSKC implies the KC. When  $\mathbb{K}$  is the class of all chains whose support is the unit interval  $[0, 1]$  with the usual ordering, the (F)SKC can be called the (finite) strong *standard* completeness, (F)SSC for short.

## 2.2. An Axiomatization for Many Hedges

A hedge may modify truth more than another [9, 10, 11]. For example, *slightly* (resp., *very*) modifies truth more than *rather* (resp., *highly*) since *slightly true* < *rather true* < *true* (resp., *true* < *highly true* < *very true*). To ease the presentation, let  $s_0, d_0$  denote the *identity* connective, i.e., for all formula  $\varphi$ ,  $\varphi \equiv s_0\varphi \equiv d_0\varphi$ , and their truth functions are the identity.

**Definition 2** [5] *Let  $L$  be a core fuzzy logic. A logic  $L_{s,d}^{p,q}$ , where  $p, q$  are positive integers, is an expansion of  $L$  with new unary connectives  $s_1, \dots, s_p$  (for truth-stressers) and  $d_1, \dots, d_q$  (for truth-depressers) by the following additional axioms, for  $i = 1, \dots, p$  and  $j = 1, \dots, q$ :*

- ( $S_i$ )  $s_i\varphi \rightarrow s_{i-1}\varphi$
- ( $S_{p+1}$ )  $s_p\overline{1}$
- ( $D_j$ )  $d_{j-1}\varphi \rightarrow d_j\varphi$
- ( $D_{q+1}$ )  $\neg d_q\overline{0}$

and the following additional deduction rule:

$(DR^h)$  from  $(\varphi \rightarrow \psi) \vee \chi$  infer  $(h\varphi \rightarrow h\psi) \vee \chi$ , for each  $h \in \{s_1, \dots, s_p, d_1, \dots, d_q\}$ .

Axiom  $(S_i)$  (resp., axiom  $(D_j)$ ) expresses that  $s_i$  (resp.,  $d_j$ ) modifies truth more than  $s_{i-1}$  (resp.,  $d_{j-1}$ ), for  $i = 2, \dots, p$  (resp.,  $j = 2, \dots, q$ ).

It is shown that  $L_{s,d}^{p,q}$  is a finitary Rasiowa-implicative logic [5], and its equivalent algebraic semantics is the class of  $L_{s,d}^{p,q}$ -algebras.

**Definition 3** [5] *An algebra  $A = \langle A, *, \Rightarrow, \cap, \cup, 0, 1, s_1, \dots, s_p, d_1, \dots, d_q \rangle$  of type  $\langle 2, 2, 2, 2, 0, 0, 1, \dots, 1 \rangle$  is an  $L_{s,d}^{p,q}$ -algebra if it is an  $L$ -algebra expanded by unary operators  $s_i, d_j : A \rightarrow A$  that satisfy, for all  $x, y, z \in A$ ,  $i = \overline{1, p}$  and  $j = \overline{1, q}$ ,*

$$s_i(x) \leq s_{i-1}(x) \quad (1)$$

$$s_p(1) = 1 \quad (2)$$

$$d_j(x) \geq d_{j-1}(x) \quad (3)$$

$$d_q(0) = 0 \quad (4)$$

$$\text{if } (x \Rightarrow y) \cup z = 1 \text{ then } (s_i^\bullet(x) \Rightarrow s_i^\bullet(y)) \cup z = 1 \quad (5)$$

$$\text{if } (x \Rightarrow y) \cup z = 1 \text{ then } (d_j^\bullet(x) \Rightarrow d_j^\bullet(y)) \cup z = 1 \quad (6)$$

In [5], it is proved that for all  $i = \overline{1, p}$ ,  $s_i$  is subdiagonal (i.e.,  $\forall x, s_i(x) \leq x$ ), and preserves 0 and 1; for all  $j = \overline{1, q}$ ,  $d_j$  is superdiagonal (i.e.,  $\forall x, d_j(x) \geq x$ ), and preserves 0 and 1. Moreover, in a chain of truth values, they are all non-decreasing.

**Theorem 1** [5] *Let  $L$  be a core fuzzy logic,  $\mathbb{K}$  a class of  $L$ -chains, and  $\mathbb{K}_{s,d}^{p,q}$  the class of the  $L_{s,d}^{p,q}$ -chains whose  $s_1, \dots, s_p, d_1, \dots, d_q$ -free reducts are in  $\mathbb{K}$ . Then: (i)  $L_{s,d}^{p,q}$  is a conservative expansion of  $L$ ; (ii)  $L_{s,d}^{p,q}$  is strongly complete w.r.t. the class of all  $L_{s,d}^{p,q}$ -chains, i.e.,  $L_{s,d}^{p,q}$  is semilinear; (iii)  $L$  has the FSSC, FS $\mathbb{K}C$ , SSC, and S $\mathbb{K}C$  iff  $L_{s,d}^{p,q}$  has the FSSC, FS $\mathbb{K}C$ , SSC, and S $\mathbb{K}C$ , respectively.*

### 2.3. Mathematical Fuzzy logic with Many Dual Hedges

It can be observed that each hedge can have a dual one, e.g., *slightly* and *rather* can be seen as a dual hedge of *very* and *highly*, respectively. Thus, there might be axioms expressing dual relations of hedges in addition to axioms expressing their comparative truth modification strength.

**Definition 4** [4] *Let  $L$  be a core fuzzy logic. A logic  $L_{s,d}^{2n}$ , where  $n$  is a positive integer, is an expansion of  $L$  with new unary connectives  $s_1, \dots, s_n$  (for truth-stressers) and  $d_1, \dots, d_n$  (for truth-depressers) by the following additional axioms, for  $i = 1, \dots, n$ :*

$$\begin{aligned}
(S_i^{dh}) \quad & s_i\varphi \rightarrow s_{i-1}\varphi \\
(S_{n+1}^{dh}) \quad & s_n\bar{1} \\
(D_i^{dh}) \quad & d_{i-1}\varphi \rightarrow d_i\varphi \\
(SD_i^{dh}) \quad & d_i\varphi \rightarrow \neg s_i\neg\varphi
\end{aligned}$$

and the following additional deduction rule:

$$(DR^{dh}) \text{ from } (\varphi \rightarrow \psi) \vee \chi \text{ infer } (h\varphi \rightarrow h\psi) \vee \chi, \text{ for } h \in \{s_1, \dots, s_n, d_1, \dots, d_n\}.$$

The logic  $L_{s,d}^{2n}$  is  $L$  expanded by  $2n$  hedges, where hedges are divided into pairs of dual ones. Axiom  $(SD_i)$  expresses the dual relation between hedges  $s_i$  and  $d_i$  and coincides with Axiom (ST2) in Vychodil's axiomatization. For the case of *very*, *slightly*, and  $\varphi = \textit{young}$ , it means “*slightly young* implies not *very old*”. It can be proved that  $(SD_i^{dh})$  is equivalent to  $s_i\varphi \rightarrow \neg d_i\neg\varphi$  which implies “*very young* implies not *slightly old*” as well.

$L_{s,d}^{2n}$  is also a finitary Rasiowa-implicative logic, and its equivalent algebraic semantics is the class of  $L_{s,d}^{2n}$ -algebras.

**Definition 5** [4] An algebra  $\mathbf{A} = \langle A, *, \Rightarrow, \cap, \cup, 0, 1, s_1, \dots, s_n, d_1, \dots, d_n \rangle$  of type  $\langle 2, 2, 2, 2, 0, 1, \dots, 1 \rangle$  is an  $L_{s,d}^{2n}$ -algebra if it is an  $L$ -algebra expanded by unary operators  $s_i, d_i : A \rightarrow A$  that satisfy, for all  $x, y, z \in A$  and  $i = 1, \dots, n$ ,

$$s_i(x) \leq s_{i-1}(x) \quad (7)$$

$$d_i(x) \geq d_{i-1}(x) \quad (8)$$

$$s_n(1) = 1 \quad (9)$$

$$d_i(x) \leq \neg s_i(\neg x) \quad (10)$$

$$\text{if } (x \Rightarrow y) \cup z = 1 \text{ then } (s_i(x) \Rightarrow s_i(y)) \cup z = 1 \quad (11)$$

$$\text{if } (x \Rightarrow y) \cup z = 1 \text{ then } (d_i(x) \Rightarrow d_i(y)) \cup z = 1 \quad (12)$$

**Theorem 2** [4] Let  $L$  be a core fuzzy logic,  $\mathbb{K}$  a class of  $L$ -chains, and  $\mathbb{K}_{s,d}^{2n}$  the class of the  $L_{s,d}^{2n}$ -chains whose  $s_1, \dots, s_n, d_1, \dots, d_n$ -free reducts are in  $\mathbb{K}$ . Then: (i)  $L_{s,d}^{2n}$  is a conservative expansion of  $L$ ; (ii)  $L_{s,d}^{2n}$  is strongly complete w.r.t. the class of all  $L_{s,d}^{2n}$ -chains, i.e.,  $L_{s,d}^{2n}$  is semilinear; (iii)  $L$  has the FSSC, FSKC, SSC, and SKC iff  $L_{s,d}^{2n}$  has the FSSC, FSKC, SSC, and SKC, respectively.

It can be seen that in a case when there is one truth-stressing (resp., truth-depressing) hedge without a dual one, we just add the axioms expressing its relations to the existing truth-stressing (resp., truth-depressing) hedges according to their comparative truth modification strength.

### 3. FIRST-ORDER MATHEMATICAL FUZZY LOGIC WITH HEDGES

We first recall the representation of first-order formalisms for core fuzzy logics [8]. Given a propositional core fuzzy logic  $L$ , the language  $\mathcal{PL}$  of the first-order core fuzzy logic  $L\forall$  is built from the propositional language  $\mathcal{L}$  of  $L$  by extending it with a non-empty set of predicate symbols  $Pred$ , a set of function symbols  $Func$  (disjoint with  $Pred$ ), a set of object variables  $Var$ , and two quantifiers  $\exists$  and  $\forall$ . The set of terms  $Term$  is the minimum set containing the elements of  $Var$  and closed under the functions. The atomic formulae are in the form of  $P(t^1, \dots, t^n)$ , where  $P \in Pred$  and  $t^1, \dots, t^n \in Term$ . The set of all formulae is obtained by closing the set of atomic formulae under combination by propositional connectives and quantification, i.e., if  $u$  is a formula and  $x$  is a variable, then  $(\forall x)u$  and  $(\exists x)u$  are formulae.

In first-order core fuzzy logics, it is usual to restrict the semantics to chains only. Given an  $L$ -chain  $\mathbf{A}$ , an  $\mathbf{A}$ -structure is  $\mathbf{M} = \langle M, (P_{\mathbf{M}})_{P \in Pred}, (f_{\mathbf{M}})_{f \in Func} \rangle$ , where  $M$  is a non-empty domain; for each  $n$ -ary predicate symbol  $P \in Pred$ ,  $P_{\mathbf{M}}$  is an  $n$ -ary fuzzy relation on  $M$ , i.e., a function  $M^n \rightarrow A$  (identified with an element of  $A$  if  $n = 0$ ); for each  $n$ -ary function symbol  $f \in Func$ ,  $f_{\mathbf{M}}$  is a function  $M^n \rightarrow M$  (identified with an element of  $M$  if  $n = 0$ ). An evaluation  $v$  of variables is a mapping  $v : Var \rightarrow M$ . Let  $x$  be a variable and  $a \in M$ . Then let  $v[x \rightarrow a]$  denote the evaluation such that  $v[x \rightarrow a](x) = a$  and  $v[x \rightarrow a](y) = v(y)$  for each variable  $y$  different from  $x$ . The values of terms and the truth values of formulae are defined inductively as follows:

$$\begin{aligned}
 \|x\|_{\mathbf{M},v}^{\mathbf{A}} &= v(x) \\
 \|f(t^1, \dots, t^n)\|_{\mathbf{M},v}^{\mathbf{A}} &= f_{\mathbf{M}}(\|t^1\|_{\mathbf{M},v}^{\mathbf{A}}, \dots, \|t^n\|_{\mathbf{M},v}^{\mathbf{A}}) && \text{for } f \in Func \\
 \|P(t^1, \dots, t^n)\|_{\mathbf{M},v}^{\mathbf{A}} &= P_{\mathbf{M}}(\|t^1\|_{\mathbf{M},v}^{\mathbf{A}}, \dots, \|t^n\|_{\mathbf{M},v}^{\mathbf{A}}) && \text{for } P \in Pred \\
 \|c(\varphi_1, \dots, \varphi_n)\|_{\mathbf{M},v}^{\mathbf{A}} &= c_{\mathbf{A}}(\|\varphi_1\|_{\mathbf{M},v}^{\mathbf{A}}, \dots, \|\varphi_n\|_{\mathbf{M},v}^{\mathbf{A}}) && \text{for } c \in \mathcal{L} \\
 \|(\forall x)\varphi\|_{\mathbf{M},v}^{\mathbf{A}} &= \inf\{\|\varphi\|_{\mathbf{M},v[x \rightarrow a]}^{\mathbf{A}} \mid a \in M\} \\
 \|(\exists x)\varphi\|_{\mathbf{M},v}^{\mathbf{A}} &= \sup\{\|\varphi\|_{\mathbf{M},v[x \rightarrow a]}^{\mathbf{A}} \mid a \in M\}
 \end{aligned}$$

if the infimum or supremum does not exist, the truth value of the quantified formula is undefined. The  $\mathbf{A}$ -structure  $\mathbf{M}$  is *safe* if  $\|\varphi\|_{\mathbf{M},v}^{\mathbf{A}}$  is defined for each formula  $\varphi$  and each evaluation  $v$ . Then the truth value of  $\varphi$  in a safe  $\mathbf{A}$ -structure  $\mathbf{M}$  is just:

$$\|\varphi\|_{\mathbf{M}}^{\mathbf{A}} = \inf\{\|\varphi\|_{\mathbf{M},v}^{\mathbf{A}} \mid v : Var \rightarrow M\}$$

For a safe  $\mathbf{A}$ -structure  $\mathbf{M}$ , if  $\|\varphi\|_{\mathbf{M}}^{\mathbf{A}} = 1$ , the pair  $\langle \mathbf{A}, \mathbf{M} \rangle$  is said to be a *model* of  $\varphi$ , denoted  $\langle \mathbf{A}, \mathbf{M} \rangle \models \varphi$ . A pair  $\langle \mathbf{A}, \mathbf{M} \rangle$  is called a *model* of a theory (a set of sentences)  $T$  if  $\langle \mathbf{A}, \mathbf{M} \rangle \models \varphi$  for each  $\varphi \in T$ .

The axioms for  $L\forall$  are obtained from those of  $L$  by substituting formulae of  $\mathcal{PL}$  for propositional variables and adding the following axioms for quantifiers:



- ( $\forall 1$ )       $(\forall x)\varphi(x) \rightarrow \varphi(t)$ ,      where the term  $t$  is substitutable for  $x$  in  $\varphi$   
 ( $\exists 1$ )       $\varphi(t) \rightarrow (\exists x)\varphi(x)$ ,      where  $t$  is substitutable for  $x$  in  $\varphi$   
 ( $\forall 2$ )     $(\forall x)(\psi \rightarrow \varphi) \rightarrow (\psi \rightarrow (\forall x)\varphi)$ ,    where  $x$  is not free in  $\psi$   
 ( $\exists 2$ )     $(\forall x)(\varphi \rightarrow \psi) \rightarrow ((\exists x)\varphi \rightarrow \psi)$ ,    where  $x$  is not free in  $\psi$   
 ( $\forall 3$ )     $(\forall x)(\psi \vee \varphi) \rightarrow (\psi \vee (\forall x)\varphi)$ ,    where  $x$  is not free in  $\psi$

The deduction rules of  $L\forall$  are modus ponens and *generalization*:

$$(Gen) \quad \text{From } \varphi \text{ infer } (\forall x)\varphi$$

The notions of proof and provability are defined in the usual way. The formula  $\varphi$  is provable in  $L\forall$  from a theory  $T$  is denoted by  $T \vdash_{L\forall} \varphi$ .

