

International Journal of Interactive Multimedia and **Artificial Intelligence**

June 2018, Vol V, Number 1, ISSN: 1989-1660



*“Artificial intelligence is growing up fast,
as are robots whose facial expressions
can elicit empathy and make
your mirror neurons quiver”*

Diane Ackerman

<http://ijimai.unir.net>

IMAI RESEARCH GROUP COUNCIL

Director - Dr. Rubén González Crespo, Universidad Internacional de La Rioja (UNIR), Spain

Office of Publications - Lic. Ainhoa Puente, Universidad Internacional de La Rioja (UNIR), Spain

Latin-America Regional Manager - Dr. Carlos Enrique Montenegro Marín, Francisco José de Caldas District University, Colombia

EDITORIAL TEAM

Editor-in-Chief

Dr. Rubén González Crespo, Universidad Internacional de La Rioja – UNIR, Spain

Associate Editors

Dr. Óscar Sanjuán Martínez, CenturyLink, USA

Dr. Jordán Pascual Espada, ElasticBox, USA

Dr. Juan Pavón Mestras, Complutense University of Madrid, Spain

Dr. Alvaro Rocha, University of Coimbra, Portugal

Dr. Jörg Thomaschewski, Hochschule Emden/Leer, Emden, Germany

Dr. Vicente García Díaz, Oviedo University, Spain

Dr. Carlos Enrique Montenegro Marín, Francisco José de Caldas District University, Colombia

Dr. Elena Verdú, Universidad Internacional de La Rioja (UNIR), Spain

Dr. Manju Khari, Ambedkar Institute of Advanced Communication Technologies and Research, India

Editorial Board Members

Dr. Rory McGreal, Athabasca University, Canada

Dr. Anys Yazidi, Oslo Metropolitan University, Norway

Dr. Nilanjan Dey, Techo India College of Technology, India

Dr. Abelardo Pardo, University of Sidney, Australia

Dr. Hernán Sasastegui Chigne, UPAO, Perú

Dr. Lei Shu, Osaka University, Japan

Dr. Ali Selamat, Malaysia Japan International Institute of Technology, Malaysia

Dr. León Welicki, Microsoft, USA

Dr. Enrique Herrera, University of Granada, Spain

Dr. Hamido Fujita, Iwate Prefectural University, Japan

Dr. Francisco Chiclana, De Montfort University, United Kingdom

Dr. Luis Joyanes Aguilar, Pontifical University of Salamanca, Spain

Dr. Ioannis Konstantinos Argyros, Cameron University, USA

Dr. Ligang Zhou, Macau University of Science and Technology, Macau

Dr. Juan Manuel Cueva Lovelle, University of Oviedo, Spain

Dr. Pekka Siirtola, University of Oulu, Finland

Dr. Francisco Mochón Morcillo, National Distance Education University, Spain

Dr. Peter A. Henning, Karlsruhe University of Applied Sciences, Germany

Dr. Vijay Bhaskar Semwal, Indian Institute of Information Technology, Dharwad, India

Dr. Miroslav Hudec, University of Economics of Bratislava, Slovakia

Dr. Walter Colombo, Hochschule Emden/Leer, Emden, Germany

Dr. Javier Bajo Pérez, Polytechnic University of Madrid, Spain

Dr. Jinlei Jiang, Dept. of Computer Science & Technology, Tsinghua University, China

Dr. B. Cristina Pelayo G. Bustelo, University of Oviedo, Spain

Dr. Cristian Iván Pinzón, Technological University of Panama, Panama

Dr. José Manuel Sáiz Álvarez, Nebrija University, Spain

Dr. Masao Mori, Tokyo Institute of Technology, Japan

Dr. Daniel Burgos, Universidad Internacional de La Rioja - UNIR, Spain

Dr. JianQiang Li, NEC Labs, China

Dr. David Quintana, Carlos III University, Spain

Dr. Ke Ning, CIMRU, NUIG, Ireland

Dr. Alberto Magreñán, Real Spanish Mathematical Society, Spain

Dr. Monique Janneck, Lübeck University of Applied Sciences, Germany
Dr. Carina González, La Laguna University, Spain
Dr. David L. La Red Martínez, National University of North East, Argentina
Dr. Juan Francisco de Paz Santana, University of Salamanca, Spain
Dr. Héctor Fernández, INRIA, Rennes, France
Dr. Yago Saez, Carlos III University of Madrid, Spain
Dr. Andrés G. Castillo Sanz, Pontifical University of Salamanca, Spain
Dr. Pablo Molina, Autonomía University of Madrid, Spain
Dr. Giuseppe Fenza, University of Salerno, Italy
Dr. José Miguel Castillo, SOFTCAST Consulting, Spain
Dr. Sukumar Senthilkumar, University Sains Malaysia, Malaysia
Dr. Juan Antonio Morente, University of Granada, Spain
Dr. Holman Diego Bolívar Barón, Catholic University of Colombia, Colombia
Dr. Sara Rodríguez González, University of Salamanca, Spain
Dr. José Javier Rainer Granados, Universidad Internacional de La Rioja - UNIR, Spain
Dr. Elpiniki I. Papageorgiou, Technological Educational Institute of Central Greece, Greece
Dr. Edward Rolando Nuñez Valdez, Open Software Foundation, Spain
Dr. Luis de la Fuente Valentín, Universidad Internacional de La Rioja - UNIR, Spain
Dr. Paulo Novais, University of Minho, Portugal
Dr. Giovanni Tarazona, Francisco José de Caldas District University, Colombia
Dr. Javier Alfonso Cedón, University of León, Spain
Dr. Sergio Ríos Aguilar, Corporate University of Orange, Spain
Dr. Mohamed Bahaj, Settati, Faculty of Sciences & Technologies, Morocco
Dr. Madalena Riberio, Polytechnic Institute of Castelo Branco, Portugal
Dr. Edgar Henry Caballero Rúa, Inforfactory SRL, Bolivia
Dr. José Miguel Rojas, University of Sheffield, United Kingdom

Editor's Note

THE International Journal of Interactive Multimedia and Artificial Intelligence (IJIMAI) provides an interdisciplinary forum in which scientists and professionals share their research results and report new advances on tools that use AI with interactive multimedia techniques.

Brief of selected papers of IJIMAI are mentioned below:

Real Time Facial Expression Recognition Using Webcam and SDK Affectiva (Slovakia) by Magdin and Prikler: the proposed system uses neural network algorithm for classification and recognizes 6 (respectively 7) facial expressions, namely anger, disgust, fear, happiness, sadness, surprise and neutral [1].

Self-Organized Hybrid Wireless Sensor Network for Finding Randomly Moving Target in Unknown Environment (India) by Nighot et al. This research paper proposes a solution for searching randomly moving target in unknown area using Mobile sensor nodes and combination of both Static and Mobile sensor nodes. Using algorithms like MSNs Movement Prediction Algorithm (MMPA), Leader Selection Algorithm (LSA), Leader's Movement Prediction Algorithm (LMPA) and follower [2].

Novel Clustering Method Based on K-Medoids and Mobility Metric (Morocco) authors (Hamzaoui et al.) propose a new algorithm of clustering based on new mobility metric and K-Medoid to distribute nodes into several clusters, to avoid the problem of negative influence of MANETS on the performance of QoS [3].

Spectral Restoration Based Speech Enhancement for Robust Speaker Identification (Pakistan) By Saleem and Tareen employed and evaluated Minimum Mean-Square-Error Short-Time Spectral Amplitude to improve performance of the speaker identification systems in presence of background noise. The identification rates are found to be higher after employing speech enhancement algorithms [4].

Hybrid Model for Passive Locomotion Control of a Biped Humanoid: The Artificial Neural Network Approach (India) by Raj et al. led to the observation that base model with learning based compensation enables the biped to better adapt in a real walking environment, showing better limit cycle behaviors [5].

Users Integrity Constraints in SOLAP Systems. Application in Agroforestry (Algeria) by Charef and Djamila, proposed a system for the implementation of user integrity constraints in SOLAP, namely "UIC-SOLAP" [6].

A Study on Persuasive Technologies: The Relationship between User Emotions, Trust and Persuasion (Malaysia) by Ahmad and Ali, concluded that emotions have significant effect on trust, and effect of emotions on persuasion using the persuasive technology was mediated by trust [7].

EEG Signal Analysis of Writing and Typing between Adults with Dyslexia and Normal Controls (Australia) authors (Perera et al.) study revealed that the extra difficulties seen in individuals with dyslexia during writing and typing compared to normal controls are reflected in the brainwave signal patterns [8].

Applying Bayesian Regularization for Acceleration of Levenberg-Marquardt based Neural Network Training (Malasia y Kazakhstan) by Suliman and Omarov: describe a method of applying Bayesian regularization to improve Levenberg-Marquardt (LM) algorithm and make it better usable in training neural networks. Results showed 98.8% correct classification when run on test samples [9].

Object Detection and Tracking using Modified Diamond Search Block Matching Motion Estimation Algorithm (India), Samdurkar et

al. observed that the MDS (modified diamond search pattern) performs better than DS (diamond search) and CDS (cross diamond search algorithms) on average search point and average computation time [10].