The syntax and semantics of first-order core fuzzy logics are bounded together by the following completeness theorem:

**Theorem 3** [12] *For any first-order core fuzzy logic  $L\forall$  with a predicate language  $\mathcal{PL}$ , any  $\mathcal{PL}$ -theory  $T$ , and any  $\mathcal{PL}$ -formula  $\varphi$ , the following are equivalent:*

- $T \vdash_{L\forall} \varphi$
- $\langle \mathbf{A}, \mathbf{M} \rangle \models \varphi$  for each model  $\langle \mathbf{A}, \mathbf{M} \rangle$  of  $T$  with  $\mathbf{A}$  being a countable  $L$ -chain.

The notions  $\text{SKC}$ ,  $\text{FSKC}$ , and  $\mathbb{K}\text{C}$  are defined similarly to the propositional case as follows:

**Definition 6** [8] *Let  $L\forall$  be a core first-order fuzzy logic and  $\mathbb{K}$  a class of  $L$ -chains.  $L\forall$  has the (finite) strong  $\mathbb{K}$ -completeness property, (F)SKC, if for every (finite) set of formulae  $T$  and every formula  $\varphi$ , it holds that  $T \vdash_{L\forall} \varphi$  iff  $\langle \mathbf{A}, \mathbf{M} \rangle \models \varphi$  for every  $L$ -algebra  $\mathbf{A}$  of  $\mathbb{K}$  and each model  $\langle \mathbf{A}, \mathbf{M} \rangle$  of  $T$ .  $L$  has the  $\mathbb{K}$ -completeness property,  $\mathbb{K}\text{C}$ , when the equivalence is true for  $T = \emptyset$ .*

When  $\mathbb{K}$  is the class of all chains whose support is the unit interval  $[0, 1]$  with the usual ordering, the (F)SKC can be called the (finite) strong *standard* completeness, (F)SSC.

Theorem 3 states that every first-order core fuzzy logic enjoys the SKC, where  $\mathbb{K}$  is the class of all countable chains.

First-order fuzzy logics with many hedges in which each hedge does not have any dual one can be defined as follows.

**Definition 7** *Given a first-order core fuzzy logic  $L\forall$ , let  $L_{s,d}^{p,q}\forall$  be the expansion of  $L\forall$  with new unary connectives  $s_1, \dots, s_p$  (for truth-stressers) and  $d_1, \dots, d_q$  (for truth-depressers), axioms  $(S_i)$ ,  $(S_{p+1})$ ,  $(D_j)$ ,  $(D_{q+1})$ , and the deduction rule  $(DR^h)$ .*

$L_{s,d}^{p,q}\forall$  can be obtained by expanding  $L_{s,d}^{p,q}$  to the first-order level in the similar way as above.

**Theorem 4** *The logic  $L_{s,d}^{p,q}\forall$  is a conservative expansion of the logic  $L\forall$ , i.e., for every set  $T \cup \varphi$  of  $\mathcal{PL}$ -formulae,  $T \vdash_{L_{s,d}^{p,q}\forall} \varphi$  iff  $T \vdash_{L\forall} \varphi$ , where  $\mathcal{PL}$  is the language of  $L\forall$ .*

*Proof.* It is trivial that if  $T \vdash_{L\forall} \varphi$  then  $T \vdash_{L_{s,d}^{p,q}\forall} \varphi$ . Now assume that  $T \not\vdash_{L\forall} \varphi$ . Then, there exists an L-chain  $\mathbf{A}$  and a model  $\langle \mathbf{A}, \mathbf{M} \rangle$  of  $T$  such that  $\langle \mathbf{A}, \mathbf{M} \rangle \not\models \varphi$ .  $\mathbf{A}$  can be expanded to an  $L_{s,d}^{p,q}$ -chain  $\mathbf{A}'$  by defining  $s_i^\bullet(1) = 1$  and  $s_i^\bullet(a) = 0$  for every  $a \in A \setminus \{1\}$  ( $i = \overline{1, p}$ ), and  $d_j^\bullet(0) = 0$  and  $d_j^\bullet(a) = 1$  for every  $a \in A \setminus \{0\}$  ( $j = \overline{1, q}$ ). It can be easily verified that the defined functions satisfy Conditions (1)-(6). Thus, in the expanded language of  $\mathcal{PL}$ ,  $\langle \mathbf{A}', \mathbf{M} \rangle$  is still safe and is a model of  $T$ , but  $\langle \mathbf{A}', \mathbf{M} \rangle \not\models \varphi$ , i.e.,  $T \not\vdash_{L_{s,d}^{p,q}\forall} \varphi$ .

Let  $\mathbf{A}, \mathbf{B}$  be two algebras of the same type with (defined) lattice operations  $\sup, \inf$ . It is said that an embedding  $f : \mathbf{A} \rightarrow \mathbf{B}$  is a  $\sigma$ -embedding if  $f(\sup C) = \sup f[C]$  (whenever  $\sup C$  exists) and  $f(\inf D) = \inf f[D]$  (whenever  $\inf D$  exists) for each countable  $C, D \subseteq A$ , i.e.,  $f$  preserves existing suprema and infima. As shown in [12], a usual way to prove SKC is to show that every non-trivial countable L-chain can be  $\sigma$ -embedded into some chain of  $\mathbb{K}$ . In this case, it is said that L has the  $\mathbb{K}$ - $\sigma$ -embedding property. As proved in [12], this is a *sufficient*, but in general *not necessary*, condition for the SKC. This method has been used to prove strong real completeness for first-order versions of a number of important core fuzzy logics, e.g., MTL, SMTL, IMTL, G, NM, and WNM.

The SSC of the logic  $L_{s,d}^{p,q}\forall$  is characterized in the following theorem.

**Theorem 5 (Strong standard completeness)** *Let  $L$  be a core fuzzy logic,  $\mathbb{K}$  the class of all standard L-chains, and  $\mathbb{K}_{s,d}^{p,q}$  the class of all standard  $L_{s,d}^{p,q}$ -chains whose  $s_1, \dots, s_p, d_1, \dots, d_q$ -free reducts are in  $\mathbb{K}$ . Then:*

(i) *If  $L$  has the  $\mathbb{K}$ - $\sigma$ -embedding property, then  $L_{s,d}^{p,q}\forall$  has the  $\text{SK}_{s,d}^{p,q}C$ , i.e., SSC.*

(ii) *If  $L$  does not have  $\mathbb{K}$ -embedding property, then  $L_{s,d}^{p,q}\forall$  does not have the  $\text{SK}_{s,d}^{p,q}C$ .*

*Proof.* For (i), it suffices to prove that any non-trivial countable  $L_{s,d}^{p,q}$ -chain  $\mathbf{A}$  can be  $\sigma$ -embedded into some standard chain of  $\mathbb{K}_{s,d}^{p,q}$ . Since L has the  $\mathbb{K}$ - $\sigma$ -embedding property, the  $s_1, \dots, s_p, d_1, \dots, d_q$ -free reduct of  $\mathbf{A}$  can be  $\sigma$ -embedded into a standard L-chain  $\mathbf{B} = \langle [0, 1], *, \Rightarrow, \cap, \cup, 0, 1 \rangle$  by a mapping  $f$  preserving existing suprema and infima. Since  $\mathbf{A}$  is countable, for each  $1 \leq k \leq p$ , we may arrange all points  $\{\langle f(x), f(s_k(x)) \rangle \mid x \in A\}$  into a sequence  $\{\langle f(x_i), f(s_k(x_i)) \rangle \mid x_i \in A, i = 1, 2, \dots\}$ , where  $0 = x_1 < x_2 < \dots$  and  $\lim_{i \rightarrow \infty} x_i = 1$ . Since  $f$  preserves the order relation, we have  $0 = f(x_1) < f(x_2) < \dots$  and  $\lim_{i \rightarrow \infty} f(x_i) = 1$ . Let  $s'_k : [0, 1] \rightarrow [0, 1]$  be the piecewise linear function connecting neighbored points from  $\{\langle f(x_i), f(s_k(x_i)) \rangle\}$ , i.e.,  $s'_k$  is continuous and  $s'_k(f(x)) = f(s_k(x))$ , and  $s'_k(1) = 1$ . Similarly, for each  $1 \leq l \leq q$ , let  $d'_l$  be the piecewise linear function connecting neighbored points from  $\{\langle f(x_i), f(d_l(x_i)) \rangle\}$ , i.e.,  $d'_l$  is continuous  $d'_l(f(x)) = f(d_l(x))$ , and  $d'_l(1) = 1$ . We will show that all  $s'_k$  and  $d'_l$  also satisfy Conditions (1)-(6). Since  $s_k$  and  $d_l$  preserve 0, clearly,  $s'_k$  and  $d'_l$  also preserve 0. By (1), we have  $s_k(x) \leq s_{k-1}(x)$ , for all  $1 \leq k \leq p$  (note that  $s_0(x) = x$ ). Thus,  $s'_k(f(x_i)) = f(s_k(x_i)) \leq f(s_{k-1}(x_i)) = s'_{k-1}(f(x_i))$  for all  $i$ . Since  $s'_k, s'_{k-1}$  are the piecewise linear functions connecting the corresponding points, we have  $s'_k(x) \leq s'_{k-1}(x)$  for all  $x \in [0, 1]$ . Similarly, we have  $d'_l(x) \geq d'_{l-1}(x)$  for all  $x \in [0, 1]$ . It remains to show that all  $s'_k$  and  $d'_l$  are non-decreasing. For all  $i \geq 1$ , since  $f$  and  $s_k$  are non-decreasing, we have  $f(x_i) < f(x_{i+1})$  and  $s'_k(f(x_i)) =$

$f(s_k(x_i)) \leq f(s_k(x_{i+1})) = s'_k(f(x_{i+1}))$ . Also, since  $s'_k$  linearly connects neighbored points from  $\{\langle f(x_i), s'_k(f(x_i)) \rangle | i = 1, 2, \dots\}$ ,  $s'_k$  is non-decreasing on  $[0, 1]$ . Similarly,  $d'_l$  is non-decreasing on  $[0, 1]$ . Hence,  $\mathbf{B}$  expanded by all  $s'_k$  and  $d'_l$  is a standard  $L_{s,d}^{p,q}$ -chain into which  $\mathbf{A}$  is  $\sigma$ -embedded.

For (ii), since  $L$  does not have the  $\mathbb{K}$ -embedding property,  $L\forall$  does not enjoy the SKC. Moreover, since  $L_{s,d}^{p,q}\forall$  is a conservative expansion of  $L\forall$ ,  $L_{s,d}^{p,q}\forall$  does not have the  $S\mathbb{K}_{s,d}^{p,q}C$  either.

It can be seen from the proof that if we only know  $L$  has the K-embedding property, we cannot conclude anything, in general, about the SSC of  $L_{s,d}^{p,q}\forall$ .

Theorem 5 can be generalized to arbitrary classes of L-chains and their  $s_1, \dots, s_p, d_1, \dots, d_q$ -expansions, which can be proved in an analogous way, as follows.

**Corollary 1 (Strong  $\mathbb{K}$ -completeness)** *Let  $L$  be a core fuzzy logic,  $\mathbb{K}$  a class of L-chains, and  $\mathbb{K}_{s,d}^{p,q}$  the class of the  $L_{s,d}^{p,q}$ -chains whose  $s_1, \dots, s_p, d_1, \dots, d_q$ -free reducts are in  $\mathbb{K}$ . Then:*

(i) *If  $L$  has the  $\mathbb{K}$ - $\sigma$ -embedding property, then  $L_{s,d}^{p,q}\forall$  has the  $S\mathbb{K}_{s,d}^{p,q}C$ .*

(ii) *If  $L$  does not have  $\mathbb{K}$ -embedding property, then  $L_{s,d}^{p,q}\forall$  does not have the  $S\mathbb{K}_{s,d}^{p,q}C$ .*

Similarly, first-order fuzzy logics with many hedges in which each hedge has its own dual one can be defined as follows.

**Definition 8** *Given a first-order core fuzzy logic  $L\forall$ , let  $L_{s,d}^{2n}\forall$  be the expansion of  $L\forall$  with new unary connectives  $s_1, \dots, s_n$  (for truth-stressers) and  $d_1, \dots, d_n$  (for truth-depressers), axioms  $(S_i^{dh})$ ,  $(S_{n+1}^{dh})$ ,  $(D_i^{dh})$ ,  $(SD_i^{dh})$ , and the deduction rule  $(DR^{dh})$ .*

$L_{s,d}^{2n}\forall$  can be obtained by expanding  $L_{s,d}^{2n}$  to the first-order level.

In an analogous way to the proofs of  $L_{s,d}^{p,q}\forall$ , we can prove the following theorems.

**Theorem 6** *The logic  $L_{s,d}^{2n}\forall$  is a conservative expansion of the logic  $L\forall$ .*

**Theorem 7** *Let  $L$  be a core fuzzy logic,  $\mathbb{K}$  a class of L-chains, and  $\mathbb{K}_{s,d}^{2n}$  the class of the  $L_{s,d}^{2n}$ -chains whose  $s_1, \dots, s_n, d_1, \dots, d_n$ -free reducts are in  $\mathbb{K}$ . Then:*

(i) *If  $L$  has the  $\mathbb{K}$ - $\sigma$ -embedding property, then  $L_{s,d}^{2n}\forall$  has the  $S\mathbb{K}_{s,d}^{2n}C$ .*

(ii) *If  $L$  does not have  $\mathbb{K}$ -embedding property, then  $L_{s,d}^{2n}\forall$  does not have the  $S\mathbb{K}_{s,d}^{2n}C$ .*

In the case that  $\mathbb{K}$  is the class of all standard L-chains, we have the corresponding results for the SSC of  $L_{s,d}^{2n}\forall$ .

## 4. LOGICS WITH HEDGES BASED ON A $\Delta$ -CORE FUZZY LOGIC

In this section, we consider logics with many hedges based on  $\Delta$ -core fuzzy logics, where the projection  $\Delta$  operator is definable.

The logic  $L_{s,d}^{p,q}$  based on a  $\Delta$ -core fuzzy logic  $L$  is defined as follows.



**Definition 9** Let  $L$  be a  $\Delta$ -core fuzzy logic. A logic  $L_{s,d}^{p,q}$ , where  $p, q$  are positive integers, is an expansion of  $L$  with new unary connectives  $s_1, \dots, s_p$  (for truth-stressers) and  $d_1, \dots, d_q$  (for truth-depressers) by the following additional axioms, for  $i = 1, \dots, p$  and  $j = 1, \dots, q$ :

$$\begin{aligned} (S_i) \quad & s_i\varphi \rightarrow s_{i-1}\varphi \\ (S_{p+1}) \quad & s_p\bar{1} \\ (D_j) \quad & d_{j-1}\varphi \rightarrow d_j\varphi \\ (D_{q+1}) \quad & \neg d_q\bar{0} \end{aligned}$$

and the following additional deduction rule:

$(DR^h)$  from  $(\varphi \rightarrow \psi) \vee \chi$  infer  $(h\varphi \rightarrow h\psi) \vee \chi$ , for each  $h \in \{s_1, \dots, s_p, d_1, \dots, d_q\}$ .

It can be proved that  $L_{s,d}^{p,q}$  is a semilinear logic, and in fact, it is the axiomatic expansion of  $L$  by the following additional axioms:

$$\begin{aligned} (S_i) \quad & s_i\varphi \rightarrow s_{i-1}\varphi \\ (S_{p+1}) \quad & s_p\bar{1} \\ (D_j) \quad & d_{j-1}\varphi \rightarrow d_j\varphi \\ (D_{q+1}) \quad & \neg d_q\bar{0} \\ (DR_\Delta^h) \quad & \Delta(\varphi \rightarrow \psi) \rightarrow (h\varphi \rightarrow h\psi), \text{ for each } h \in \{s_1, \dots, s_p, d_1, \dots, d_q\} \end{aligned}$$

We can obtain the first-order logic  $L_{s,d}^{p,q}\forall$  by expanding  $L_{s,d}^{p,q}$  to the first-order level in an analogous way as in Section 3, and we can have similar completeness results.

Concerning the logic  $L_{s,d}^{2n}$  based on a  $\Delta$ -core fuzzy logic  $L$ , we have the following definition.

**Definition 10** Let  $L$  be a  $\Delta$ -core fuzzy logic. A logic  $L_{s,d}^{2n}$ , where  $n$  is a positive integer, is an expansion of  $L$  with new unary connectives  $s_1, \dots, s_n$  (for truth-stressers) and  $d_1, \dots, d_n$  (for truth-depressers) by the following additional axioms, for  $i = 1, \dots, n$ :

$$\begin{aligned} (S_i^{dh}) \quad & s_i\varphi \rightarrow s_{i-1}\varphi \\ (S_{n+1}^{dh}) \quad & s_n\bar{1} \\ (D_i^{dh}) \quad & d_{i-1}\varphi \rightarrow d_i\varphi \\ (SD_i^{dh}) \quad & d_i\varphi \rightarrow \neg s_i\neg\varphi \end{aligned}$$

and the following additional deduction rule:

$(DR^{dh})$  from  $(\varphi \rightarrow \psi) \vee \chi$  infer  $(h\varphi \rightarrow h\psi) \vee \chi$ , for  $h \in \{s_1, \dots, s_n, d_1, \dots, d_n\}$ .

Also, it can be proved that  $L_{s,d}^{2n}$  is a semilinear logic, and in fact, it is the axiomatic expansion of  $L$  by the following additional axioms:

$$\begin{aligned} (S_i^{dh}) \quad & s_i\varphi \rightarrow s_{i-1}\varphi \\ (S_{n+1}^{dh}) \quad & s_n\bar{1} \\ (D_i^{dh}) \quad & d_{i-1}\varphi \rightarrow d_i\varphi \\ (SD_i^{dh}) \quad & d_i\varphi \rightarrow \neg s_i\neg\varphi \\ (DR_\Delta^{dh}) \quad & \Delta(\varphi \rightarrow \psi) \rightarrow (h\varphi \rightarrow h\psi), \text{ for each } h \in \{s_1, \dots, s_n, d_1, \dots, d_n\} \end{aligned}$$



Also, we can obtain the first-order logic  $L_{s,d}^{2n,\forall}$  by expanding  $L_{s,d}^{2n}$  to the first-order level in a similar way as in Section 3, and we can have similar completeness results.

## 5. CONCLUSION

In this paper, we propose first-order fuzzy logics with many hedges based on a first-order core fuzzy logic for two cases: (i) each hedge does not have a dual one, and (ii) each hedge can have its own dual one. More precisely, we expand two axiomatizations for propositional mathematical fuzzy logic with many hedges in previous works to the first-order level and prove a number of completeness results for the resulting logics w.r.t. the underlying logic. In fact, we prove that the new first-order logics are a conservative expansion of the underlying first-order core fuzzy logic and gives a characterization of their strong completeness, especially, the strong standard completeness. We also consider logics with many hedges based on  $\Delta$ -core fuzzy logics, which are extensions of core fuzzy logics by  $\Delta$  connective. They can be obtained in an analogous way, and similar completeness results can be proved.

## ACKNOWLEDGEMENTS

Funding from HUMG under grant number T16-02 is acknowledged.

## REFERENCES

- [1] P. Hájek, “On very true,” *Fuzzy Sets and Systems*, vol. 124, no. 3, pp. 329–333, 2001.
- [2] V. Vychodil, “Truth-depressing hedges and bl-logic,” *Fuzzy Sets and Systems*, vol. 157, no. 15, pp. 2074–2090, 2006.
- [3] F. Esteva, L. Godo, and C. Noguera, “A logical approach to fuzzy truth hedges,” *Inf. Sci.*, vol. 232, pp. 366–385, 2013.
- [4] V. H. Le, F. Liu, and D. K. Tran, “Mathematical fuzzy logic with many dual hedges,” in *Proceedings of the Fifth Symposium on Information and Communication Technology, SoICT '14, Hanoi, Vietnam*, 2014, pp. 7–13.
- [5] V. H. Le and D. K. Tran, “Linguistic logics with hedges,” in *Proceedings of the 2nd International Workshop on Semantic Technologies (IWOST2), Bangkok, Thailand*, 2015.
- [6] V. H. Le and F. Liu, “Tabulation proof procedures for fuzzy linguistic logic programming,” *Int. J. Approx. Reasoning*, vol. 63, pp. 62–88, 2015.
- [7] P. Hájek, *Metamathematics of Fuzzy Logic*. Dordrecht, The Netherlands: Kluwer, 1998.
- [8] P. Cintula, P. Hájek, and C. Noguera, Eds., *Handbook of Mathematical Fuzzy Logic*, ser. Studies in Logic, Mathematical Logic and Foundations. College Publications, 2011.
- [9] L. A. Zadeh, “A fuzzy-set-theoretic interpretation of linguistic hedges,” *Journal of Cybernetics*, vol. 2, no. 3, pp. 4–34, 1972.

- [10] C. H. Nguyen and W. Wechler, “Hedge algebras: An algebraic approach to structure of sets of linguistic truth values,” *Fuzzy Sets and Systems*, vol. 35, pp. 281–293, 1990.
- [11] C. H. Nguyen and W. Wechler, “Extended hedge algebras and their application to fuzzy logic,” *Fuzzy Sets and Systems*, vol. 52, pp. 259–281, 1992.
- [12] P. Cintula, F. Esteva, J. Gispert, L. Godo, F. Montagna, and C. Noguera, “Distinguished algebraic semantics for t-norm based fuzzy logics: Methods and algebraic equivalencies,” *Ann. Pure Appl. Logic*, vol. 160, no. 1, pp. 53–81, 2009.

## AUTHOR

Van-Hung Le received a B.Eng. degree in Information Technology from Hanoi University of Science and Technology (formerly, Hanoi University of Technology) in 1995, an M.Sc. degree in Information Technology from Vietnam National University, Hanoi in 2001, and a PhD degree in Computer Science from La Trobe University, Australia in 2010. He is a lecturer in the Faculty of Information Technology at Hanoi University of Mining and Geology. His research interests include Fuzzy Logic, Logic Programming, Soft Computing, and Computational Intelligence.



# A NOVEL ENF EXTRACTION APPROACH FOR REGION-OF-RECORDING IDENTIFICATION OF MEDIA RECORDINGS

Majed El Helou, Abdel Wahab Turkmani, Rawan Chanouha and  
Samer Charbaji

Department of Electrical and Computer Engineering,  
American University of Beirut, Beirut, Lebanon  
{mhe33, ast11, rsc08, sac13} @mail.aub.edu

## ABSTRACT

*The electric network frequency (ENF) of power lines leaves its trace in nearby media recordings. The ENF signals vary in a consistent way in a given power grid. Therefore, it is possible to develop signal processing and machine learning techniques to identify the grid of origin by extracting attributes of the embedded ENF signal in recorded audio. This paper presents a model based on a novel ENF extraction technique with training on audio and power recordings from different grids. The proposed approach is based on correcting erroneously selected peaks from the Short Time Fourier Transform (STFT) by leveraging time correlations. These peaks are mistakenly taken for the frequency component belonging to the embedded ENF signal of the power grid and are corrected by the algorithm. Results on a test set of 50 recordings from nine different locations demonstrate the effectiveness of the proposed approach with an overall accuracy of 88%.*

## KEYWORDS

*ENF, Region of Recording, Media Recordings, Power Grids, Noise in Audio Recordings.*

## 1. INTRODUCTION

Electric network frequency (ENF) signals, which are embedded in recordings made close to power lines, can be used to identify the origin of these media recordings. This serves the function of allowing audio recordings to be validated as unaltered in legal applications. Moreover, the uniqueness [1] of the embedded ENF components to its grid of origin allow for the localisation of media recordings, allowing for applications in domains of authentication, security and crime fighting [2, 3].

State of the art methods involve examining the Fourier Transform over overlapping time frames of the data and combining identified frequency components, or their harmonics, to construct an estimate of the ENF signal variation, which is captured in the recording [4].

In this paper, a model for ENF extraction and grid of origin prediction is proposed. Particularly, the paper suggests a novel algorithm for improving the accuracy of the estimation of ENF values extracted from audio recordings. This algorithm comes as a solution to the presence of noise, which cannot be easily removed by using conventional noise cancellation algorithms since the

David C. Wyld et al. (Eds) : DBDM, CICS, CSIP, AI&FL, SCOM, CSE, CCNET-2016  
pp. 97–108, 2016. © CS & IT-CSCP 2016 DOI : 10.5121/csit.2016.60510

ENF itself is considered as noise inside an audio recording. Particularly, the paper proposes a solution whereby correction is done by studying the strongest frequency component, checking if it satisfies certain noise thresholds, and then replacing it by the next strongest frequency component should it fail to do so.

In addition, the paper presents the design and implementation of a hardware circuit to capture power mains reference recordings. A simple circuit is used in tandem with an NI MyDAQ in order to record the reference signal. Finally, the data is analysed using our proposed algorithm and we show that our reference recordings belong to one of the recordings in the training set.

The rest of this paper is organised as follows. Section II covers some work done related to our topic. In Section III, details of the extraction of the electric network frequency signals from audio recordings and power reference recordings are discussed. Section IV tackles the details of the extraction of features of the ENF signal and discusses the machine-learning model generated. Section V covers the recording of local power reference signals and presents an analysis of the recordings. Section V concludes the paper.

## 2. RELATED WORK

The ENF signal constitutes a pattern observed in audio recordings. It represents the most significant frequency component of the voltage on a specific power grid at each second [18]. The analysis of ENF signals was motivated by many forensic applications. Multiple studies were conducted to explore the possibility of determining either the time of recording of the audio [19] or the location where it was recorded through the analysis of the ENF signal embedded in it [17]. Many means of extracting this signal were proposed in the literature. For example in [20], the authors divide the audio into overlapping frames and estimate the prevailing instantaneous frequency in each frame. However, in this paper, the authors focused on recaptured audio recordings which may contain multiple ENF signals. In [21], two main challenges faced in ENF estimation were tackled. The authors propose a special short time Fourier transform (STFT) approach to solve the issues of noise and frequency resolution. The latter is chosen to be equal to the threshold within which the true peak value can be recovered from the noise interference. Another algorithm for ENF extraction was discussed in [22]. The approach relies on binary search to narrow the search band until an accurate estimate of the ENF information is obtained. Another method proposed in [23] estimates the ENF signal by tracking it across different frames. The number of peaks in each frame is selected to minimize the Bayesian Information Criterion (BIC) model. Once the ENF signal is obtained, one needs to match it to other present signals in the database as presented in [23, 22, and 17]. In [17], it is proven that there exists an approximate linear relationship between the correlation coefficient between the observed ENF and a signal in the database and the distances between the two recording locations. In [23], this is done through minimizing a squared error and equivalently maximizing the correlation coefficient between the estimated ENF signal and the corresponding signal from the database. In [22], a penalty coefficient is introduced which will use conventional methods under good conditions (weak noise and long recording time) but adopt a different approach in case a non-trivial noise component is present by relying on the fact that the noise is not dominant during all time instances. In this paper, we present a novel method which relies on extracting features from the ENF signal and applying classical machine learning techniques in order to make an accurate guess on each of the extracted ENF signals.

## 3. ENF EXTRACTION

### 3.1. Power grid operating frequency detection

The ENF signal is defined as the sequence of the changing values over time of the instantaneous ENF, which is in itself the supply frequency of power distribution grids. The extraction of ENF

begins with a close study of the Fast Fourier Transform (FFT) of the signals. Looking at the signal in the frequency domain, the aim was to find the component corresponding to the trace of the power grid, which is expected to be around 50Hz or 60Hz, and then to estimate the exact instantaneous ENF of that grid.

The algorithm to extract the ENF signal across time starts by identifying whether the grid ENF is around 50Hz or 60Hz. To achieve this, the proposed algorithm starts by taking the magnitude of the FFT of the recording. Since the recording is a real signal its FFT is even, and it is enough to study one of the two symmetric halves of the magnitude of the FFT. When studying the frequency components, it is noticeable that in some signals some harmonics of the 50Hz or 60Hz component are relatively stronger than others, whereas some are relatively very weak. Because most harmonics were negligible compared to the strongest one in the majority of cases, and mostly to keep the complexity minimal, the algorithm considers the ENF at the strongest harmonic

The algorithm computes the distance between the highest frequency component  $F$  in the FFT and each of the first 20 harmonics of 50/60Hz (For ex: 50, 100, 150, 200Hz etc. and 60, 120, 180, 240Hz etc.). The closest harmonic to the strongest frequency value is thus determined. Let  $d$  be the distance between that harmonic and  $F$ . This closest harmonic,  $H_{\text{number}}$ , is taken as the correct one corresponding to the power grid if  $d$  is less than a chosen frequency filter half width  $F_{w/2}$  (half the width of a rectangular window filter chosen to be applied in the frequency domain, which will be discussed shortly). The width of the window is always scaled to the corresponding harmonic. This procedure can be expressed as follows:

$$d < F_{w/2} * H_{\text{number}} \quad (1)$$

### 3.2. ENF Estimation

After determining whether the grid ENF is around 50Hz or 60Hz, the algorithm creates the aforementioned bandpass filter centred at  $F$ :

$$(F - F_{w/2}, F + F_{w/2}) \text{ Hz} \quad (2)$$

Once the filter is created, the algorithm computes a short-time Fourier Transform (STFT). This is done by first choosing a padding factor ( $P$ ), time window ( $Tw$  (sec) and a per cent overlap ( $O$ ) then moving in time along the recording. The FFT is taken across consecutive sections of the recording, after padding the section with zeros to increase the length in time of each section by the padding factor and thus have a more precise frequency domain signal after taking the FFT [5, 6]. The sampling frequency of all the down-sampled signals used being 1kHz, the FFT bin size is equal to:

$$0.5 \log_2(Tw * P * 1kHz) \quad (3)$$

Each section will have a width equal to  $Tw$  seconds, and will be overlapping with the previous section over a duration of time equal to  $Tw * O$  seconds.

The bandpass filter is applied in the frequency domain to the consecutive FFT signals computed on these time windows to get rid of as much noise as possible [7].

From the output of the filter, the strongest frequency component is determined. Using this frequency value as well as the one right before it and the one right after it (the two adjacent FFT frequency bins) the algorithm estimates the value of the strongest frequency component by quadratic interpolation [5, 8, 9] which makes up for the precision loss of the discrete Fourier transform.

Finally, the ENF signal is scaled back to the base harmonic by regular division in case it was extracted from a higher order harmonic, simply for the conservation of consistency in all readings.

### 3.3. Peak Correction in noisy signals

At this stage, the ENF signals extracted from power recordings are in conformity with ENF signals seen in the literature [10]. However, those extracted from audio signals remain highly noisy i.e. having many fast variations above and below the moving mean of the ENF, in comparison with the ENF signals extracted from power recordings taken from the same power grid. Simple smoothing algorithms, including median filtering [8] and outlier rejection algorithms do not fix the incorrect readings in addition to making the initially accurate readings deviate from the correct frequency values.