Spatial and Textural Aspects for Arabic Handwritten Characters Recognition (Morocco), Boulid, et al. purpose was the recognition of handwritten Arabic characters in their isolated form, and had 2.82% error rate [11].

A Novel Smart Grid State Estimation Method Based on Neural Networks (Egypt) Abdel-Nasser et al. presents a novel method called SE-NN (state estimation using neural network) for smart grid state estimation using artificial neural networks (ANNs) and is a very fast tool to estimate voltages and re/active power loss with a high accuracy compared to the traditional methods [12].

Conceptual model development of big data (Malaysia) Adrian et al. aims to identify and analyze the affecting factors of BDA implementation and to propose a conceptual model for effective decision-making through BDA implementation assessment [13].

MSA for Optimal Reconfiguration and capacitor Allocation in Radial/Ring Distribution Networks (Egypt y Japan) Mohamed et al. presents a hybrid heuristic search algorithm called Moth Swarm Algorithm (MSA) in the context of power loss minimization of radial distribution networks (RDN). Results state that MSA can achieve optimal solutions for losses reduction and capacitor locations with finest performance compared with many existing algorithms [14].

Dr. Manju Khari

REFERENCES

- [1] M. Magdin and F. Prikler, "Real Time Facial Expression Recognition Using Webcam and SDK Affectiva", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 7-15, 2018.
- [2] M. Nighot, A. Ghatol and V. Thakare, "Self-Organized Hybrid Wireless Sensor Network for Finding Randomly Moving Target in Unknown Environment", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 16-28, 2018.
- [3] Y. Hamzaoui, M. Amnai, A. Choukri and Y. Fakhri, "Novel Clustering Method Based on K-Medoids and Mobility Metric", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 29-33, 2018.
- [4] N. Saleem and T. G. Tareen, "Spectral Restoration Based Speech Enhancement for Robust Speaker Identification", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 34-39, 2018.
- [5] M. Raj, V. Bhaskar-Semwal and G. C. Nandi, "Hybrid Model for Passive Locomotion Control of a Biped Humanoid: The Artificial Neural Network Approach", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 40-46, 2018.
- [6] A. B. Charef and H. Djamila, "Users Integrity Constraints in SOLAP Systems. Application in Agroforestry", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 47-56, 2018.
- [7] W. N. W. Ahmad and N. M. Ali, "A Study on Persuasive Technologies: The Relationship between User Emotions, Trust and Persuasion", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 57-61, 2018.
- [8] H. Perera, M. F. Shiratuddin, K. W. Wong and K. Fullarton. "EEG Signal Analysis of Writing and Typing between Adults with Dyslexia and Normal Controls", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 62-67, 2018.
- [9] A. Suliman and B. S. Omarov, "Applying Bayesian Regularization for

- Acceleration of Levenberg-Marquardt based Neural Network Training”, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 68-72, 2018.
- [10] A. S. Samdurkar, S. D. Kamble, N. V. Thakur, and A. S. Patharkar, “Object Detection and Tracking using Modified Diamond Search Block Matching Motion Estimation Algorithm”, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 73-85, 2018.
- [11] Y. Boulid, A. Souhar, and M. M. Ouagague, “Spatial and Textural Aspects for Arabic Handwritten Characters Recognition”, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 86-91, 2018.
- [12] M. Abdel-Nasser, K. Mahmoud and H. Kashef, “A Novel Smart Grid State Estimation Method Based on Neural Networks”, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 92-100, 2018.
- [13] C. Adrian, R. Abdullah, R. Atan and Y. Y. Jusoh, “Conceptual model development of big data”, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 101-106, 2018.
- [14] E. A. Mohamed, A.-A.A. Mohamed and Y. Mitani, “MSA for Optimal Reconfiguration and Capacitor Allocation in Radial/Ring Distribution Networks”, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, no. 1, pp. 107-122, 2018.

TABLE OF CONTENTS

EDITOR'S NOTE	4
REAL TIME FACIAL EXPRESSION RECOGNITION USING WEBCAM AND SDK AFFECTIVA	7
SELF-ORGANIZED HYBRID WIRELESS SENSOR NETWORK FOR FINDING RANDOMLY MOVING TARGET IN UNKNOWN ENVIRONMENT.....	16
NOVEL CLUSTERING METHOD BASED ON K-MEDOIDS AND MOBILITY METRIC.....	29
SPECTRAL RESTORATION BASED SPEECH ENHANCEMENT FOR ROBUST SPEAKER IDENTIFICATION	34
HYBRID MODEL FOR PASSIVE LOCOMOTION CONTROL OF A BIPED HUMANOID: THE ARTIFICIAL NEURAL NETWORK APPROACH.....	40
USERS INTEGRITY CONSTRAINTS IN SOLAP SYSTEMS. APPLICATION IN AGROFORESTRY	47
A STUDY ON PERSUASIVE TECHNOLOGIES: THE RELATIONSHIP BETWEEN USER EMOTIONS, TRUST AND PERSUASION.....	57
EEG SIGNAL ANALYSIS OF WRITING AND TYPING BETWEEN ADULTS WITH DYSLEXIA AND NORMAL CONTROLS	62
APPLYING BAYESIAN REGULARIZATION FOR ACCELERATION OF LEVENBERG-MARQUARDT BASED NEURAL NETWORK TRAINING	68
OBJECT DETECTION AND TRACKING USING MODIFIED DIAMOND SEARCH BLOCK MATCHING MOTION ESTIMATION ALGORITHM.....	73
SPATIAL AND TEXTURAL ASPECTS FOR ARABIC HANDWRITTEN CHARACTERS RECOGNITION.....	86
A NOVEL SMART GRID STATE ESTIMATION METHOD BASED ON NEURAL NETWORKS	92
CONCEPTUAL MODEL DEVELOPMENT OF BIG DATA ANALYTICS IMPLEMENTATION ASSESSMENT EFFECT ON DECISION-MAKING	101
MSA FOR OPTIMAL RECONFIGURATION AND CAPACITOR ALLOCATION IN RADIAL/RING DISTRIBUTION NETWORKS	107

OPEN ACCESS JOURNAL

ISSN: 1989-1660

COPYRIGHT NOTICE

Copyright © 2018 UNIR. This work is licensed under a Creative Commons Attribution 3.0 unported License. Permissions to make digital or hard copies of part or all of this work, share, link, distribute, remix, tweak, and build upon ImaI research works, as long as users or entities credit ImaI authors for the original creation. Request permission for any other issue from support@ijimai.org. All code published by ImaI Journal, ImaI-OpenLab and ImaI-Moodle platform is licensed according to the General Public License (GPL).

<http://creativecommons.org/licenses/by/3.0/>

Real Time Facial Expression Recognition Using Webcam and SDK Affectiva

M. Magdin*, F. Prikler

Constantine the Philosopher University in Nitra, Faculty of Natural Sciences, Department of Computer Science, Tr. A. Hlinku 1, 949 74 Nitra (Slovakia)

Received 23 June 2017 | Accepted 3 November 2017 | Published 9 November 2017



ABSTRACT

Facial expression is an essential part of communication. For this reason, the issue of human emotions evaluation using a computer is a very interesting topic, which has gained more and more attention in recent years. It is mainly related to the possibility of applying facial expression recognition in many fields such as HCI, video games, virtual reality, and analysing customer satisfaction etc. Emotions determination (recognition process) is often performed in 3 basic phases: face detection, facial features extraction, and last stage - expression classification. Most often you can meet the so-called Ekman's classification of 6 emotional expressions (or 7 - neutral expression) as well as other types of classification - the Russell circular model, which contains up to 24 or the Plutchik's Wheel of Emotions. The methods used in the three phases of the recognition process have not only improved over the last 60 years, but new methods and algorithms have also emerged that can determine the ViolaJones detector with greater accuracy and lower computational demands. Therefore, there are currently various solutions in the form of the Software Development Kit (SDK). In this publication, we point to the proposition and creation of our system for real-time emotion classification. Our intention was to create a system that would use all three phases of the recognition process, work fast and stable in real time. That's why we've decided to take advantage of existing Affectiva SDKs. By using the classic webcam we can detect facial landmarks on the image automatically using the Software Development Kit (SDK) from Affectiva. Geometric feature based approach is used for feature extraction. The distance between landmarks is used as a feature, and for selecting an optimal set of features, the brute force method is used. The proposed system uses neural network algorithm for classification. The proposed system recognizes 6 (respectively 7) facial expressions, namely anger, disgust, fear, happiness, sadness, surprise and neutral. We do not want to point only to the percentage of success of our solution. We want to point out the way we have determined these measurements and the results we have achieved and how these results have significantly influenced our future research direction.

KEYWORDS

Facial Expression, Emotion Recognition, Real Time Recognition, Face Detection, Classification.

DOI: 10.9781/ijimai.2017.11.002

I. INTRODUCTION

At present, an issue of great interest is the correspondence between information reflected in and conveyed by the human face and the person's concurrent emotional experience.

There are several recent studies reporting findings that facial signs and expressions can provide insights into the analysis and classification of emotional states [27], [28].

Many research works have been developed which are related to recognizing the actions of human. Human actions can be recognized in the form of skeleton [31], silhouettes, bio-metric gait [26] and in the form of images [17]. Visual surveillance technique is used for identifying packages of human actions [13]. Researchers used different techniques to recognize human actions such as hierarchical probabilistic approach [1], multi-modality representation of joints [18], [15].

Already long ago Darwin argued that facial expressions are universal, i.e. most emotions are expressed in the same way on the

human face regardless of race or culture [5], [11].

According to Darwin, facial expression can influence the course of communication with another person when we express our emotions, opinions, as well as our intentions in an efficient way. In addition, in his studies of human behavior he explicitly states that such expressions also provide information about the cognitive state of a person. This includes conditions such as boredom, stress, confusion and others.

Darwin's work is remarkable because at the time when he lived he had already assumed that emotional expressions are multimodal patterns of behavior, and thus established his own detailed descriptions of the manifestations of more than 40 emotional states [14].

Although face expressions doesn't necessarily express emotions (e.g. the physiological character of the face after stroke), when performing face detection (if we are talking about the intelligent systems research area), most of the authors of professional publications refer to the classification of facial features that was introduced by Ekman. This is because unlike other emotions (in his period Darwin identified more than 40 expressions), they are easily and unambiguously identifiable. Classification model of 6 emotions by Ekman [6]:

1. happiness,

* Corresponding author.

E-mail address: mmagdin@ukf.sk

2. sadness,
3. surprise,
4. fear,
5. anger,
6. disgust.

Later Ekman [7] added a neutral expression to the previously mentioned six emotions. This method of classifying emotions has gained a relatively high popularity. The main advantage is that facial expressions related to basic emotions are easily recognizable and they can be described even by ordinary people – non psychologists. This model of emotions is currently the most widespread one, used for their detection, and in the last four decades a number of facial recognition and facial parts recognition studies have been developed. Subsequent Russell's critical point of view on the issue (compiling the circular model) in 1994 raised questions about the degree of universality in recognizing these emotional (prototypical) facial states and what these expressions actually signify [24].

In psychological literature, emotions are defined as targeted responses to relevant individuals' suggestions, which include behavior, physiological and experiential components. Various techniques have been developed to enable computers to understand human affects or emotions, in order to predict human intention more precisely and provide better service to users to enhance user experience [16]. If we compare individual approaches, we find that ubiquitous hardware, available in most computer systems (classic webcam), allows to obtain important information about human facial expression. Webcam has become a de facto device thanks to the popularity of interactive social networking applications. A human can oftentimes deduce the emotion of a person sitting in front of a webcam to a certain degree of accuracy. Recent research (in last decade) in video processing and machine learning has demonstrated that human affects can be recognized via webcam video, noticeably via human facial features [32] and eye gaze behaviors [12], [16].

In the publication we will point out our designed and implemented solution, which with the help of a webcam can identify the subject's emotions in real time. When designing and building our system, we encountered many issues that will be described in detail.

II. RELATED WORK

Systems that could not only detect the face in the picture, but also evaluate the emotional states of the subject from individual facial features, have been the subject of research of psychologists as well as pedagogical and information technology specialists for a very long time.

We can say that their interest in this area dates back to the 1960s. One of the first documented studies is Woodrow W. Bledsoe's work for Panoramic Research, Inc. However, this software had been developed for face recognition and facial parts recognition for the US Department of Defense. Therefore, the work is not well accessible to the public. However, with the available materials we know that during 1964-1965, the Bledsoe, Helen Chan and Charles Bisson triad were able to create the first algorithm that ensured facial recognition. Bledsoe was the first to propose a semi-automatic system that worked on the principle of selecting and choosing the coordinates of some parts of the face (those were entered by a human operating a computer), and then these were used as recognition co-ordinates. As early as 1964, Bledsoe has described most of the problems faced even by contemporary facial recognition systems and their parts: change in illumination, turning the head in the recognition process, inappropriate expressions such as grimaces, ageing, and the like. That's why he was looking for other procedures and found that some parts of the face change in adulthood only for a minimum - ear size or distance between eye pupils. His

approaches and techniques were later used in Bell Laboratories, where Goldstein, Harmon and Lesk described a vector that can be used to identify 21 subjective functions such as ear lobe, eyebrow size, nose length, which was later the basis for designing other techniques (Extractions) used to recognize the face and its parts [23].

The first functional full-automatic system was implemented by Kanade in 1973. His algorithm was able to automatically evaluate 16 different facial parameters by comparing extractions obtained from the computer with extractions delivered by humans. His program, however, did correct identification in the range of only 45-75%.

In 1978, Ekman and Friesen developed Facial Action Coding System (FACS) to encode facial expressions, where the facial movements are described by a set of action units (AU) [8]. Each AU is based on the affinity of the muscles. This mimic coding system is manually executed using a set of rules. Entries are static images of the mimetic, often representing the peak of the expression. This process is time-consuming.

Ekman's work inspired many researchers (L. Davis, T. Otsuk, J. Ohyu, Y. Yacoob and others) to analyze facial expressions using an image and video processing. By tracking facial features and measuring the amount of movement on the face, he tried to sort out different facial expressions into individual categories [23].

Further work from this period drew the very ideas from the system developed by Kanade and tried to improve the results he had previously achieved, using methods that provide measurements using subjective functions. An example may be Nixon's work (Nixon in 1985), which proposed a method of measuring the geometric distribution of the eyes. This suggested approach has been improved several times by other authors (for example - Kobayashi & Hara in 1997) [23].

Pantic and Rothkrantz in their work provide a detailed, in-depth recapitulation of many researches on automatic face recognition from 1980 to 2000 [22].

As time went on, the methods for the individual phases of the recognition process (detection, extraction and classification) developed complete solutions (FILTWAM), the so-called SDK (Software Development Kit), various libraries (OpenCV, FaceTracker) and support api appeared [35].

III. AUTOMATIC SYSTEM FOR CLASSIFICATION EMOTION – THEORETICAL BACKGROUND

From the historical point we can say that there are various methods for face recognitions. Such a system for automatic facial expression recognition has to deal with such problems: facial detection and location in a chaotic environment, extraction of facial features and correct classification of the expression on face [34].

Drawing upon various published works [4], [10], [31], [2], [30], main system elements from various similar system have been identified. The main elements of each system to evaluate facial expression are:

1. facial detection,
2. extraction of face features,
3. classification, expression evaluation.

Our main goal was to design and implement a system for real-time recognition and evaluation of human emotions based on a facial expression using a webcam so that the resulting application provides the ability to analyze the collected data. In the future, such a system would be included (e.g. in the form of a plugin) in the LMS Moodle learning environment, so we could better understand the problems that students face while studying. We have already had experience with the design and implementation of such a system, since starting from 2015, the Department of Computer Science, FNS CPU in Nitra, has

been undertaking intensive research in this field. It resulted in a fully functional system which was described in detail from the perspective of the classification success degree in the International Journal of Interactive Multimedia and Artificial Intelligence [19]. However, it was impossible for this system to recognize and evaluate emotional states of more than one person simultaneously with 1 camera. The system worked Real Time, but stored the photos on the disk and there, on the basis of the ViolaJones detector (Haar Cascade Principle), detected the face in the image, extracted the facial features and eventually classified the emotions. In this way, we achieved relatively high accuracy (around 78%), but the system was slow, and the photos took up disk space (even if automatic photo deletion after the correct classification of the expression was set). That's why we decided to design and implement a completely new system that would be faster and work more accurately. At this point, we have got help from establishing the partnership with IBM Slovakia, which resulted in the realization of a project with a name *Pilot project with UKF Nitra University: Modelling the behavior of users based on data mining with a support of IBM Bluemix*.

To determine the level of face detection success and to determine the emotional state of the user with a help of our solution and SDK from Affectiva, we have established the following research question:

RQ1: Can our software unambiguously detect a face, regardless of the distance of a person, surrounding light conditions, rotation of the face or its part, or other non-standard conditions?

Based on the established research question, which is also a research problem for us, it is also possible to think about the hypothesis:

H1: The ability of our software to detect faces in the image and evaluate the user's emotional state is completely independent of the distance of the person being shot, the surrounding light conditions and the rotation of the person's face.

The principles of emotional states analysis through facial expression have a common basis with the principles used in face recognition. The whole analysis process can be divided into three consecutive phases [9]. Those are the detection and localisation of the face in the image with a complex background, including its normalisation, the extraction of appropriate symptoms describing the given expression in the face and finally the determination of the corresponding expression based on the extracted symptoms using the selected classifier. The choice of the observed symptoms (classifier) is always dependent on the approach used and the desired outputs.

As mentioned at the beginning of the publication, we do not need to use only the well-known procedures and methods for the detection, extraction, and classification phases when designing and implementing the system, but we can use available libraries, apis and SDKs that are verified. We decided to use the Affectiva SDK. The advantage of using such a solution is the fact that we do not have to deal with the training dataset in the final phase of the recognition process (classification).