Looking carefully at the frequency domain equivalent of time windows where readings that are deemed incorrect and that cause a noisy ENF were taking place, the following was observed: in the frequency domain, the strongest frequency component around the chosen harmonic is correctly identified, however, the problem is this peak is not in fact the ENF component, but simply a noise frequency [4]. As can be seen in the Figure 1, there is another strong local peak inside the bandpass filter region, which is slightly weaker than the highest peak first found.

The two border labels in the figure delimit the bandpass filter frequency limits, while the two middle labelled points correspond to the local peaks mentioned above.

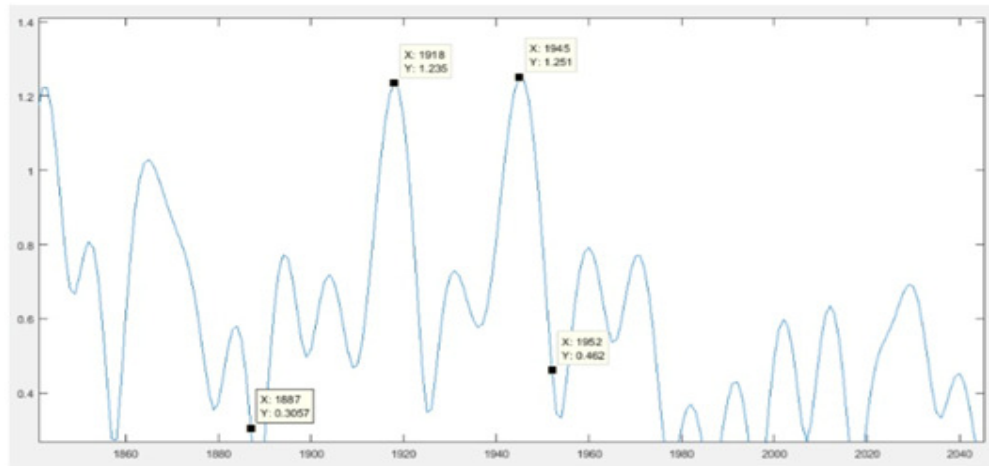


Figure 1. Magnitude vs. Frequency in the FFT of a Time Window

The global peak ( $P_G$ ) at frequency  $F_G$ , which causes the extracted ENF to be noisy, is very close to the filter's upper limit, while the next local peak ( $P_2$ ) located at frequency  $F_2$  is centred in the middle of the bandpass filter. Through the inspection of many similar cases of noisy points in the final ENF, the conclusion that this second largest local peak inside the bandpass filter is actually the ENF component was drawn. This conclusion holds in a wide majority of the noise cases. An algorithm was developed to identify these second strongest local peaks while extracting the ENF signal. Imposing a noise threshold ( $N_o$ ) and considering the previously computed instantaneous ENF value  $F_{old}$ , if the conditions (4) and (5) hold, then the second local peak is taken to be the ENF component and not the global peak.

$$P_2 > N_o * P_G \quad (4)$$

$$d(F2, \text{Fold}) < \text{distance}(F1, \text{Fold}) \quad (5)$$

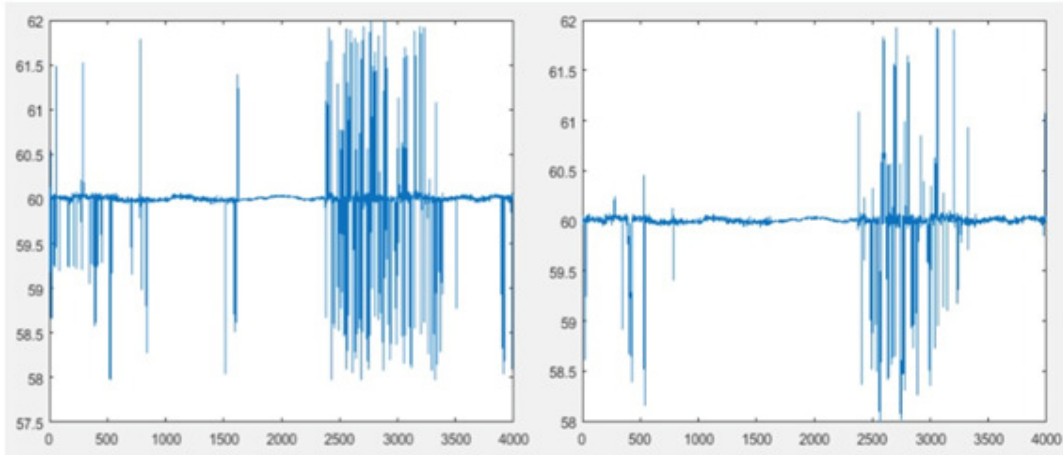


Figure 2. Audio ENF Post-Processing

Figure 2 shows the extracted ENF without applying the novel post-processing detection algorithm (left figure) and the ENF extracted with the post-processing algorithm applied (right figure). As can be noted, the right figure ENF has significantly less noisy points than the ENF extracted without the novel algorithm, and, even more importantly, points that are not originally noisy are not affected. There remains a fraction of the noisy point readings in the right figure for the following reasons:

- 1) The algorithm is not applied very aggressively by setting the *noise threshold* to relatively high percentage values. This ensures that the algorithm does not affect previously correctly predicted ENF values. Thus correct estimates remain unaffected with high probability.
- 2) Additionally, the algorithm only accounted here for the second strongest peak and corrected between that peak and the global maximum peak. However, it is possible to concurrently consider the third, fourth, etc. peaks in correcting the original one. Nevertheless, this comes at a higher computational cost.

The algorithm leverages the fact that there is an inherent time correlation between consecutive estimated instantaneous ENF values, and thus the variability in ENF values cannot be, on average, too large. This is because the ENF of the grid at time  $t_n$  is not independent from that at time  $t_{n-1}$ , the distribution could actually be approximated by a random walk if the variations on the grid were assumed to be Gaussian noise. The algorithm thus takes advantage of that information to single out estimations that are altered by noise and estimates again a more probabilistically plausible value of the ENF at that time instant.

Since the problem at hand does not occur in the case of power recordings, this extra processing is only done for audio recordings to reduce the running time. To differentiate whether a given recorded signal is a power recording or an audio recording, the power spectral density (PSD) of the signal is compared to a chosen threshold: *PSD threshold*. The power spectral density of a power recording is far larger than that of an audio recording because the power is highly focused at the ENF component only, unlike with audio recordings. Audio recordings contain many more frequency components aside from the ENF, the latter being simply a trace. This makes this decision accurate and highly confident.



Finally, all of the aforementioned parameters were chosen after thorough testing and studying of the processing steps as well as the literature (for first four parameters) and are grouped in Table 1.

The *frequency filter half width* ( $F_w/2$ ) was set to 2 Hz at the base harmonic to filter out as much noise as possible (the ENF signals not having a range going beyond 4 Hz), while not affecting the frequency components inside the bandpass filter by making it too narrow, since no rectangular filter in the frequency domain is ideal as its counterpart requires infinite length in the time domain (infinite sinc function) [11].

The *time window* ( $T_w$ ) was initially chosen to be 5 seconds [10] but was then reduced to only 3 seconds through testing to achieve higher accuracies. Additionally shorter window times result in a better fit with the autoregressive model, which would fit a shorter signal in general more accurately than a lengthier one. The same was done with the *per cent overlap* ( $O$ ) which, initially set to be 50%, ended up being an 85% overlap after tweaking. For the *padding factor*, the chosen value was based on the observations from the maximum-bias curves for “rect” windows noted in [5], in a way to reduce the final bias ratio. Lastly, the *noise threshold* ( $N_o$ ) of the newly proposed algorithm was first chosen by inspection of cases similar to Figure 1, and then tweaked for best model accuracy.

Table 1. Parameter Values of the Proposed Model

Parameter	Chosen Value
<i>frequency filter half width</i>	2 Hz
<i>time window</i>	3 seconds
<i>padding factor</i>	16
<i>percent overlap</i>	85%
<i>noise threshold</i>	0.7

## 4. ORIGIN CLASSIFICATION

### 4.1 Selected Features

After careful examination of the extracted ENF signals, a list of features were identified in order to differentiate between the signals coming from different grids by training a machine learning model on those features extracted from training data. The first feature chosen is the mean of the ENF signal. As explained in [10], the mean not only allows us to identify whether the signal belongs to a 50 Hz grid or a 60 Hz grid but also within the same type of grids the mean is never exactly equal to its ideal value and this shift from the ideal frequency is reasonably consistent in the recordings within the same grid.

Other important statistical measures chosen are the variance, the median and the mode, which are additional descriptors of the ENF signal, noted to be rather consistent for a given grid. The range as well as the minimum and the maximum values of the ENF have also proven to provide a useful insight for the classification process by helping with the distinction between some grids. A log base 10 operator is applied to the range and the variance instead of directly using their initial values to emphasize the fact that only a significant change in these parameters is wanted to affect the classification model [10]. The skewness, which is defined as a measure of the degree of symmetry in the data, and the kurtosis, describing the sharpness of the peaks near the mean, both were used as features in classifying the recordings, with the kurtosis being a notably relevant feature characterizing ENF signals coming from different grids.



In addition, a wavelet transform is applied to the original ENF signal in order to extract statistical characteristics of the data at different resolution levels [10]. Specifically, a 9-Level wavelet decomposition was chosen and the log base 10 of the variance of each of the 9 resulting detailed levels was used as a statistical feature from every resolution level. Also the variance of the approximated signal is used to reflect the general variation pattern of the signal ignoring relatively high- frequency variations in the ENF signals. An autoregressive (AR) model, of order 2, was then developed.

$$y[n] = x_1 y[n - 1] + x_2 y[n - 2] + v[n] \quad (6)$$

Three features were obtained from this new model, notably the two AR parameters as well as the log base 10 of the variance of the model. These features give insight on how the signal can be fitted into an AR model [10].

Additionally, spectrum features are also famous parameters used in audio analysis. Among these features, the spectral centroid, a measure of the centre of gravity of the STFT's magnitude [12], and the spectral roll-off, which represents the frequency below which a fraction (commonly 85%) of the total energy is present, proved to give better results in terms of the accuracy of the classification when integrated into the model. The mean crossing was also computed to capture how many times the ENF signal crosses its mean value, going over and under it, and also, the zero-crossings of the derivative of the ENF signal as well the maximum value of the derivative were added as classification features to portray how fast the ENF signal changes its direction of variation (increasing, decreasing) and how abrupt is its fastest change in frequency value.

Finally, the outlier ratio of the ENF signal, which is a measure of instances that do not fall within two standard deviations of the mean value, also turned out to be a suitable criterion for the classification. These 28 features, mentioned earlier, constitute the training set for the classifier and are summarized in the Table 2.

Additionally all the features which are affected by ENF signal length were normalized by dividing their value by the signal's length, although all training data was finally chosen to have equal length of 20 minutes per sample as explained shortly.

Table 2. Summary of Selected Features

Index	Feature
1	Mean of ENF signal.
2	Log10 (var) of ENF signal.
3	Log10 (range) of ENF signal.
4-12	Log10 (var) of the 9 detailed levels of the ENF signal obtained by applying a wavelet transform on the original signal.
13	Log10 (var) of the approximated signal obtained after applying the wavelet transform (with 9 levels) on the original ENF.
14-15	Auto-regression parameters.
16	Log10 (var) of the autoregressive model innovation signal.

17	Median of the ENF signal.
18	Mode of the ENF signal.
19	Skewness of the ENF signal.
20	Kurtosis of the ENF signal.
21-22	The minimum and maximum value of the ENF signal respectively.
23	Mean crossing of the ENF signal.
24-25	Spectral centroid and spectral roll-off respectively.
26	Maximum derivative.
27	Outlier ratio.
28	Derivative zero-crossing.

## 4.2 Classification Model

Prior to training the classifier, the training data was split into 20-minute frames, with a 50% overlap across frames. This increased the total training set from 111 to 790 entries and contributed substantially to increasing the accuracy of the classifier. Among all the classifiers tested, Random Forests have proven to be the most successful. They outperformed SVM, AdaBoost, and many other widely used algorithms. This is due to the inherent randomness in selecting the data to train its weak learners on, and in selecting the features to provide its weak learners (decision trees) to split on [13].

These characteristics, combined with splitting the training data into overlapping smaller components, which resulted in large quantities of data samples to train on, ensued in remarkable gains in accuracy. The chosen bagging tree classifier is trained over 30 cycles, and actually 255 such classifiers are generated. Each of the 255 classifiers is trained to predict whether or not a sample belongs to a chosen subset of the 9 grid classes, and for prediction, all of the 255 classification results are combined into a single one. There are 255 possible subsets from the 9 grids, 255 being  $2^{\# \text{of grids}} - 1 - 1$  (-1 because the subset containing all of the 9 grids is meaningless).

It is noteworthy to mention that the trained classifier does not account for a none-of-the-above option to keep the focus exclusively on classifying samples belonging to grids the classifier was trained on as previous works have shown this approach to actually improve overall accuracy [10]. Lastly, the classification model is used to generate posterior probabilities corresponding to each grid class, which are used as confidence measures in every prediction.

The proposed algorithm was tested on a test dataset with 50 recordings, comprising a combination of audio and power recordings coming from more than the 9 grids in the training set. As mentioned above, the constructed model does not account for out-of-training-set (N) classes and thus will classify them erroneously. Tables 3 and 4 contain the performance results of the full model on the test dataset. Table 3 portrays the accuracy of the model without using the novel ENF estimation algorithm in noisy audio recordings.

Table 3 Result of Classification without Applying the Novel Peak-Correction Algorithm

Accuracy (%)	Power	Audio	Power and Audio
Including N	89.2857	54.5454	74

This accuracy clearly improves with the application of the peak-correction algorithm, through significantly improving the system's performance on audio recordings which are naturally more noisy signals. Final accuracies are reported in Table 4.

With the first row indicating the accuracies when the signals belonging to class N (out of training set class) are included and the second row indicating the accuracy only on signals belonging to classes the model was trained on. As can be seen, the algorithm performs exceptionally well in predicting the origin of signals it was trained on, with the accuracy of audio signals displaying a tremendous improvement using the novel peak-correction technique.

Table 4. Results of Classification on Test Dataset with the Novel Algorithm

Accuracy (%)	Power	Audio	Power and Audio
Excluding N	96.4285	100	98
Including N	89.2857	86.3636	88

## 5. DATA ACQUISITION AND ANALYSIS

To obtain reference power recordings, a simple circuit, shown in the diagram below, was devised based on the one used in [14]. The focus in this circuit is on its simplicity and its portable design as well as on the accessibility of its components, which consist of a transformer, operational amplifier, capacitor, and resistors. The circuit is easily portable and can be directly plugged into different wall outlets in any location, which makes collecting recordings at different locations within the grid a quite simple task.

The step down transformer is used to bring down the 220VAC signal from the power outlet to a 25V peak-to-peak signal. The signal is then passed through a low pass filter, which consists of a simple RC circuit with the output signal taken across the capacitor. The RC circuit serves as a low-pass filter and performs anti-aliasing to allow for proper acquisition. In order to remove the loading effect and further bring down the voltage to readable values, a unity gain amplifier with a voltage divider at the buffer output is utilised, resulting in a final 2V peak-to-peak signal biased at 0V [15]. The resulting output signal is read by a NI MyDAQ hardware and fed into LabVIEW where it is recorded at a sampling rate of 1000 samples/sec and number of samples equal to 1000. The resulting signal is then exported as a .wav file to the machine running LabVIEW.

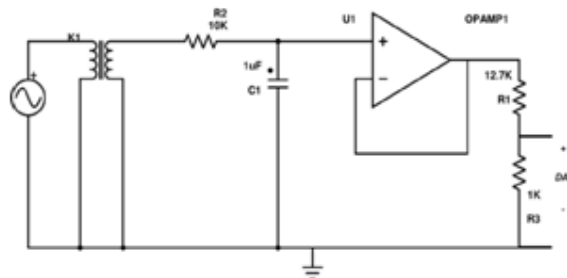


Figure 3 Schematic of the Chosen Circuit Design.

Over the duration of our work, we collected night onto 14 hours, in one-hour samples, of clean power recordings from the power grid in Lebanon. Much care was taken to filter out recordings where mains power was cut due to country-wide regular power outages and instead back-up generators were used to supply electricity. Next, the ENF component of the signals is extracted, and a sample is shown in Figure 5.

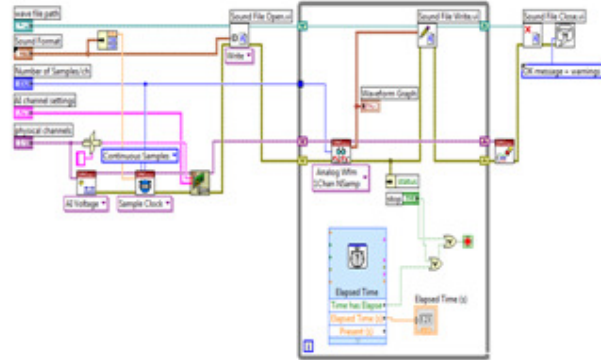


Figure 4 LabVIEW Data Acquisition VI.

Across all the acquisitions, it can be noted that on average the ENF was centred around 50.6Hz to 50.8Hz. Additionally, a common trend observed across recordings are occasional plunges of magnitude around 1Hz to 1.5Hz in the ENF, with drops from  $\sim 50.7\text{Hz}$  down to 49.3Hz to be the trend. A sample ENF is shown in Figure 5, highlighting one of the plunges.

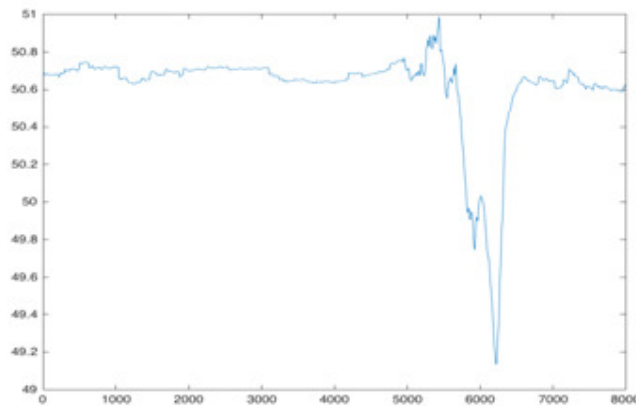


Figure 5 ENF of Signal Recorded using MyDAQ.

## 6. CONCLUSIONS

In this paper, we document our work for building a model for the extraction of ENF from media signals and for the classification of the origin of media recordings and power mains reference signal recording. Additionally, we propose a novel technique for ENF extraction that considerably improves the accuracy of the extraction in the presence of noise, as indicated by the difference in accuracies obtained with and without the ENF extraction method and as portrayed in the discussions. This method leverage the existence of correlation across time between ENF values, and thus estimates instantaneous ENF values with aim a reduction in overall variability frequency, thus having a similar yet much more accurate effect than smoothing. Moreover, we present the multi-class machine learning algorithm utilized to classify each grid as well as the features that were employed in the classification model. Lastly, for the purpose of ensuring reproducibility the hardware system developed is also portrayed in the paper.

**REFERENCES**

- [1] M. H. Bollen and I. Gu, "Signal processing of Power Quality Disturbances," Hoboken, NH, USA: Wiley, 2006.
- [2] C. Grigoras, "Applications of ENF Analysis in Forensic Authentication of Digital Audio and Video Recordings," J. Audio Engineering Society, vol. 57, no. 9, pp. 643-66. Sept. 2009
- [3] N. Fechner and M. Kirchner, "The Humming Hum: Background Noise as a Carrier of ENF Artifacts in Mobile Device Audio Recordings," in Proc. Of 2014 8th International Conference on IT Security Incident Management & IT Forensics, Munster, Germany, pp. 3-13. May 2014,
- [4] A. Hajj-Ahmad, R. Garg and M. Wu, "ENF-Based Region-of-Recording Identification for Media Signals", in Proc. Of IEEE Trans.Inform.Forensic Secur., vol. 10, no. 6, pp. 1125-1136, 2015.
- [5] M. Abe, & J.O Smith III, "Design criteria for simple sinusoidal parameter estimation based on quadratic interpolation of FFT magnitude peaks." In Audio Engineering Society Convention 117. Audio Engineering Society 2004.
- [6] A. Cooper, "An automated approach to the Electric Network Frequency (ENF) criterion - Theory and practice", IJSL, vol. 16, no. 2, 2010.
- [7] J. O. Smith and X.Serra, "PARSHL: An analysis/synthesis program for nonharmonic sounds based on sinusoidal representation," in Proc. Of Int. Computer Music Conf. pp. 290–297. 1987
- [8] A. Hajj-Ahmad, R. Garg and M. Wu, "Spectrum Combining for ENF Signal Estimation", in Proc. Of IEEE Signal Processing Letters, vol. 20, no. 9, pp. 885-888, 2013.
- [9] O. Ojowu, J. Karlsson, J. Li and Y. Liu, "ENF Extraction From Digital Recordings Using Adaptive Techniques and Frequency Tracking", in Proc. Of IEEE Trans.Inform.Forensic Secur., vol. 7, no. 4, pp. 1330-1338, 2012.
- [10] Y. Yang and H. Chen, Music emotion recognition. Boca Raton, Fla.: CRC, pp.44-45, 2011.
- [11] Y. Zhang and D. LV, "Selected Features for Classifying Environmental Audio Data with Random Forest", TOAUTOJ, vol. 7, no. 1, pp. 135-142, 2015.
- [12] P. Top, M.R. Bell, E. Coyle, O. Wasynczuk, "Observing the Power Grid: Working Toward a More Intelligent, Efficient, and Reliable Smart Grid with Increasing User Visibility," in Proc. Of Signal Processing Magazine, IEEE , vol.29, no.5, pp.24-32, Sept. 2012
- [13] Y. Liu; C. Yao; C. Sun; Y. Liu, "The Authentication of Digital Audio Recordings Using Power System Frequency," in Proc. Of Potentials, IEEE , vol.33, no.2, pp.39-42, March-April 2014
- [14] T. Yardibi, J. Li, P. Stoica, M. Xue, and A. B. Baggeroer, "Source localization and sensing: A nonparametric iterative adaptive approach based on weighted least squares," in Proc. Of IEEE Trans. Aerospace Electron. Syst., vol. 46, no. 1, pp. 425–443, Jan. 2010.
- [15] B. Kosko, Noise. New York: Viking, 2006
- [16] M. Stamm, Min Wu and K. Liu, "Information Forensics: An Overview of the First Decade", IEEE Access, vol. 1, pp. 167-200, 2013.
- [17] R. Garg, A. Hajj-Ahmad and M. Wu , "Geo-location estimation from electrical network frequency signals", in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. , pp.1 -3
- [18] T. Baksteen, "The ENF criterion: A literature study", 2015.

- [19] M. Huijbregtse and Z. Geradts, "Using the ENF criterion for determining the time of recording of short digital audio recordings", *Computational Forensics*, pp. 116-124, 2009.
- [20] H. Su, R. Garg, A. Hajj-Ahmad and M. Wu, "ENF analysis on recaptured audio recordings", *Acoustics, Speech and Signal Processing (ICASSP)*, 2013 IEEE International Conference on., pp. 3018-3022, 2013.
- [21] G. Hua, J. Goh and V. Thing, "A Dynamic Matching Algorithm for Audio Timestamp Identification Using the ENF Criterion", *IEEE Trans.Inform.Forensic Secur.*, vol. 9, no. 7, pp. 1045-1055, 2014.
- [22] Ling Fu, P. Markham, R. Connors and Yilu Liu, "An Improved Discrete Fourier Transform-Based Algorithm for Electric Network Frequency Extraction", *IEEE Trans.Inform.Forensic Secur.*, vol. 8, no. 7, pp. 1173-1181, 2013.
- [23] O. Ojowu, J. Karlsson, J. Li and Y. Liu, "ENF Extraction From Digital Recordings Using Adaptive Techniques and Frequency Tracking", *IEEE Trans.Inform.Forensic Secur.*, vol. 7, no. 4, pp. 1330-1338, 2012.

## AUTHORS

**Majed El Helou** is currently pursuing a B.E. degree in computer and communication engineering with a math minor from the American University of Beirut, Beirut, Lebanon, expected to graduate in 2016. He was with the Computer Science and Artificial Intelligence Lab at the Massachusetts Institute of Technology as a Research Intern in the Wireless Net group, in summer 2015. His current research interests include signal and image processing, machine learning and statistical analysis, biomedical engineering and communication. He was listed on the Faculty of Engineering and Architecture Dean's Honor List for every semester of his undergraduate studies.

**Abdel Wahab Turkmani** is currently pursuing a B.E. degree in electrical and computer engineering from the American University of Beirut, Beirut, Lebanon, expected to graduate in 2016. He was with the Applications Engineering team at National Instruments as an Electrical Engineering Intern during the summer of 2015. His current research interests include machine learning, statistical analysis, computer vision and computational finance. He was listed on the Faculty of Engineering and Architecture Dean's Honor List for every semester of his undergraduate studies.

**Rawan Chanouha** is currently pursuing a B.E. degree in electrical and computer engineering with a math minor from the American University of Beirut, Beirut, Lebanon, expected to graduate in 2016. She was with the electrical engineering department at Veolia Water Technologies as an Electrical Engineering Intern, in summer 2015. Her current research interests include machine learning, big data, antenna design and biomedical engineering. She was listed on the Faculty of Engineering and Architecture Dean's Honor List for every semester of her undergraduate studies.

**Samer Charbaji** is currently pursuing a B.E. degree in computer and communications engineering from the American University of Beirut, Beirut, Lebanon, expected to graduate in 2016. He was with the Information Visualization and Management Lab in Purdue Polytechnic as a Research Intern, in summer 2015. His current research interests include machine learning, big data, supply chain management and strategic IT adoption. He was listed on the Faculty of Engineering and Architecture Dean's Honor List for every semester of his undergraduate studies.

# AUTOMATIC TRANSLATION OF ARABIC SIGN TO ARABIC TEXT (ATASAT) SYSTEM

Abdelmoty M.Ahmed<sup>1</sup>, Reda Abo Alez<sup>2</sup>, Muhammad Taha<sup>3</sup> and  
Gamal Tharwat<sup>4</sup>

<sup>1, 2, 4</sup>Department of Systems and Computer Engineering,  
Faculty of Engineering /Al-Azhar University, Egypt  
<sup>3</sup>Faculty of Science /Al-Azhar University, Cairo, Egypt

## ABSTRACT

*Sign language continues to be the preferred tool of communication between the deaf and the hearing-impaired. It is a well-structured code by hand gesture, where every gesture has a specific meaning, In this paper has goal to develop a system for automatic translation of Arabic Sign Language. To Arabic Text (ATASAT) System this system is acts as a translator among deaf and dumb with normal people to enhance their communication, the proposed System consists of five main stages Video and Images capture, Video and images processing, Hand Signs Construction, Classification finally Text transformation and interpretation, this system depends on building a two datasets image features for Arabic sign language gestures alphabets from two resources: Arabic Sign Language dictionary and gestures from different signer's human, also using gesture recognition techniques, which allows the user to interact with the outside world. This system offers a novel technique of hand detection is proposed which detect and extract hand gestures of Arabic Sign from Image or video, in this paper we use a set of appropriate features in step hand sign construction and classification of based on different classification algorithms such as KNN, MLP, C4.5, VFI and SMO and compare these results to get better classifier.*

## KEYWORDS

*Sign language, Hand Gesture, Hand Signs Construction, gesture recognition, hand detection*

## 1. INTRODUCTION

Sign language is the prominent means of communication among the deaf and the hearing-disabled. Two approaches are mainly followed in sign language; they are vision-based and bio mechanical (Gloves sensor). The advantage of vision-based systems over the counterpart is that users do not need to use complex equipments but in the pre-processing stage requires sufficient computations in place of cameras which are used in vision based systems , sensor-based systems use sensor enabled instrumented gloves this paper presents computer vision-based gesture interface that is part of a sign language recognition system, and also explains computerized sign language recognition system for the vocally disabled (deaf and dumb) that uses sign language for communication.

Generally there are three levels of image based ATASAT, they are continuous recognition, alphabet recognition and isolated-word recognition. The input of the vision based methods is a set of images or video sequence of the signs. The signers are asked to have pause between the signs to isolate the signs which is done manually. This paper will present research progress and findings on techniques and algorithms for hand detection as it will be used as an input for gesture recognition process.

The ATASAT system using image and pattern recognition technology is presented in Section I. Related works are described in section 2. Overview of the proposed system is discussed in Section 3. In Section 4, the computer simulation results of the system are presented, finally Conclusions and Future Directions are summarized in Section 6.

## 2. RELATED WORKS

Deaf and dumb or Hearing impairment people cannot talk and hear like normal people; so they have to depend on some types of visual communications in most of the time. [1]

The communication between hearing impairment and normal people depends only on the sign language, while the vast majority of normal people don't know this language. Sign language is not universal; it varies according to the country or regions, class of people [2].

A finger spelling recognition method by using distinctive features of hand shape was proposed by Tabata et al [3]. An Arabic sign language translation system on mobile devices was introduced by Halawani [4]. To use input device of wearable computer, a new glove-based input device was proposed by Tsukada et al [5]. Statistical template matching was used to recognize Pakistan sign language based on data glove by Khalid Alvi et al [6]. Arabic sign language recognition system was developed using an instrumented glove proposed by Al-Buraiky et al [7]. Arabic sign language (ArSL) has recently been recognized and documented. Many efforts have been made to establish the sign language used in Arabic countries. are trying to standardize the sign language and spread it among members of the deaf community and those concerned. However, Arabic speaking countries deal with the same sign alphabets [8], [9].

Feris et al [10] proposed an approach to exploit depth discontinuities for finger spelling recognition to differentiate between similarities of some signs by using multi flash camera. On the other hand, Tanibata et al. [11] provided a prototype approach based on feature extraction to solve hand occlusion problem for Chinese sign language recognition. Mohandes [12], [13] introduced a prototype system to recognize the Arabic sign language based on Support Vector Machine (SVM) and also an automatic Translation system to translate Arabic Text to Arabic Sign Language. Foong et al [14] proposed A Sign to Voice system prototype which is capable of recognizing hand gestures by transforming digitized images of hand sign language to voice using Neural Network approach.

## 3. OVERVIEW OF THE PROPOSED SYSTEM

In this section an overview of the proposed Automatic Translation Arabic Sign to Arabic Text (ATASAT) System consists of five main stages Video and Images capture, Video and images processing, Hand Signs Construction, Classification and Text transformation and interpretation, cycle of proposed system is given in Figure 1.