Affectiva is a company that has been operating since 2009 at MIT University, Media Lab, and its main focus is "to teach computers to understand human emotions". Their solutions are mainly used in the advertising sector. During their research, for example, they exposed people to video ads and, with the help of the webcam, captured the people's reactions and analysed the emotions they were experiencing. Based on such observations, they could then evaluate responses to the same ad in different demographic backgrounds. Affectiva also introduced a wearable bio-sensor that can reveal emotions based on skin input.

For the first time, the company released the Software Development Kit for Developers in 2015. The current version of SDK 3.4.1 for the Windows operating system was released in April 2017. The SDK aims to detect facial expressions and classify emotions associated with them. The SDK includes components that enable integration with the

NET framework and C++ written applications.

The recommended hardware requirements are as follows:

- Processor: 2GHz,
- RAM: 1 GB,
- Free disk space: 950 Mb.

SDK supports Windows 7 and above. For proper operation, Visual C ++ Redistributable runtime for VisualStudio 2013 and Microsoft .NET Framework 4.0 is required. The Affectiva SDK supports multiple platforms, including, in addition to Windows, Android (Java), iOS (Objective-C), Linux (C ++), MacOS (Objective-C), Unity and Web.

SDK solution from Affectiva works on the principle of the extended facial expression model at the following points:

1. Facial detection and detection of significant facial features,
2. Extraction of facial features,
3. Classification of facial action points,
4. Emotional prototype modeling using EMFACS.

Facial detection and extraction of significant facial features - performed using the Viola-Jones Detection algorithm [32]. Landmark detection is then applied to each rectangle bounding the face by 34 identified landmarks. If the benchmark detection confidence is below the threshold, the boundary is ignored. Extracted facial points, head rotation, and a center point between the eyes for each face are displayed in the SDK.

Facial actions - Histogram of Oriented Gradient (HOG) is extracted from the image area of interest defined by the face orientation points. Support Vector Machine Algorithm (SVM), trained on 10,000 manually encoded face images collected from around the world, is used to generate a score from 0 to 100 for each face action. For details on the training and testing system, see [25].

Emotional expressions - in this case, anger, fear, joy, sadness, surprise, and contempt are based on combinations of face action. This encoding was based on the emotional EMFACS face coding system. Emotional expressions get a score of 0 (the expression is missing) to 100 (the expression is present).

Computer classification of facial expressions requires a large amount of data to reflect the diversity of conditions occurring in the real world. Public databases (listed in the Success Rating section) help speed up research progress by providing a reference source for researchers. They represent a complexly identified set of valid spontaneous face reactions recorded in the natural environment over the Internet. In the case of Affectiva's SDK for data collection, viewers watched one of three deliberately entertaining ads on the "Super Bowl" over the Internet while uploading them using a webcam. They answered three evaluation questions about their experience. In addition, some audiences have given their consent to make their data publicly shared with other researchers. This subgroup consists of 242 video footage (168,359 images) recorded in real conditions. The database is completely marked with the following data:

1. Single-frame labels that represent 10 symmetric FACS action units, 4 asymmetric (single-sided) FACS action units, 2 head movements, smile, general expressivity, face and gender facial extraction faults.
2. Placement of 22 automatically recognized landmarks.
3. Participants' own answers to questions: Did you see the video before? Do you want to watch stimulation videos again?
4. Index of basic performance of detection algorithms on this database.

This data is available for distributing to researchers through the Internet [20], [21].

Our user interface (Fig. 1) is very simple, but it is sufficient for the determination of emotional states of the subject.



Fig. 1. The preview of the application's user interface (own creation).

The system is capable, under normal lighting conditions, of detecting a face correctly at a distance of 7.5 meters from the recording device, while setting the parameter `Affdex.FaceDetectorMode` to `SMALL_FACES`. The minimum distance for successful face detection is 20 cm.

All controls are located in the main window (Fig. 2). The largest part of the content is the webcam output. In addition to the "Start" and "Stop" buttons, the following elements were added to the beta testing experience:

- "Reset" button - to resume image capture in the event of an error.
- "Show / Hide Metric" button - allows you to view and turn off real-time emotional status evaluation.
- "Show / Hide Points" button - to switch between displaying and turning off the display of points that show facial features.
- "Statistics" button - this element allows you to look at the statistics that the program recorded during the shooting.
- "Exit" button - to turn off the program.

The program is set to start scanning immediately after startup and initial initialization. The graphical interface was defined using XAML (eXtensible Application Markup Language). In the picture above there is the application's user interface in the paused state.



Fig. 2. Demonstration of the application's user interface (own creation).

IV. EXPERIMENT – EVALUATION OF SUCCESS RATE (WITH CONTINUOUS DISCUSSION)

One of the most important aspects of creating a new system to detect or recognize emotions is a database choice that would be used for new system testing. If only one database was used for each research, then new system testing, compared to other top-class systems and efficiency

testing, would be rather trivial tasks [3].

Therefore, we have used various databases that are generally respected by professionals [36]:

1. Bao Face Database,
2. CMU/VASC Image Database,
3. Caltech Faces 1999 Database,
4. NIST Mug-shot Images Database,
5. Yale Face Database,
6. Vision Group of Essex University Face Database.

The testing environment (Fig. 3) was constructed in the following way: on one side, a notebook with software designed to classify emotions was placed at a distance of 1 meter (which is the standard distance of the subject by the webcam). On the other side, a monitor was placed on which test photographs were displayed.



Fig. 3. Testing environment.

Our proposed solution was subjected to evaluation using individual databases. Since each of these databases specializes in different conditions, such as light, turning of the subject's face, the amount of simultaneously captured faces in the image, the result of such an experiment should be a comprehensive evaluation not only of the success rate of face detection but also of the subsequent classification of the emotional state. The evaluation was carried out gradually through the mentioned databases.

Bao Face Database - This database includes lots of face images, mostly of people from Asia (Fig. 4). It is divided into two subdirectories: single face pictures and picture with many faces. The total number of images with one face is 149. Photographs are created in low resolution in bad light conditions. The resolution is most often from 149x207 to 192x273 pixels with DPI 96. Out of this number of images our software detected all 149 faces in the image, but could not classify emotional status with 21 photos. The subjects tested on these 21 photographs had a face rotated by more than 15 degrees. We can say that the software in this case worked with 86% detection success.



Fig. 4. Selection of photos on which a face was detected but the emotional state was not classified.

The second part of the database contained 221 photos. There was a lot of people in each photo. Our goal was to determine if the software can detect faces in a group. Photos are created in different resolutions (from 348x432 to 800x600, DPI 72) in poor lighting conditions. Out of a total of 221 photos, the software was able to detect all the faces on each of the photos and at the same time classify the emotional state of the subject. So he worked with 100% success (Fig. 5).



Fig. 5. Example of face detection and emotional status classification in a picture with multiple people (own creation).

CMU/VASC Image Database - The image dataset is used by the CMU Face Detection Project and is provided for evaluating algorithms for detecting frontal views of human faces. This particular test set was originally assembled as part of work in Neural Network Based Face Detection. It combines images collected at CMU and MIT. The database contains 4 subdirectory: *newtest*, *rotated*, *test*, *test-low*.

The *newtest* subdirectory contains - 65 pictures of movies and fairytales, as well as covers from CDs, cartoons or cards. Images are created in different resolutions (from, for example, 60x75 to the high resolution 1280x1024, DPI 600) in different lighting conditions. The software was able to detect the face and evaluate the emotional state even in non-standard images (Fig. 6).



Fig. 6. Face detection of nonstandard subjects.

We only noticed the problem with two pictures where the system did not detect the face and could not determine the emotional state of the subject (Fig. 7).

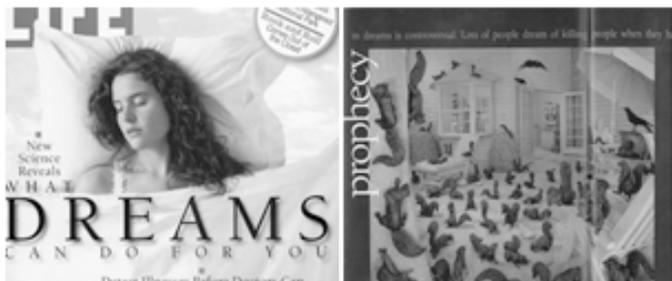


Fig. 7. Problem of face detection.

So the software was able to detect the subject's face with a 96% success rate.

The directory *rotated* contains 50 different images (resolution from 109x145 to 2615x1986, DPI 600). These pictures were not rotated. Rotated was the face of a scanned subject. The system, out of a total 50 shots, was able to detect face and classify emotional status in only two shots (Fig. 8). The software only worked with 4% success.



Fig. 8. Determination of the slightly inclined face of the subject and determination of the emotional state.

The directory *test* contains 42 various pictures (resolution from e.g. 142x192 to 716x684, DPI 600). System out of a total of 48 frames failed to detect face and classify emotional status in 3 frames (Fig. 9). The software in this case worked with a 93.75% success rate.



Fig. 9. Demonstration of problem detection of the subject's face.

The directory *testlow* contains 23 various pictures (resolution from e.g. 160x100 to 628x454, DPI 600). The system was able to correctly detect the face and classify the emotional state for each frame (and non-standard, Fig. 10). The software was 100% successful.



Fig. 10. Demonstration of subject's non-standard face detection.

Caltech Faces 1999 Database – Frontal face dataset. Collected by Markus Weber at California Institute of Technology. Contains 450 face images in great resolution e.g. 896 x 592 pixels, bad lights conditions, Jpeg format. This folder contains 27 jpg files of unique people under with different lighting, expressions and backgrounds (Fig. 11). On all 450 photos, the system was able to detect the face and classify the emotional state of the subject, so we achieved 100% success.



Fig. 11. Demonstration of subject's face detection under poor lighting conditions.

NIST Mug-shot Images Database – contains 2 subdirectory: *single* and *multiple*. For a better idea of the internal structure of these subdirectories, we present the chart in Fig. 12.

The subdirectory *f1_p1* contains 24 subdirectories under the label *sing01* to *sing24*. Each of these subdirectories contains various count of pictures:

- sing01* – 68 frontal face pictures, 58 right rotated face pictures, 10 left rotated face pictures
- sing02* – 50 frontal face pictures, 42 right rotated face pictures, 8 left rotated face pictures
- sing03* – 50 frontal face pictures, 46 right rotated face pictures, 4 left rotated face pictures
- sing04* – 50 frontal face pictures, 47 right rotated face pictures, 3 left rotated face pictures
- sing05* – 50 frontal face pictures, 44 right rotated face pictures, 6 left rotated face pictures
- sing06* – 50 frontal face pictures, 48 right rotated face pictures, 2 left rotated face pictures
- sing07* – 50 frontal face pictures, 46 right rotated face pictures, 6 left rotated face pictures
- sing08* – 50 frontal face pictures, 48 right rotated face pictures, 2 left rotated face pictures
- sing09* – 50 frontal face pictures, 40 right rotated face pictures, 10 left rotated face pictures
- sing10* – 50 frontal face pictures, 45 right rotated face pictures, 5 left rotated face pictures
- sing11* – 50 frontal face pictures, 43 right rotated face pictures, 7 left rotated face pictures
- sing12* – 50 frontal face pictures, 40 right rotated face pictures, 10 left rotated face pictures
- sing13* – 50 frontal face pictures, 45 right rotated face pictures, 5 left rotated face pictures
- sing14* – 50 frontal face pictures, 48 right rotated face pictures, 2 left rotated face pictures
- sing15* – 50 frontal face pictures, 43 right rotated face pictures, 7 left rotated face pictures
- sing16* – 50 frontal face pictures, 42 right rotated face pictures, 8 left rotated face pictures
- sing17* – 50 frontal face pictures, 45 right rotated face pictures, 5 left rotated face pictures
- sing18* – 50 frontal face pictures, 44 right rotated face pictures, 6 left rotated face pictures
- sing19* – 50 frontal face pictures, 42 right rotated face pictures, 8 left rotated face pictures

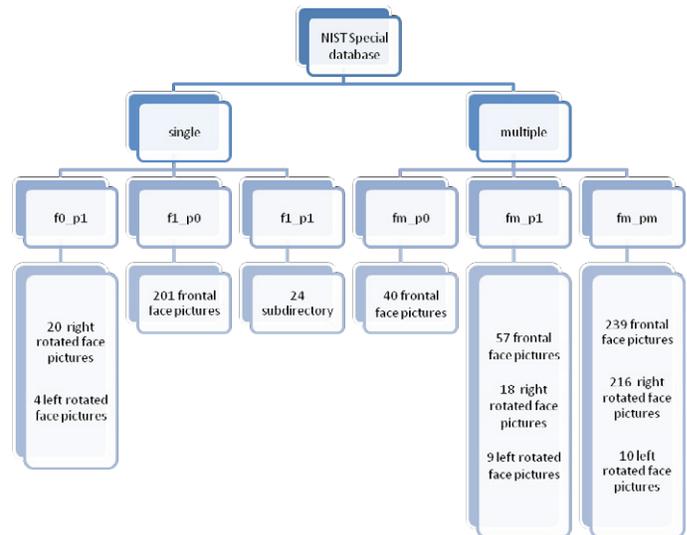


Fig. 12. Internal structure of NIST Special database.

- sing20* – 50 frontal face pictures, 45 right rotated face pictures, 5 left rotated face pictures
- sing21* – 50 frontal face pictures, 45 right rotated face pictures, 5 left rotated face pictures
- sing22* – 50 frontal face pictures, 41 right rotated face pictures, 9 left rotated face pictures
- sing23* – 50 frontal face pictures, 56 right rotated face pictures, 4 left rotated face pictures
- sing24* – 49 frontal face pictures, 45 right rotated face pictures, 4 left rotated face pictures

With a detailed evaluation of the images from this database we found that the software can detect the face and also determine the emotional state of the subject only in frontal face (100% success rate). In the case of rotated the head to the right or to the left, the software was unable to detect the face and thus not even determine the emotional state (0% success rate).

Yale Face Database – this directory database contains 3 subdirectories: *centered*, *faces* and *rotated*. This database concentrates to the illumination variations. All pictures in subdirectory with a name *centered* are in gray scale, resolution is 195x231 pixels and 600 DPI. Count of pictures is 165. These pictures are divided to the 3 sections: *centerlight*, *glasses/noglasses* and type of *emotional state*. All pictures in subdirectory with a name *faces* is in gray scale, resolution is 320x234 pixels and 600 DPI. Count of pictures is 165. The face of the subject itself is not rotated, rotated is only picture.

With a detailed evaluation of the images from this database we found that the software can detect the face and also determine the right emotional state of the subject in all cases from 3 subdirectories: *centered*, *faces* and *rotated*. The software worked with 100% success rate (Fig. 13).

Vision Group of Essex University Face Database - this robust database is created by Dr. Libor Spacek. Database contains 4 subdirectories: *face94*, *face95*, *face96* and *grimace*. The subdirectory *face94* is divided to 3 subdirectories with name *female*, *male* and *malestaff*. Subdirectory with name *female* contains 20 subdirectories. In each of these subdirectories there are 20 face pictures. We tested therefore 400 face pictures. Subdirectory with name *male* contains 113 subdirectories. In each of these subdirectories there are 20 face pictures. We tested therefore 2260 face pictures. Subdirectory with name *malestaff* contains 20 subdirectories. In each of these subdirectories there are 20 face pictures. We tested therefore 400 face pictures. With

a detailed evaluation of the images from this database we found that the software can detect the face and also determine the emotional state of the subjects in all cases. The software worked with 100% success rate (Fig. 14).



Fig. 13. Face Detection of a subject with a classified emotional state.



Fig. 14. Demonstration of subject's face detection (directory face94).

The subdirectory face95 contains 72 subdirectories. In each of these subdirectories there are 20 face pictures. We tested therefore 1440 face pictures. These pictures are not divided to the subdirectory female or male. Our software solution recognised all pictures also with slight tilt of head subject (to 15%). The software worked with 100% success rate (Fig. 15).

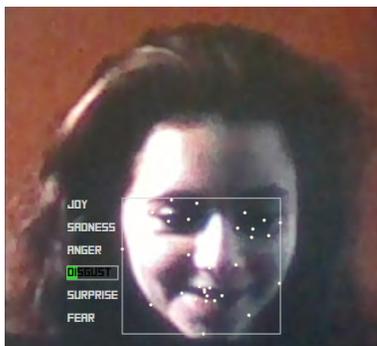


Fig. 15. Demonstration of subject's face detection (directory face95).

The subdirectory face96 contains 152 subdirectories. In each of these subdirectories there are 20 face pictures. We tested therefore 3040 face pictures. These pictures are not divided into the subdirectory female or male. Resolution of each picture is 196x196 pixels with 600

DPI, bad light conditions. With a detailed evaluation of the images from this database we found that the software can detect the face and also determine the emotional state of the subject in all cases. The software worked with 100% success rate (Fig. 16).

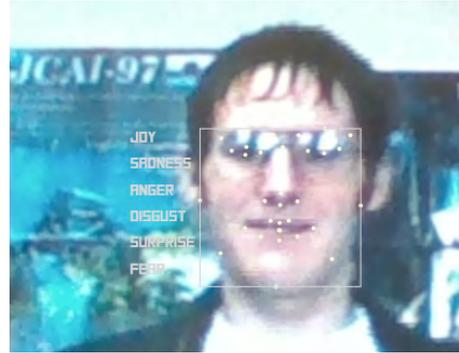


Fig. 16. Demonstration of subject's face detection (directory face96).

The subdirectory grimace contains 18 subdirectories. In each of these subdirectories there are 20 face pictures. We tested therefore 360 face pictures. Resolution of each picture is 180x200 pixels with 600 DPI, bad light conditions. With a detailed evaluation of the images from this database we found that the software can detect the face and also determine the emotional state of the subject in all cases. The software worked with 100% success rate (Fig. 17).