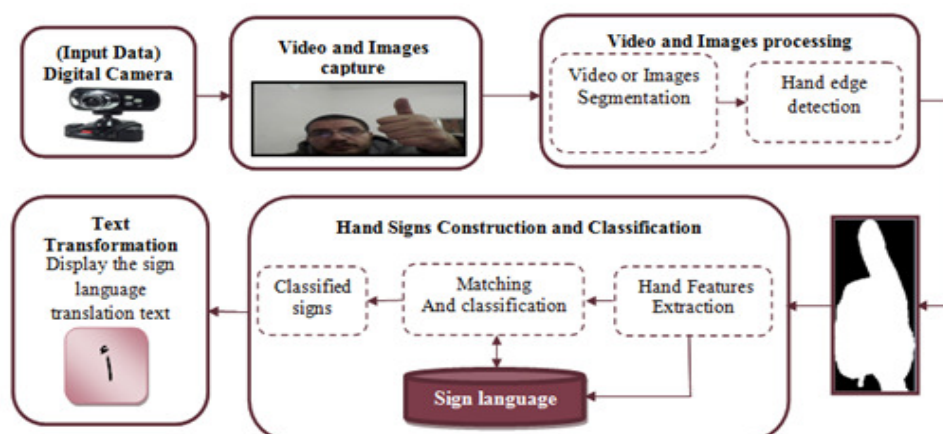


Figure. 1 The cycle of proposed system ATASAT System

The ATASAT system goes through many steps which are illustrated in the following flow chart shown in Figure. 2, the next sections explain in details every step.

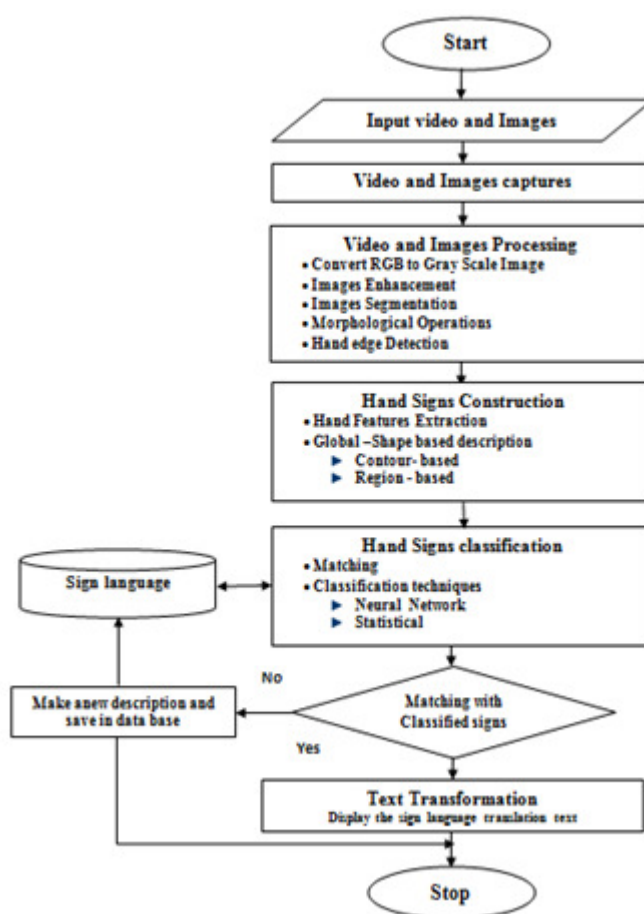


Figure.2 The ATASAT System flow chart

### 3.1 Video and Images capture

In first step in ATASAT system, the visual data is captured from the environment using input device like camera, Data entered for this stage give rise for gestures carried out by a number of indicators by wearing the glove dark color in different lighting environments with a light background (first data set) or without wearing a glove (natural color of the skin) with a dark background (second data set) so that it output of this stage is a set of colored image (RGB) representing the hand gestures corresponding to each one letter of the Arabic sign language the Figure 3 describes the image capture process.



Figure.3 The main screen of the proposed ATArSaT System

### 3.2 Video and images pre-processing

In the second step was the use of image processing techniques for initial processing of images captured gestures in order to hand detection of the image and isolate it from the background, it has been using a number of algorithms to discover it and choose the appropriate adjectives to describe and distinguish the form of a gesture hand and draw these qualities for use in computer training to get to know the letters, and the corresponding characters in the Arabic language.

Once the frame image is captured, it requires RGB to Gray conversion as our taken image is in RGB form. After the conversion, we are having Gray scale image and now use some different techniques to perform operations in order to enhance image, image segmentation and conduct morphological image processing to remove noise, and revealing important areas (hand shape) in the image.

To complete image processing operation, we designed a new algorithm is called Arabic Sign Hand Detection (ArSHD) this algorithm has been proven effective in the process of image processing and distinguish easily understanding and implementation, the implementation of this algorithm proposed several methods to detect a hand such as Sobel ,canny edge, thresholding

and Adjust the contrast method ,these are similar in technologies used to convert the image to gray, extract hand after detection of the image and these are different in technologies to enhance image, image segmentation and morphological image.

Use a combination of Captured images in the first stage around 700 pictures at the rate of 25 images for each letter an indicative of letters Arabic alphabet 28 and the type of images extracted is black and white Each image in the form of a gesture hand after the isolation, detection and the following figure is represents an example for each letter an indicative before the primary image processing process.[15]

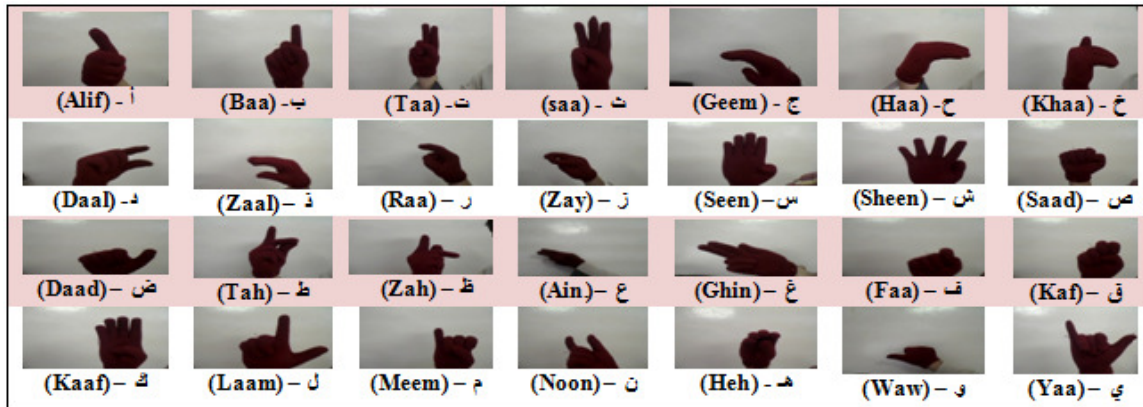


Figure 3 the ATASAT system dataset of alphabets letters before pre-processing

The following figure shows some images of characters indicative, and explains how to extract shape of the hand from the background in images indicative characters after the processing, which has to one of the previous methods such as Sobel, Canny, Threshold and adjust the contrast.

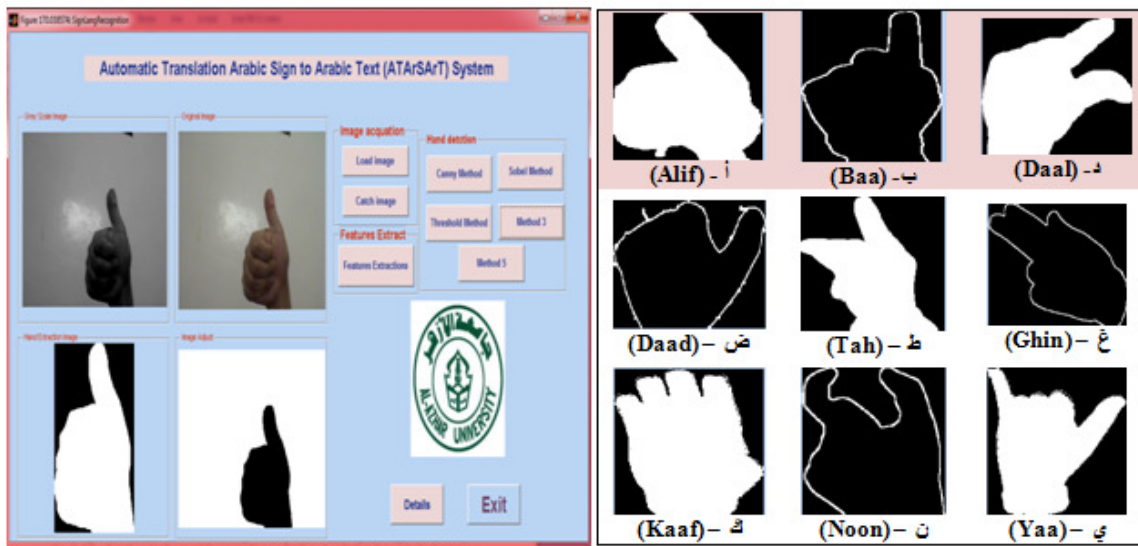


Figure.5 (a) the result of image processing (b) Examples of hand detection by different methods

### 3.3 Hand Signs Construction

In this step conducted Hand is described in the image caused by hand detection step according to one of the methods described based on the outer frame and the interior forms,, after that select the best features in the description gesture and distinguish them from other gestures, then, feature extraction, which have been selected for the processing of income data, which are then used in the training or testing processes.

Feature extraction stage depends on the success of hand gesture recognition in the previous stage; this is achieved through five basic steps data collection, features selection, select model of classification, classifiers training and evaluation. These steps can be summarized in two steps:

- a. Choose a good description of shape to be recognizable through analysis of data in this method and it collects data and choose the features of object must take into account the features of shape learned several things as follows[16]
  - The many small differences between the features of same classifier that is convergent.
  - Choose a smaller number of features for class to get a high percentage of accuracy
  - Discrimination varieties from each other by choosing the big differences between the features of each class and other.
  - Identify potential feature points for each class.
  - Identify attributes and features that are helpful for distinguish each class even if change his place, his change any factor of its factors, such as displacement, rotation, size, its coordinates and area.
- b. Choose the method for effective classification and there are several well-known to identify the patterns and the following table illustrates these methods and how to represent the patterns, recognition function, typical criterion that use with each method [17] , [18]

Table1. The important characteristics of the pattern recognition methods.

<b>Classification methods</b>	<b>Statistical</b>	<b>structural</b>	<b>Template matching</b>	<b>Artificial neural network</b>
<b>Representation</b>	Features	Primitives	Sample, pixel, curves	Samples, pixels, features
<b>Recognition function</b>	Discriminate function	Rules, grammar	Correlation, distance measures	Network function
<b>Typical Criterion</b>	Classification Error	Classification Error	Classification Error	Mean square Error

A lot of shape description and similarity measurement techniques have been developed in the past. A number of new techniques have been proposed in recent years. There are 3 main different classification methods [19]

1. Contour-based methods and region-based methods [20]. This is the most common and general classification and also is based on the use of shape boundary points as opposed to shape interior points.
2. Space domain and transform domain. Methods in space domain match shapes on point (or point feature) basis, while feature domain techniques match shapes on feature (vector) basis.
3. Information preserving (IP) and non-information preserving (NIP). IP methods allow an accurate reconstruction of a shape from its descriptor, while NIP methods are only capable of partial ambiguous reconstruction.

We have adopted in this paper on the first method to describe a shape that called Contour-based and region-based method, in order to get easy features account and fixed for rotation, displacement and change the size of the order to the same gesture.



Figure.6 The representation and description of shapes techniques

### 3.3.1 Hand Features Extraction Implementing

The proposed system success is heavily dependent on the choice of a good description of the gestures of the hand so that it becomes It can be distinguished from other gestures of hand.

As we mentioned in section 3.4 we use shape-based description in the present research Depending on the global method that calculate features through Contour –based or Region -based We have chosen a set of features for describing hand gesture.

- a) Object area (The number of pixels that compose the object (hand gesture))
- b) Object Perimeter (Minimum bounding rectangle and corresponding parameters)
- c) The center of gravity is also called centroid.
- d) Object Major Axis Length
- e) Object Minor Axis Length
- f) Bounding box
- g) Eight terminal points of the object coordinates (Extrema)

The following figure shows the illustration description features of hand shape that has been study [21]

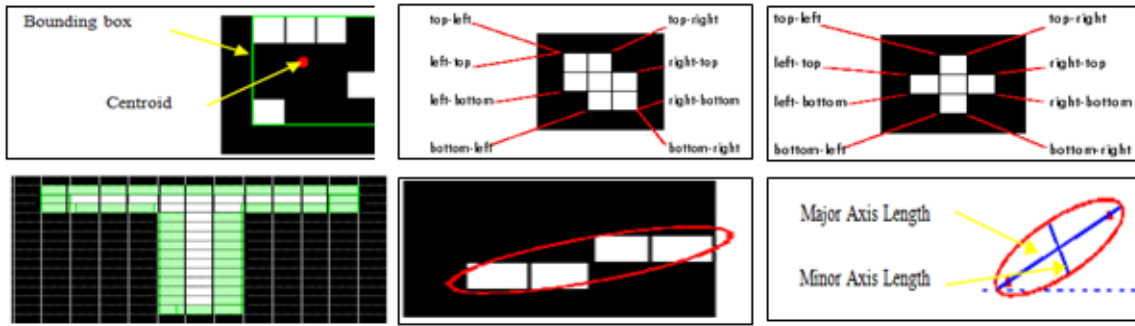


Figure.7 The Illustration of hand shape features

The calculation the studied features hand in a gesture of characters indicative table below illustrates some of the values of the features of these gestures.

Table 2 the calculated values for some features characters indicative gestures

The features		(Alif) - أ		(Seen) - س	
Object area		20195		30945	
Object Perimeter		904.4579		1.1479e+003	
Object centroid		CX=69.3843 CY=136.4365		CX=76.7294 CY=130.7228	
Object Major Axis Length		222.6721		254.5715	
Object Minor Axis Length		129.8120		165.1430	
Bounding box		x.width =136 , y.length =232		x.width=165 , y.length =236	
Extrema	Top –Left	54.5000	0.5000	76.5000	0.5000
	Top –Right	69.5000	0.5000	165.5000	0.5000
	Right-Top	136.5000	165.5000	165.5000	0.5000
	Right-Bottom	136.5000	232.5000	165.5000	236.5000
	Bottom-Right	136.5000	232.5000	165.5000	236.5000
	Bottom-Left	0.5000	232.5000	0.5000	236.5000
	Left- Bottom	0.5000	232.5000	0.5000	236.5000
	Left- Top	0.5000	167.5000	0.5000	55.5000

By study and calculate the features of all the gestures and the examples of it b we have noted that the traits important to distinguish the hand gesture, however, the group captured images are real image, and result in a significant difference to the values of features examples of one class per gesture, the interference with the values of the features between two classes of hand gestures for many reasons.:-

- Resize the image as a result of changing the distance from the camera.
- Change angle image capture for the same type of gesture.
- Rotation and displacement of the image, or one of its parts as a result of moving the hand.
- Distortions of some captured images gestures.
- The similarity of some of the characters gestures.
- Human differences in some methods of represent gestures per class.



Due to these reasons, we have the work of modifying the values of these features to improve the description of shaped gesture for the characters in order to obtain the highest accuracy in the classification stage.

For example, it can be modified a number of selected features by deleting endpoints (Bottom-right, Bottom-Left) because it is useless values to distinguish hand shape because it reflects where to place the palm of the hand.

The following table describes the modified features and methods of calculating the value of each it.