Fig. 17. Demonstration of subject's face detection (directory grimace).

V. GENERAL DISCUSSION

Testing our solution designed using the Affectiva SDK [37] was thoroughly tested with 6 robust databases [36].

1. Bao Face Database,
2. CMU/VASC Image Database,
3. Caltech Faces 1999 Database,
4. NIST Mug-shot Images Database,
5. Yale Face Database,
6. Vision Group of Essex University Face Database.

While experiment, we focused not only on determining the success rate of our solution, but we were wondering if this solution can correctly detect the face in the image under different lighting conditions, regardless of the distance or the face of the subject being shot. At the same time, we were interested, whether the software can not only detect but also evaluate the emotional state of the subject. From these listed databases, only one (Yale) contained images with rated emotional states of the subject. Therefore, other databases have

been used mainly to determine the ability to detect face in the image under different conditions: low resolution, different DPI, poor lighting conditions, head rotation and so on...

We determined the following research question to determine the level of face detection success and determine the user's eco-condition using our Affectiva solution and SDK solution:

RQ1: Can the software designed by us uniquely detect the face, regardless of the subject's distance, ambient light, face or facial part, or other nonstandard conditions?

Based on this research question, which was also a research problem for us, we considered the hypothesis:

H1: The ability of our software to detect faces in the image and evaluate the user's emotional state is completely independent of the distance of the subject being shot, ambient light conditions, and the face of the subject.

The research question was answered by conducting an experiment. We can say that the percentage of face detection in the image and the ability to classify the emotional state is strongly dependent on the rotation of the subject's head. At the beginning of the experiment, we measured the monitor's distance from the webcam and tested the ability to detect the subject's face at different distances. We found that the minimum distance for Face Detection is 20cm, maximum is 7.5m. For the implementation of the test environment, a uniform distance of 1 meter was chosen subsequently. The answer to the research question is that the ability to detect the subject's face by using our Affectiva SDK solution is:

1. Independent of the distance of the scanned object,
2. Ambient lighting conditions,
3. Dependent on facial rotation (if rotation is greater than 15%),
4. Independent from image rotation (face rotation of the object is not realised).

Thus, hypothesis 1 can be partially accepted. The results of the experiment did not need to be statistically verified because the software either detects or doesn't detect the face in the image and statistically calculates the level of emotional expression. Our task, however, was not to determine the percentage of successful classification of the correct emotional expression, but the overall average percentage of success of the detection of the proposed solution using the Affectiva SDK. This is 84.27%.

While conducting the experiment, however, we realised that it would be appropriate to add the solution to the software's ability:

1. Identify the sex of the subject,
2. Determine the age of the subject,
3. Show the duration of individual emotions,
4. Show the duration of detection of at least 1 face of the subject,
5. Show the number of faces successfully detected,
6. Graphically represent the statistical results,
7. To provide the results of the emotional state classification for further processing (save in a suitable format),
8. Detect the face of the subject that is rotated by more than 15°.

VI. CONCLUSION

Face recognition in real-time using a webcam, face detection itself, as well as the classification of the emotional state of the subject, is currently of great importance in various areas of our lives. Emotions affect our lives, they are a manifestation of how we feel. The emotional status classification can be used in different sectors of our life, from education (determining learner feelings), through industry (employee's

feelings), trade (the area of neuromarketing) through the automotive industry (controlling the aggressiveness of motor vehicle drivers).

In the publication, we pointed out our proposed solution using Affectiva SDK. This solution has been thoroughly tested. The total average detection rate is 84.27%. However, with the front view of the subject (if the subject does not have a rotating head above 15°), we achieve values around 100%. Looking to the left or right, when the subject's head is over 15° face detection fails as well as the classification of the emotional state. We are aware of this problem and it will be the subject of our further research.

REFERENCES

- [1] P. Augustyniak, et al. Seamless tracing of human behavior using complementary wearable and house-embedded sensors. *Sensors*, 14(5), 2014, p. 7831-7856.
- [2] K. Bahreini, R. Nadolski, & W. Westera. Towards multimodal emotion recognition in e-learning environments. *Interactive Learning Environments*, no. Ahead-of-print, 2014, p. 1-16.
- [3] V. Bettadapura. *Face Expression Recognition and Analysis: The State of the Art*. arXiv: Tech Report, (4), 2012, p. 1-27.
- [4] Y. Chen, R. Hwang, & C. Wang. Development and evaluation of a web 2.0 annotation system as a learning tool in an e-learning environment. *Computers and Education*, 58(4), 2011, p. 1094-1105.
- [5] C. Darwin. *The Expression of the Emotions in Man and Animals*. Oxford University Press, USA (1998).
- [6] P. Ekman. All Emotions Are Basic. In Ekman & Davidson (Eds.), *The Nature of Emotion*. New York: Oxford University Press, 1992, p. 15-19.
- [7] P. Ekman. Strong evidence for universals in facial expressions: A reply to russell's mistaken critique. *Psychological Bulletin*, 115(2), 1994, p. 268-287.
- [8] P. Ekman, & W. Friesen. *Facial Action Coding System: Investigator's Guide*. Palo Alto, CA: Consulting Psychologists Press, 1978.
- [9] B. Fasel, & J. Luetin. Automatic facial expression analysis: A survey. *Pattern Recognition*, 36(1), 2003, p. 259-275.
- [10] M. Feidakis, T. Daradoumis, & S. Caballe. Emotion Measurement in Intelligent Tutoring Systems: What, When and How to Measure. *Third International Conference on Intelligent Networking and Collaborative Systems*, 2011, p. 807-812.
- [11] G. Giannakakis, M. Padiaditis, D. Manousos, E. Kazantzaki, F. Chiarugi, P.G. Simos, & M. Tsiknakis. Stress and anxiety detection using facial cues from videos. *Biomedical Signal Processing and Control*, 31, 2017, p. 89-101. doi:10.1016/j.bspc.2016.06.020.
- [12] M.X. Huang, J. Li, G. Ngai, & H. V. Leong. StressClick: Sensing stress from gaze-click patterns. Paper presented at the MM 2016 - Proceedings of the 2016 ACM Multimedia Conference, 2016, p. 1395-1404. doi:10.1145/2964284.2964318.
- [13] Y. Juan, S. Dobson, & S. McKeever. Situation identification techniques in pervasive computing: A review. *Pervasive and mobile computing*, 8(1), 2012, p. 36-66.
- [14] D. Keltner. *Born to be good: The science of a meaningful life*. New York: WW Norton & Company, 2009.
- [15] A. Kumar, A. Kumar, S.K. Singh, & R. Kala. Human Activity Recognition in Real-Times Environments using Skeleton Joints. *International Journal of Interactive Multimedia and Artificial Intelligence*, 3(7), 2016, p. 61-69.
- [16] J. Li, G. Ngai, & V. Hong. Multimodal Human Attention Detection for Reading from Facial Expression, Eye Gaze, and Mouse Dynamics. *Applied Computing Review*, 16(3), 2016, p. 37-49.
- [17] Y.F. Li, J. Zhang, & W. Wang. *Active sensor planning for multiview vision tasks*. Vol. 1. Heidelberg: Springer, 2008.
- [18] A.A. Liu, et al. Coupled hidden conditional random fields for RGB-D human action recognition. *Signal Processing*, 2015, p.74-82.
- [19] M. Magdin, M. Turcani, L. Hudec. Evaluating the Emotional State of a User Using a Webcam. *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(1), Special Issue, 2016, p. 61-68.
- [20] D. McDuff, E.R. Kaliouby, T. Senechal, M. Amr, J.F. Cohn, & R. Picard. Affectiva-mit facial expression dataset (AM-FED): Naturalistic and spontaneous facial expressions collected 'in-the-wild'. Paper presented at the IEEE Computer Society Conference on Computer Vision and Pattern

- Recognition Workshops, 2013. p. 881-888. doi:10.1109/CVPRW.2013.130
- [21] D. McDuff, A. Mahmoud, M. Mavadati, M. Amr, J. Turcot, & E.R. Kaliouby. AFFDEX SDK: A cross-platform real-time multi-face expression recognition toolkit. Paper presented at the Conference on Human Factors in Computing Systems - Proceedings, 07-12-May-2016, 2016, p. 3723-3726. doi:10.1145/2851581.2890247
- [22] M. Pantic, & L.J. Rothkrantz. Automatic Analysis of Facial Expressions: The State of Art. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 22(12), 2000, p.1424-1445.
- [23] F. Pasquale, G. Percannella, & M. Vento. Graph matching and learning in pattern recognition in the last 10 years. *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 28, 2014, p. 1450001 (40 pages).
- [24] J.A. Russell. Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological Bulletin*, 115(1), 1994, p. 102-141.
- [25] T. Senechal, D. McDuff, & E.R. Kaliouby. Facial action unit detection using active learning and an efficient non-linear kernel approximation. Paper presented at the Proceedings of the IEEE International Conference on Computer Vision, 2016, p. 10-18. doi:10.1109/ICCVW.2015.11.
- [26] V. B. Semwal, et al. An optimized feature selection technique based on incremental feature analysis for bio-metric gait data classification. *Multimedia Tools and Applications*, 76 (22), 2016, pp. 24457-24475.
- [27] N. Sharma, & T. Gedeon. Modeling observer stress for typical real environments. *Expert Systems with Applications*, 41(5), 2014, p. 2231-2238. doi:10.1016/j.eswa.2013.09.021.
- [28] N. Sharma, & T. Gedeon. Objective measures, sensors and computational techniques for stress recognition and classification: A survey. *Computer Methods and Programs in Biomedicine*, 108(3), 2012, p. 1287-1301. doi:10.1016/j.cmpb.2012.07.003.
- [29] M. Suk, & B. Prabhakaran. Real-time facial expression recognition on smartphones. Paper presented at the Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision, WACV 2015, 2015, p. 1054-1059.
- [30] M. Trojahn, F. Arndt, M. Weinmann, & F. Ortmeier. Emotion recognition through keystroke dynamics on touchscreen keyboards. Paper presented at the ICEIS 2013 - Proceedings of the 15th International Conference on Enterprise Information Systems, 3, 2013, p. 31-37.
- [31] P. Viola, & M. Jones. Rapid object detection using a boosted cascade of simple features. Paper presented at the Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, p. 511-518.
- [32] L. Xia, Ch.Ch. Chen, & J.K. Aggarwal. View invariant human action recognition using histograms of 3d joints. *Computer Vision and Pattern Recognition Workshops*, IEEE Computer Society Conference on IEEE, 2012.
- [33] Z. Zeng, M. Pantic, G.I. Roisman, & T.S. Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1), 2009, p. 39-58. doi:10.1109/TPAMI.2008.52.
- [34] Z. Zhang. Feature-Based Facial Expression Recognition: Sensitivity Analysis and Experiment with a Multi-Layer Perceptron. *International Journal of Pattern Recognition Artificial Intelligence*, 13 (6), 1999, p. 893-911.
- [35] Nordic Apis. (2017, November 3). Nordic Apis. Retrieved from: <http://nordicapis.com/20-emotion-recognition-apis-that-will-leave-you-impressed-and-concerned/>
- [36] Face Detection Home Page. (2017, November 3). Face Detection and Recognition. Retrieved from <http://facedetection.com>
- [37] Affectiva. (2017, November 3). Developer Portal. Retrieved from <http://developer.affectiva.com>



M. Magdin

M. Magdin works as a professor assistant at the Department of Computer Science. He deals with the theory of teaching informatics subjects, mainly implementation interactivity elements in e-learning courses, face detection and emotion recognition using a webcam. He participates in the projects aimed at the usage of new competencies in teaching and also in the projects dealing with learning in virtual

environment using e-learning courses.



F. Prikler

F. Prikler was student at the Department of Computer Science. He deals with the theory of teaching informatics subjects, mainly face detection and emotion recognition using a webcam. At present he works as programmer in the company Muhlbauer Technologies.

Self-Organized Hybrid Wireless Sensor Network for Finding Randomly Moving Target in Unknown Environment

Mininath Nighot*, Ashok Ghatol, Vilas Thakare

SGB Amravati University, Computer Department, 444602, Maharashtra (India)

Received 16 April 2016 | Accepted 14 July 2017 | Published 29 September 2017



ABSTRACT

Unknown target search, in an unknown environment, is a complex problem in Wireless Sensor Network (WSN). It does not have a linear solution when target's location and searching space is unknown. For the past few years, many researchers have invented novel techniques for finding a target using either Static Sensor Node (SSN) or Mobile Sensor Node (MSN) in WSN i.e. Hybrid WSN. But there is a lack of research to find a solution using hybrid WSN. In the current research, the problem has been addressed mostly using non-biological techniques. Due to its complexity and having a non-linear solution, Bio-inspired techniques are most suited to solve the problem.

This paper proposes a solution for searching of randomly moving target in unknown area using only Mobile sensor nodes and combination of both Static and Mobile sensor nodes. In proposed technique coverage area is determined and compared. To perform the work, novel algorithms like MSNs Movement Prediction Algorithm (MMPA), Leader Selection Algorithm (LSA), Leader's Movement Prediction Algorithm (LMPA) and follower algorithm are implemented. Simulation results validate the effectiveness of proposed work. Through the result, it is shown that proposed hybrid WSN approach with less number of sensor nodes (combination of Static and Mobile sensor nodes) finds target faster than only MSN approach.

KEYWORDS

Particle Swarm Optimization, Self-Organization, Target Finding, Hybrid Wireless Sensor Network.

DOI: 10.9781/ijimai.2017.09.003

I. INTRODUCTION

WIRELESS Sensor Networks have gained worldwide attention due to its potential applications in area surveillance such as disaster monitoring, animal monitoring, underwater monitoring etc. [1],[2]. Different sensors have their own physical properties like temperature, moisture, smoke, light, odor, etc. As per the demands of application specific sensors are recommended to be used. The main challenges in WSN are its low bandwidth, memory limitation, and processing power. Researchers need to consider these limitations of WSN to provide a solution. In recent work, very few researchers focused on hybrid WSN due to communication hurdle between SSN to MSN.

Numbers of computations are required in analytical optimization methods. The number of computations depends on the size of the problem. If problem size increases, then computations also increase exponentially. Bio-inspired optimization techniques can be another alternative to analytical optimization. It is more efficient for the increased problem size or when the problem is complex [3],[4],[5].

The objectives of the paper are a) to simulate random moving target searching in an unknown environment, with minimum sensor nodes in hybrid WSN (SSN and MSN), b) to efficiently use, PSO (Particle Swarm Optimization) technique to achieve group movements of MSNs for target searching c) to compare area coverage of all approaches.

A. PSO - Particle Swarm Optimization

Self-organization is one of the important features of Swarm Intelligence (SI). Self-organization is a nonlinear distributed system which cannot have a linear solution and is not controlled by any single particle. It is a continuous process in which particles interact with each other locally [6], [7].

Initially, self-organized systems are predictable, but after some iteration or some time instances, these may be predictable, neutral or unpredictable.

There are mainly five features:

- Positive feedback
- Negative feedback
- Amplification
- Multiple iterations
- Balance of exploitation & exploration

The system has positive and negative feedback in which positive feedback inspires for the creation of convenient structure while negative feedback neutralizes the positive feedback [8],[9],[10].

Multiple iterations are required to reach to the goal. All particles find their own best position (i.e. Local best position). Among the local best positions of all particles, a best position is chosen (i.e. global best position) and all particles use best global position for next movement.

* Corresponding author.

E-mail address: imaheshnighot@gmail.com

B. PSO Algorithm

Generally, PSO performs searching operations using swarm particles. To get the optimal position each move in the direction to their best local position (pbest) and best global position (gbest) [11], [12], [13].

$$pbest(i, (k-1)) = \arg \min_{k=1,2,\dots,t} [f(P_i(k))], \quad (1)$$

$$gbest(k-1) = \arg \min_{i=1,2,\dots,N, k=1,2,\dots,t} [f(P_i(k))] \quad (2)$$

Where t is number of iterations and N is the total number of particles (i.e. swarm size).

Each particle in the swarm decides the movement by an objective function

$$f(x_1, x_2, \dots, x_n) \text{ where } f: \mathcal{R}^n \rightarrow \mathcal{R} \quad (3)$$

The fitness of particle is calculated from its $pbest$ position in the searching area. The particle where $pbest$ is closer to the $gbest$ have lower cost and vice versa. PSO determines minimization of a fitness function. After each iteration, the position of a particle and its velocity are modified to achieve lower cost or higher fitness value. The notations used in PSO are shown in Table I.

TABLE I. THE NOTATIONS USED IN PSO

Symbol	Meaning
$c1$	Self-confidence factor
$c2$	Swarm confidence factor
$rand1$ & $rand2$	Random numbers
ω	Inertia weight
X_i^d	Particle's position
V_i^d	Particle's velocity
$k, k-1$	Current and previous iterations respectively (movement of particle)
$pbest_i^d$	Particles best position
$gbest^d$	Swarm's best position
P_i	Position of agents in the solution space
t	Total number of iterations
d	Dimensions of solution space

Velocity and position of every particle are modified after k iteration and is shown as:

$$V_i^d(k) = \omega V_i^d(k-1) + c1 rand1_i^d(k) + c2 rand2_i^d(k) (pbest_i^d - X_i^d(k)) \quad (4)$$

$$X_i^d(k) = X_i^d(k-1) + V_i^d(k) \quad (5)$$

Here $rand1$ & $rand2$ are random numbers in the range $[0,1]$ for good coverage. ω is $0.2 < \omega < 1.2$ an inertia weight manipulates the trade-off between exploitation & exploration abilities of the object. Flowchart of PSO is shown in Fig. 1.