Table 3 The modified features to characters indicative gestures

Features	feature description	feature equation
F1	The ratio coordinates of the hand center of mass on the x-axis, to the width of around box by hand	$F1 = CX / x.width$
F2	The ratio coordinates of the hand centre of mass on the y-axis, to the length of around box by hand	$F2 = CY / y.length$
F3	The ratio between the width of around box by hand to the length of around box by hand	$F3 = x.width / y.length$
F4	The ratio pixels of between the object(Hand) area to the area of around box by hand	$F4 = \text{area (object)} / \text{area of bounding box}$
F5	The ratio between object(Hand) Perimeter to object (Hand) area	$F5 = \text{Perimeter (object)} / \text{area (object)}$
F6	The ratio between Object Major Axis Length to area of bounding box	$F6 = \text{Object Major Axis Length} / \text{area of bounding box}$
F7	The ratio between Object Minor Axis Length to area of bounding box	$F7 = \text{Object Minor Axis Length} / \text{area of bounding box}$
F8-F15	Extrema (Endpoints of hand shape)	$F8 = X \text{ Top -Left}$
		$F9 = X \text{ Top -Right}$
		$F10 = X \text{ Right-Top}$
		$F11 = Y \text{ Right-Top}$
		$F12 = X \text{ Right-Bottom}$
		$F13 = Y \text{ Right-Bottom}$
		$F14 = Y \text{ Left- Bottom}$
		$F15 = Y \text{ Left- Top}$

### 3.4 Hand Signs classification

In the classification process we will use two methods statistical classification and classification through neural networks in order to know the alphabet indicative fingers in the translation system into Arabic.

In the stage classification, choose one of the statistical classification algorithms or neural networks to design a classifier, and his teaching under the supervision of the training data (database formed in the previous phase of the system), to classify a new gesture (hand shape) is not represented in the training set to one of the existing categories.

The main objective of the ATASAT System is to create a matching between the indicative letters played by dumb people, and the signs of letters were stored in database signs that generated in the previous step.

This process will be designed classifiers uses one of the Education algorithms supervised, to do the task of classification a new character indicative to the corresponding product was his depending on digital data base resulting from the application extracting hand shape features (which was exposed when applying hand detection of the resulting images of the stage capture images database), which represent training data set for the classifiers and sometimes test data set.

In the implementation of the classification process based on statistical classification algorithms (C4.5, Naïve-Bayesian, K-NN), Multilayer perceptron (MLP) network algorithm Sequential Minimal Optimization ( SMO ) and Voting Feature Intervals (VFI) algorithm and compare the results in order to choose the best seed in terms of accuracy and speed.

Classification algorithms were applied using a WEKA program to compare the results with the results of classification using the ATASAT system to the same data base the results are presented in the fourth section of this paper.

### **3.5 Text transformation and interpretation**

Text transformation stage, to display the Arabic letter that corresponding and matching with the image of the character stored in the given data base.

According to the built-in database which contains a image signs features describing a pattern, the obtained matching pattern is presented through its text. Repeating this step for each input image and video frame allows integration between their descriptions, and leads to formulating the text representing the translation of the input image or video. The features obtained from the gesture sign language database for every key frame are concatenated to transform the image gestures into a text.

## **4. RESULT EVALUATION AND DISCUSSION**

Results of Implementing Classification in Weka Software has been conducting a number of experiments are described as follows:

### **4.1 Role of the methods of estimate error classification**

Two methods were used to estimate the error classification, the first method is holdout for the selection of samples for testing and training data sets and second method is cross validation used to measure performance evaluation.

The following tables presents the results of comparing the performance of the six classifiers studied in the research about the number of letters gestures classified correctly according to the group for the first dataset features images.



Table 4 Recognizing gestures letters accuracy by Holdout in dataset-1

Classifiers	Holdout Percentage				
	50%	66%	75%	80%	90%
<b>C4.5</b>	94	97.0588	97.7143	99.2857	98.5714
<b>SMO</b>	96.8571	98.3193	97.7143	99.2857	98.5714
<b>VFI</b>	95.7143	95.3782	98.2857	98.5714	98.5714
<b>MLP</b>	100	100	100	100	100
<b>K-NN</b>	100	100	100	100	100
<b>Naïve-Bayesian</b>	100	100	100	100	100

Table 5 Recognizing gestures letters accuracy by Cross-Validation in dataset-1

Classifiers	Cross-Validation value			
	Fold=5	Fold=10	Fold=15	Fold=20
<b>C4.5</b>	97.1429	98.1429	97.5714	98.1429
<b>SMO</b>	99.1429	99.1429	99.1429	99.1429
<b>VFI</b>	97.4286	97.2857	97.4286	97.4286
<b>MLP</b>	100	100	100	100
<b>K-NN</b>	99.8571	99.8571	99.8571	99.8571
<b>Naïve-Bayesian</b>	99.7143	99.7143	99.8571	99.8571

The following tables presents the results of comparing the performance of the six classifiers studied in the research about the number of letters gestures classified correctly according to the group for the second dataset features images of characters signs by using two these methods.

Table 6 Recognizing gestures letters accuracy by Holdout in dataset-2

Classifiers	Holdout Percentage				
	50%	66%	75%	80%	90%
<b>C4.5</b>	76.2857	80.6723	85.7143	86.4286	85.7143
<b>SMO</b>	70.5714	70.1681	78.2857	77.8571	82.8571
<b>VFI</b>	68.5714	69.3277	73.1429	73.5714	67.1429
<b>MLP</b>	86.8571	88.6555	88.00	92.1429	92.8571
<b>K-NN</b>	88.5714	91.1765	93.1429	95.7143	94.2857
<b>Naïve-Bayesian</b>	86.8571	84.4538	86.2857	83.5714	84.2857

Table 7 Recognizing gestures letters accuracy by Cross-Validation in dataset-2

Classifiers	Cross-Validation value			
	Fold=5	Fold=10	Fold=15	Fold=20
<b>C4.5</b>	81.00	83.857	84.8571	84.143
<b>SMO</b>	83.00	82.1429	83.00	81.8571
<b>VFI</b>	72.4286	70.2857	71.7143	70.7143
<b>MLP</b>	90.143	91.000	92.00	90.429
<b>K-NN</b>	92.286	93.000	93.00	93.286
<b>Naïve-Bayesian</b>	87.286	88.000	87.00	88.286

#### 4.2 Role value of the neighbouring K in the performance of the k-NN algorithm

In this section presents the results performance of classifier K-NN according to the value of neighbour by choose values has been increasingly for a neighbour from 1 to 10, the following table describe this results.

Table 8 The role of K value in the performance of the K-NN algorithm

K- Neighbour value	K- value			
	K=1	K=3	K=5	K=10
The proportion of correct classified examples –data set -1	99.8571	99.7143	99.143	98.5714
The proportion of correct classified examples –data set -2	93.00	88.5714	86.8571	81.00

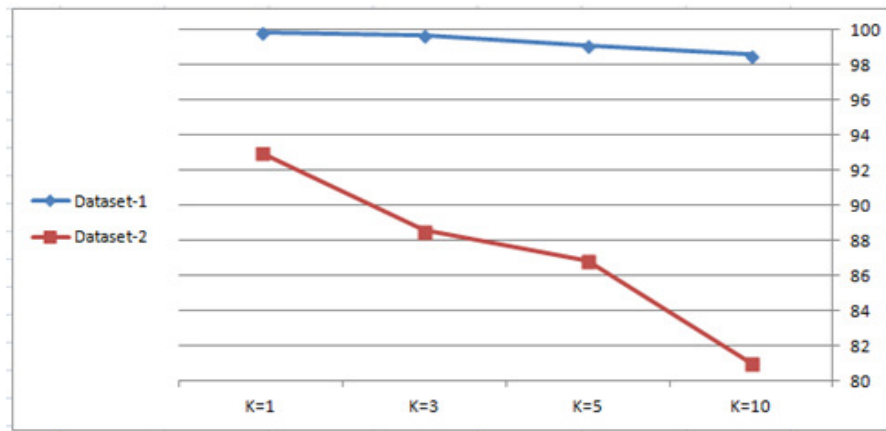


Figure.8 A graph describe role of K value in the performance of the K-NN algorithm

## 5. CONCLUSIONS AND FUTURE DIRECTIONS

### 5.1 Conclusions

The results obtained after performing the cross validation method (fold=10) shows much better performance compared to all other existing algorithms taken for the comparison. The value of k used in our approach (KNN) is very important and it has a significant role in the classification process. Its highly recommend to use the value of k=1 which gives the best performance results. When holdout percentage values for the image features which is included in dataset2 is performed, our algorithm (K-NN) is giving better results than the other approaches (C4.5, SMO, VFI, MLP and Naïve-Bayesian).

Same as the other approaches like MLP and Naïve-Bayesian, K-NN also use images contains glove weared gestures with white background and one-signer. When the holdout percentage values for different percentages ( 50%,66%,75%,80% and 90%) are calculated for the featured images of dataset-1, our algorithm along with MLP and Naïve-Bayesian shows far better performance compared to other classifiers. Results are shown in Table 4.

Two data sets are introduced for the analysis, containing the image features for gesture recognition. Six classifiers based on statistical and neural network approaches are used for the comparison. MLP, K-NN and Naïve-Bayesian are based on neural topology, C4.5, SMO, VFI are based on statistical approach. We used to estimate the error classification is holdout for the selection of samples for testing and training data sets and second method is cross validation used to measure performance evaluation. In cross validation method the experimental results shows that the MLP classifier gives better accuracy results using two datasets (dataset-1 and dataset-2). In Estimation error classification method, K-NN classifier has remarkable holdout value when compared with all other classifiers.

## 5.2 Future works and directions

The researcher hopes to use an initial model in educational programs for the deaf by Computer, As well as a means to communicate with the deaf and understand their language. So far in this area single signer means of gesture recognition is taken into account. High researches scopes are there if double hand gesture recognition approaches are put forwarded. Sign language to voice recognizer methods can also be implemented; even mobile phone platforms can also be used for this. Social networking like instant messaging services for the hearing disabled people can also be possible, more researches need to be carried out in this regard, so that it can make a revolutionary change in the lives of people who all are deaf and dump.

## ACKNOWLEDGEMENTS

We thank Bijoy Babu of College of engineering, King Khalid University, Kingdom of Saudi Arabia for his assistance and valuable comments that greatly improved the manuscript.

## REFERENCES

- [1] Ravikiran J; Kavi Mahesh, "Finger Detection for Sign Language Recognition", Proceedings of the International Multi Conference of Engineers and Computer Scientists, vol.1, 2009.
- [2] Joyeeta Singha, Karen Das, " Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 2,2013
- [3] Tabata Y, Kuroda T. Finger spelling recognition using distinctive features of hand shape. In: 7th ICDVRAT with Art Abilitation. Maia, Portugal: 2008:287-292.
- [4] Halawani S. Arabic sign language translation system on mobile devices. IJCSNS. 2008; 8(1):251.
- [5] Tsukada K, Yasumura M. Ubi-Finger: a Simple Gesture Input Device for Mobile and Ubiquitous Environment. Journal of Asian Information, Science and Life (AISL). 2004; 2(2): 111-120.
- [6] Khalid Alvi A, Azhar M, Usman M, Mumtaz S, Rafiq S, Rehman R, et al. Pakistan Sign Language Recognition Using Statistical Template Matching. In: Proceedings of world academy of science, engineering and technology. Rome, Italy: 2005;3:52-55.
- [7] AI-Buraiky S. Arabic sign language recognition using an instrumented glove. Master Thesis, King Fahd University of Petroleum & Minerals, Saudi Arabia. 2004.

- [8] Nashwa El-Bendary, Hossam M. Zawbaa, Mahmoud S. Daoud, Aboul Ella Hassanien: "ArSLAT: Arabic Sign Language Alphabets Translator", International Journal of Computer Information Systems and Industrial Management Applications, Vol.3, 2011,
- [9] Samaa M. Shohieb, Hamdy K. Elminir, A.M. Riad" Signs World Atlas; a benchmark Arabic Sign Language database" Journal of King Saud University - Computer and Information Sciences Volume 27, Issue 1, January 2015, Pages 68–76
- [10] Feris R, Turk M, Raskar R, Tan K, Ohashi G. Exploiting depth discontinuities for vision-based finger spelling recognition. In: Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04) Vol 10. IEEE Computer Society; 2004.
- [11] Tanibata N, Shimada N, Shirai Y. Extraction of hand features for recognition of sign language words. In: The 15th International Conference on Vision Interface. 2002:391-398.
- [12] Mohandes M. Arabic sign language recognition. In Proceedings of the Int. Conf. on Imaging Science. 2001. Las Vegas, Nevada, USA: CSREA PRESS; 2001 ;(1):25-28.
- [13] Mohandes M. Automatic Translation of Arabic Text to Arabic Sign Language. ICGST International Journal on Artificial Intelligence and Machine Learning, AIML. 1996; 6(4):15-19.
- [14] Foong OM, Low TJ, Wibowo S. Hand Gesture Recognition: Sign to Voice System. International Journal of Electrical, Computer and Systems Engineering (IJECSSE). 2009; 3(4):198-202.
- [15] Abdelmoty M. Ahmed, Reda Abo Alez , Muhammad Taha , Gamal Tharwat. Propose a New Method for Extracting Hand using in the Arabic Sign Language Recognition (Arslr) System. International Journal of Engineering Research & Technology (IJERT). Volume. 4 - Issue. 11 , November – 2015 e-ISSN: 2278-0181
- [16] Priyanka Sharma, Manavjeet Kaur. Classification in Pattern Recognition: A Review. International Journal of Advanced Research in computer Science and Software Engineering (IJCARCSSE). Volume 3, Issue 4, April 2013- ISSN: 2277 128X
- [17] Mingqiang Yang, Kidiyo Kpalma, Joseph Ronsin. A Survey of Shape Feature Extraction Techniques. Peng-Yeng Yin. Pattern Recognition, IN-TECH, pp.43-90, 2008.
- [18] Jie Liu, Jigui sun Pattern recognition: An overview, 1st edn. Wang. College of computer science & technology, china. IJCSNS, International Journal of Computer Science & Network Security. VOL.6. NO.6. June 2006.
- [19] Navdeep Kaur, Usvir Kaur, Survey of Pattern Recognition Methods, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 2, February 2013.
- [20] Kidiyo Kpalma and Joseph Ronsin, An Overview of Advances of Pattern Recognition Systems in Computer Vision Source: Vision Systems: Segmentation and Pattern Recognition, ISBN 987-3-902613-05-9, IETR, UMRCNRS 6164.
- [21] <http://ch.mathworks.com/help/images/ref/regionprops.html?requestedDomain=www.mathworks.com>

# E- LEARNING IN NIGERIA : CURRENT IMPLEMENTATIONS AND CHALLENGES

Oluwatosin Temitope Aderinola<sup>1</sup>, Ojonukpe Egwuche<sup>2</sup> and  
Abimbola Afolayan<sup>3</sup>

<sup>1,2,3</sup>Department of Computer Science,  
The Federal University of Technology, Akure, Ondo State, Nigeria

<sup>1</sup>toosinaderinola@yahoo.com

<sup>2</sup>osegwuche@futa.edu.ng

<sup>3</sup>ahafolayan@futa.edu.ng

## ABSTRACT

*E-learning is a concept which came to limelight about two decades ago in Nigeria. It has gone through several metamorphoses since then ranging from open skepticism to acceptance. It was accepted by most people because the percentage of Conventional Universities in Nigeria to the number of candidates seeking admission into these Universities were getting small every year and such Conventional Universities cannot cope with the pressure from the candidates. This is the main reason why many people embraced E- learning as option and a welcome supplement to normal university learning. This study is designed to evaluate the current implementation of E-learning in Nigeria and the current challenges militating against the full implementation of E-learning as an option in Nigeria.*

## KEYWORDS

*E-learning, Conventional Universities, Candidates.*

## 1. INTRODUCTION

Education has become increasingly important for occupational and personal success, therefore the demand for higher education is becoming greater with each passing day. Higher Institutions have found themselves in a position which they are unable to admit the greater percentage of the students who qualify for admission in their various Institutions. They are therefore seeking ways to offer their courses to a larger group of people. One promising way is to apply the method of distance learning where one can take courses, complete them and collect the certificate without actually going to the classroom to attend lectures.

The Internet has become one of the vital ways to make available resources for research and learning for both teachers and students to share and acquire information [1]. Technology-based e-learning encompasses the use of the internet and other important technologies to produce materials for learning, teach learners, and also regulate courses in an organization [2].

There has been extensive debate about a common definition of the term e-learning. Existing definitions according to [3] tend to reveal the specialization and interest of the researchers. E-learning as a concept covers a range of applications, learning methods and processes [4]. It is therefore difficult to find a commonly accepted definition for the term e-Learning, and according to [8] and [3], there is even no common definition for the term. [7] also made a comment on these inconsistencies by saying that there may be as many definitions of the term e-learning as there are academic papers on the subject [3] in trying to find a common meaning of the term e-learning went on to ask the following questions: Is e-learning an on-line coursework for students at a distance? Does it mean using a virtual learning environment to support the provision of campus based education? Does it refer to an on-line tool to enrich, extend and enhance collaboration? OR is it a totally on-line learning or part of blended learning? [3].

Some of the definitions of the term e-learning as given by different researchers and institutions are reviewed below. In some definitions e-Learning encompasses more than just the offering of wholly on-line. [8] noted that e-Learning has transformed from a fully-online course to using technology to deliver part or all of a course independent of permanent time and place. Also [6] describes, e-Learning as the use of new multimedia technologies and the Internet to increase learning quality by easing access to facilities and services as well as distant exchanges and collaboration. The following are also different definitions of e-learning. E-learning refers to the use of information and communication technologies to enable the access to online learning/teaching resources. In its broadest sense, [5] defined E-learning to mean any learning that is enabled electronically. They however narrowed this definition down to mean learning that is empowered by the use of digital technologies.

E-learning is a method of learning which employs the use of digital Technology to learn. It means the application of internet technology to 'receive lectures', perform assignments, and write tests and examinations. The students are graded by the Computer and there is no bias. However, in Nigeria as a developing Country, we still have a long way to go as far as E-learning is concerned. Some of these challenges can be resolved or made better and the others will improve in synchronization with development in the Country.

E-learning in Nigeria was adopted about a decade ago, at first it was a welcome distraction from the normal classroom lectures. But as time went on, the numerous challenges such as power issue, internet connectivity issues, the system began to encounter at first skepticism then outright antagonism from the would be users.

The objective of this research is to examine the current level of implementation of distance e-learning in Nigeria and the factors militating against its total challenge, the attendant challenges and possible ways of distance learning as related to e-learning and its attendant challenges using National Open University of Nigeria (NOUN) as a case study.

NOUN was established initially in 1983 and suspended in 1984. Since it's resuscitation in 2001, this School has taken Tertiary Education beyond the boundaries of Conventional Classroom method.

## 2. CURRENT IMPLEMENTATION OF E-LEARNING IN NIGERIA

Today, technology enhanced learning, including distance and online instruction, is recognized as a viable tool necessary for preparing citizens to participate in the technologically driven global environment. A multidisciplinary approach to online pedagogical research recognizes the value of technology enhanced teaching and learning as critical in the mix of diverse strategies. Its centrality in the global marketplace has been enhanced by a new culture shared by many educators.

At present, the method employed in National Open University of Nigeria (NOUN) is through internet-generated mode where the course materials would be uploaded on the Institution website and the students are expected to log on with their unique user id and password and download the materials meant for them, study it, do assignments and interact with their lecturers and classmates via the internet.

Likewise writing of Examination in NOUN is through the internet. Taking online exams in NOUN is a simple process. It does not require extensive computing skills. The basic skill of reading text on the computer screen and clicking on or typing the right answers is what is required to take exam.

The steps to follow are outlined below:

1. Open your web-browser – (Internet explorer or Mozilla Firefox ....)
2. Log on to the National Open University of Nigeria website; [www.nou.edu.ng](http://www.nou.edu.ng)
3. Click on the e-exam link
4. You will be directed to a page where you will have to choose among the following options:
  - i. Instructions on taking the Electronic Test
  - ii. Log in to take an exam or practice test
  - iii. Download student help
5. Click on the “Log in to take an exam or practice test” link to take a practice test or to take an actual examination. Take note that you will only be able to take an actual examination if it has been scheduled for you.
6. A login screen similar to that in figure 1 below displays
7. Enter your Matriculation Number into the user login section.
8. Enter your Matriculation Number as password (lower case only).

9. The System may ask you to verify that your name and other details have been entered correctly. If so check through to see that the details are correct then click on OK.
10. You will be taken to the system homepage
11. Click on the class (course) whose exam you want to take under the class name
12. Click on the exam you want to take under the assignment name header to begin the examination.
13. Once you have finished answering the questions click on grade. You will be warned if there are unanswered questions.
14. You will get an instant feedback similar to figure 5.
15. To view your results, at the top of the page click View Details.
16. Once you have finished reviewing your results, click Quit & Save.
17. To log out of Maple T.A., click the log out link.

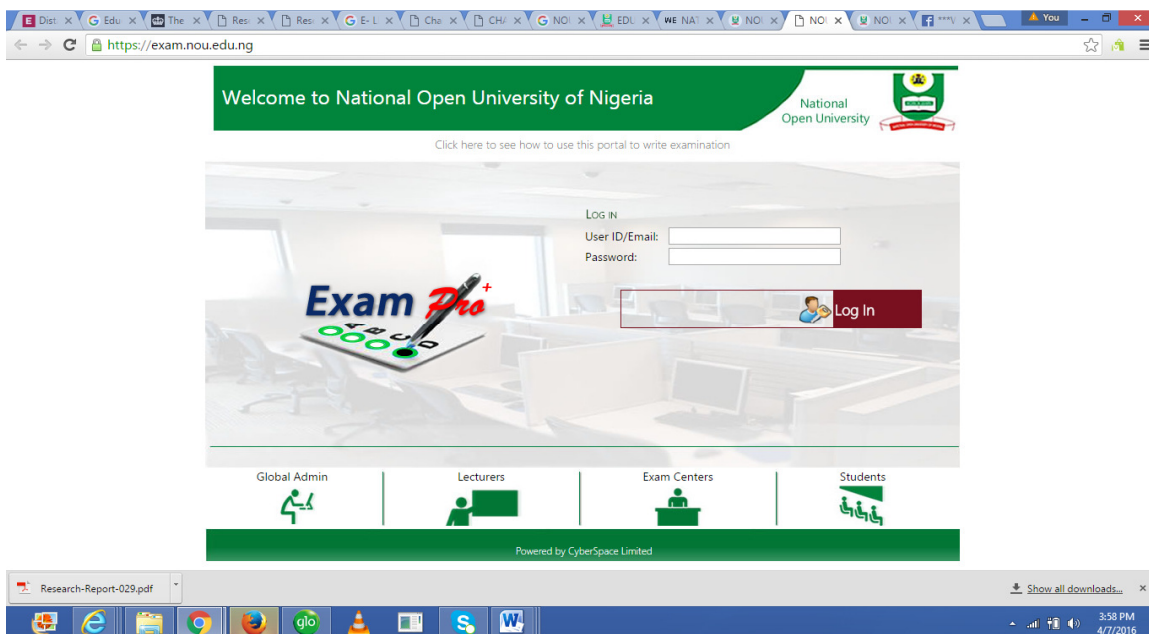


Figure 1 : Login screen to MapleTA – NOUN's e-examination platform

### 3. ISSUES AGAINST E- LEARNING

#### (a) Lack of monitoring

The usual method is for students to log on to the school's website, download and read the course materials. But logging on might not necessarily means that the student is doing any useful thing



on the internet and neither is there any measure to monitor the student at any time unlike the conventional classroom practice. Loopholes exist in the system as per lack of monitoring. There is actually no way to monitor when a student should log on or what the student should do when he logs on to the institution website.

#### **(b) Stack overload**

The normal method employed as at the present day is the upload of study materials onto the net. Pulling from various centers makes downloading extremely slow and this can be frustrating to the student who barely has enough money for buying the airtime to download course materials. Sometimes, some might end up not being able to download within the time allocated for a particular internet usage.

#### **(c) Discrimination**

One major issue against distance learning in Nigeria is Discrimination against graduates of University that utilizes e- learning. Employers, People and Institutions believe that the quality of Education obtained from E-learning is low when compared to graduates that receive their lecture in normal classroom settings. Hence graduates are discriminated against by prospective employers, Other Universities and some Government bodies.

Graduates of e-learning seeking Post Graduate admissions into Conventional Universities are discriminated against because it is believed that the standard of education they've had is too low for them to adapt. However, this discrimination is more intense because of the competition for admissions into such post graduate programs.

A particular instance of this discrimination is the inability of the Nigerian Law School to admit Law Graduates of the National Open University of Nigeria. They claimed that Law is a course which cannot be learned effectively outside normal classrooms. Questions of whether or not the graduates would pass or fail in law school were not even considered.

### **4. CHALLENGES OF DISTANCE LEARNING IN NIGERIA**

#### **(a) Light/ Power**

Nigeria is a developing Country which has difficulties providing stable power supply to citizens. This situation has grown worse over the years and with the advent of democracy in the year 1999, stable power supply has become a mirage. With Power supply being a mirage, e-learning is almost impossible to run or sustain. The Computer system with the issue of power supply will break down sooner than expected and most internet facilities rely on power supply to function well. The main private means of power generation in Nigeria is the gasoline generator which has several hazards to its use. Some of the hazards of generator usage is noise and air pollution. Another hazard is the cost implication.

#### **(b) Internet**

Internet facilities are not available to the common man in Nigeria. The readily available options of Internet that we have are too expensive and not affordable to the common man. In Nigeria

where the average man lives on less than two dollar per day, It costs an average of 5 dollar per day to be on the internet per hour. The average Nigeria cannot afford to be on the internet for 10 hours at a stretch.

### **(c) Economy**

The poor economic situation of Nigeria is another factor militating against distance learning in Nigeria. The average Nigerian lives on less than 2 dollar a day and as such most cannot afford to own a Computer System, pay for the needed internet facilities and the other financial implications.

### **(d) Computer Illiteracy Issue:**

An average Nigeria doesn't have a Computer System not to think of knowing how to use the Computer. Candidates of the Computer Based Examinations are disadvantaged because of their inability to use the system effectively.

## **5. BENEFITS OF E-LEARNING**

- (a) Significant reduction or elimination of costs associated with instructor fees and materials.
- (b) Reduction of learning time and the amount of time employees are away from their jobs.
- (c) Increased retention by users and enhanced hands-on application to the job over traditional training methods.
- (d) Instruction and progress are managed within the framework of a company web portal.
- (e) Short videos or hands-on exercises offer practice and assessment activities designed to confirm that a learner has mastered the performance objectives of the course or lesson.
- (f) A user's progress is automated and can be monitored by or shared with one's supervisor
- (g) Interactivity engages users, pushing them rather than pulling them through training.
- (h) Advanced learners may choose to speed through a course and bypass instruction that is redundant.
- (i) Beginners or slower learners may determine their own pace and progress, eliminating frustration with themselves, their fellow learners, and the subject matter.
- (j) Knowledge is cumulative; lessons build upon one another and may be taken in any order.
- (k) Anytime-anywhere learning can greatly increase knowledge retention. – Learners may also view or print simple, one-page "quick start" Job Aids that have stepby-step procedures and graphic workflow charts illustrating tasks to be performed
- (l) Testing is self-paced.

## 6. CONCLUSION

There is a greater potential for sharing information through the Internet than through other means of transmitting and receiving information. Not everyone can be well suited to distance learning programs. Successful participants must be highly motivated and self-disciplined. Because the course may be unmonitored, the learners themselves have full responsibilities for proceeding with the course and evaluating their mastery of a skill or subject.

Although distance education is very flexible and convenient, it still cannot provide the 'college experience.' Working with other learners, being part of a total educational environment, and collaborate closely with academic mentors is still valuable to many learners. In addition, some courses cannot be taught on the Internet and this causes some limitations of distance learning.

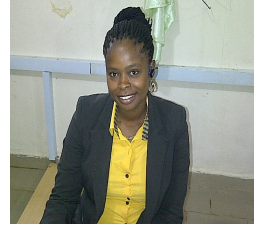
E-learning is a concept which has come to stay. The internet is gaining ground every day and the whole world is gradually becoming a global village. The issues of discrimination should be addressed by Government so that we don't make life difficult for graduates who have gone to a University that utilizes e-learning by making them feel inferior in the society.

## REFERENCES

- [1] Richard, H., & Haya, A. (2009) "Examining student decision to adopt web 2.0 technologies: theory and empirical test", *Journal of computing in higher education*, 21(3), 183-198.
- [2] Fry, K. (2001), "E-learning markets and providers: some issues and prospects", *Education Training*, 233-239.
- [3] Dublin, L. (2003), "If you only look under the street lamps.....Or nine e-Learning Myths", *The eLearning developers journal*.<http://www.eLearningguild.com>
- [4] Rossi.P.G. (2009), "Learning environment with artificial intelligence elements", *Journal of e-learning and knowledge society*, 5(1), 67-75.
- [5] Abbad, M. M., Morris, D., & de Nahlik, C. (2009), "Looking under the Bonnet: Factors Affecting Student Adoption of E-Learning Systems in Jordan", *The International Review of Research in Open and Distance Learning*.
- [6] European Commission (2001). "The eLearning Action Plan: Designing tomorrow's education", <http://www.elearningeuropa.info>. *International Journal of Education and Research* Vol. 2 No. 12 December 2014
- [7] Holmes, B. & Gardner, J. (2006), "E-Learning: Concepts and Practice, London", SAGE Publications.
- [8] Oblinger, D. G., & Hawkins, B. L. (2005) "The myth about E-learning. Educause review", *International Journal of Education and Research* Vol. 2 No. 12 December 2014

**AUTHORS**

Oluwatosin Temitope Aderinola is a Female Technologist from Nigeria. She obtained an HND in Computer Science from the Federal Polytechnic, Ado Ekiti, Nigeria in 2003 and a Post Graduate Diploma from the Federal University of Technology, Akure, Nigeria from the Federal University of Technology, Akure. She currently works in the Computer Science Department of the Federal University of Technology, Akure. Her research interest includes Computer Networks, Human Computer Interaction, Artificial Intelligence and Wireless Applications.



Egwuche, Ojonukpe Sylvester obtained B.Sc (Hons) in Computer Science from Kogi State University, Anyigba, Nigeria in 2008 and M.Tech. degree in Computer Science from the Federal University of Technology, Akure, Nigeria in 2014. He is currently a Research Fellow and a P.hD student in the department of Computer Science, Federal University of Technology, Akure, Nigeria. He has published many papers in peer-reviewed national and international journals. His research interests among others include Access Control, Power Optimization in Wireless Sensor Networks, Mobile Agents Technologies and Embedded Computing.



Abimbola Helen Afolayan is an Assistant Lecturer in the Computer Science Department of the Federal University of Technology, Akure, Nigeria



## **AUTHOR INDEX**

- Abdel Wahab Turkmani* 97
- Abdelkrim Haqiq* 01
- Abdelmoty M.Ahmed* 109
- Abimbola Afolayan* 123
- AissaTaibi* 11
- Ameur Touir* 75
- Aram AlSedrani* 75
- Baghdad Atmani* 11
- Basma Boukenze* 01
- Chengliang Huang* 57
- Fang Wang* 57
- Fatih Korkmaz* 29, 37
- Gamal Tharwat* 109
- Hajar Mousannif* 01
- Izdihar Al-muwaffaq* 67
- Majed El Helou* 97
- Muhammad Taha* 109
- Ojonukpe Egwuche* 123
- Oluwatosin Temitope Aderinola* 123
- Rawan Chanouha* 97
- Reda Abo Alez* 109
- Samer Charbaji* 97
- Sulaiman S. AlDahri* 45
- Van-Hung Le* 85
- Xiao-Ping Zhang* 57
- Zeki Bozkus* 67