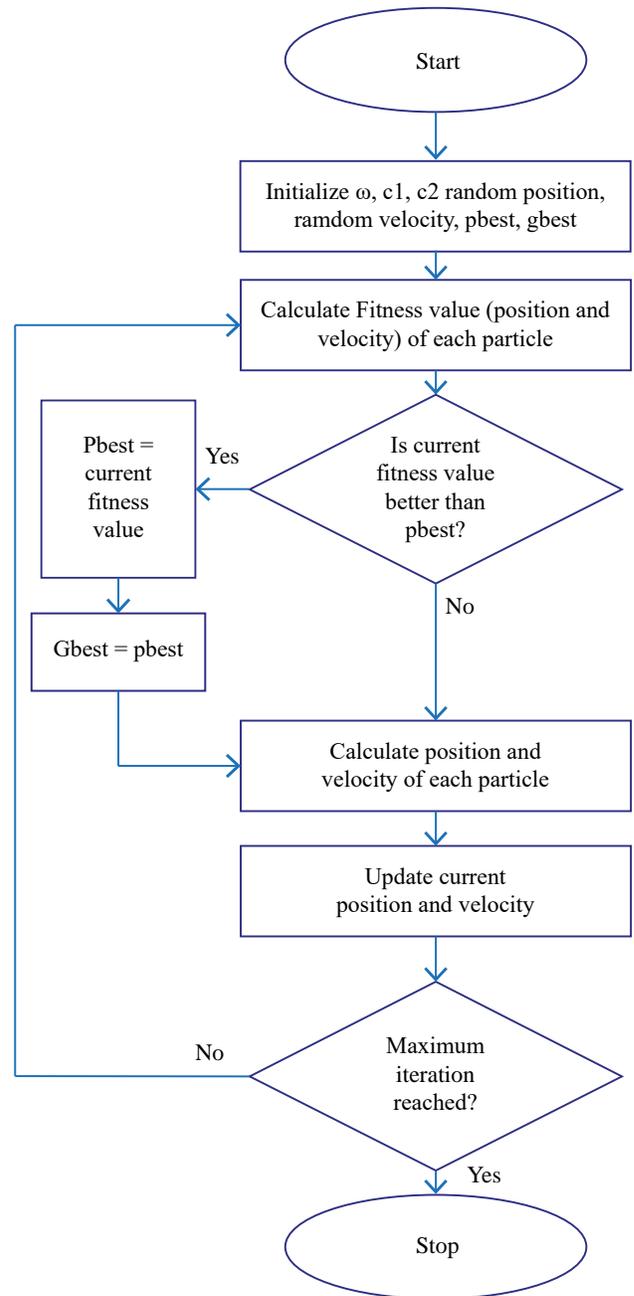


Fig. 1. Working of PSO.

C. Global Positioning System (GPS) with Real Time Kinematic (RTK)

GPS is a location receiving device used worldwide. At least four GPS Satellites and a GPS receiver are required to get the location of objects in 2-D space. But its accuracy is 5 meters to 100 meters and precision is 5 meters to more than 20 meters. So use of GPS is not possible in proposed work. It needs more accuracy and precision.

More accurate position can be calculated using Real Time Kinematic (RTK) [14], [15]. Its accuracy is up to 2 centimeters. It consists of one GPS base station and multiple rovers. Setup positions are shown in Fig. 2. GPS base station is positioned on known location. It takes measurements from satellites in view and sends it with its known position to the rovers. Rover receiver also collects measurement from satellite in view, and process it with the base station information. Rover calculates their locations with relative to the base station.

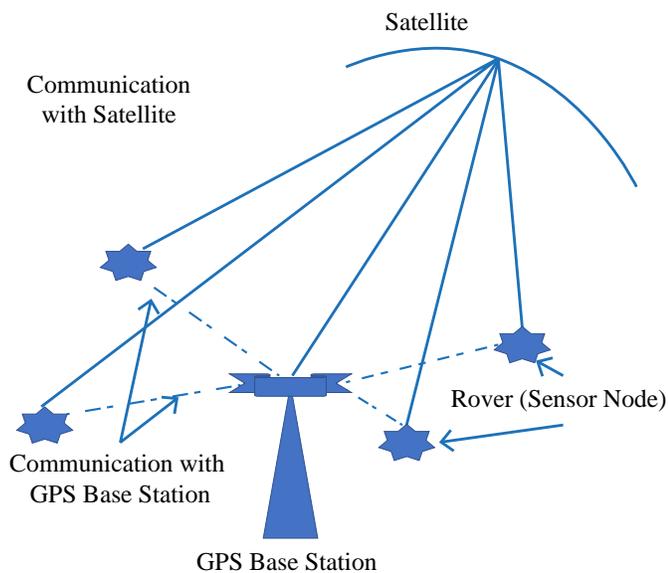


Fig. 2. Real Time Kinematic (RTK) setup.

Equipments of RTK are costlier hence GPS less technique can be used to get the position of sensor nodes. The proposed work is simulation based so not much focused on how to get locations of sensor nodes.

The rest of the paper is organized as follows: A literature survey is explained in section II. In the section III proposed work with algorithms are introduced and explained in detail. Results analysis is given in section IV and in section V conclusion and future scope is given.

II. LITERATURE SURVEY

In recent years, several researchers proposed new techniques to find the object. Most of the techniques are SSN based and very few are MSN and hybrid based. Bio-inspired optimization strategies are also implemented to optimize the performance of their work [16],[17]. Some of the techniques are now discussed.

Zhuofan Liao et al [18] proposed algorithms to solve MSN deployment (MSD) problem. Multiple MSN & Multiple Static Targets are considered in the system. The study focuses on overall energy consumption by minimizing MSN's movement to track the target & Network connectivity. MSD is divided into two subproblems 1. TCOV (Target Coverage) and 2. NCON (Network Connectivity) problems. TCOV problem is solved using two Heuristic algorithms- Basic algorithm and TV Greedy algorithm. (Target based Voronoi Greedy algorithm) in which Basic algorithm selects one MSN for one target, and TV-Greedy algorithm minimizes the total number of movements using a Voronoi diagram. For NCON problem ECST & ECST-H algorithms are proposed to move coverage MSN near to moved MSN. So that moved MSN will be able to communicate with sink node. Matlab simulation is carried out to support the proposed work.

In dynamic transportation system, Ning Zhu. et al [19] proposed a system which is used for collecting traffic information by MSN. Two problems are studied: choosing route link and staying time on the route link. To tackle these problems Ant Colony builds the route for MSN and Particle Swarm Optimization (PSO) determines the stay time on each route link. The proposed work is studied analytically and proved that MSNs are more effective than SSN in transport network surveillance.

Dusade A. et al [20], proposed Moving Object Tracking using Support Vector Machine (MOT-SVM) for finding a movable object by SSN. The proposed algorithm requires fewer communication resources and less amount of communication computations. Because each sensor

node needs to send one bit of information to the central processing unit to indicate that a moving object is going far or coming near. Moving target far from observing SSN indicates '-' and near indicates '+'. Observer SN calculates + and - using RSS (Received Signal Strength). If RSS is less, it is going far & if RSS is more it is coming near. Experimental analysis is carried out and compared with Aslam's work & shown MOTSVN performs well in terms of accuracy, precision, and robustness to data errors.

Jia Wei Tang et al [21], proposed a technique to track moving objects using image processing technique. UAV (Unmanned Aerial Vehicle) with vision capability is used to detect the movement of objects. FPGA (Field Programmable Gate Array Algorithm) is used to develop the UAV. For motion estimation and object segmentation block (area based) matching & RANSAC (Random Sample Consensus) algorithms are used. The entire work is supported by self-experimentation and analytical calculations.

Hamid Maboubi et al [22], [23], focused on problem of tracking and monitoring a mobile target in the field with multiple obstacles for an increasing network lifetime. MSN's are located using relocation technique. The proposed work finds near-optimal relocation strategy for MSN. It also finds the energy efficient path to send information from movable target to destination. The proposed technique is proved by simulation result & shown network life time increases.

Enyang Xu. et al [24] proposed work for mobile target tracking by MSN. MSN controller gets the location of MSN & Target continuously by anchor nodes. After analysis of Time of Arrival (TOA), controller guides movement to MSN. MSN navigation strategy, target localization, MSN localization and joint target & MSN localization are formulated & calculated. Based on Time of Arrival (TOA), convey optimization algorithm is developed for localization. Cubic law is used for routing MSN.

Yifan Cai et al [25], [26], proposed a couple of algorithms to track the target in a totally unknown physical environment. Multiple robots search the target cooperatively to get the parameter ranges of cooperation method in multiple robots using a combination of HRL and MAXQ algorithms.

All parameters which are required for cooperation are obtained through learning approach and new tasks can be performed by the multiple robots. With simulation study, it is studied and shown that multiple robots in the unknown environment can search the target. Summary of studied literature is shown in Table II. (Table II (a) & Table II (b)).

III. PROPOSED WORK

Proposed technique shows self-organization Mobile Sensor Node (MSN) and Static Sensor Node (SSN) to find the moving target (T) using Particle Swarm Optimization (PSO). MSN & SSN are equipped with Global Positioning System (GPS) which is used to tell their exact location coordinates in two-dimensional spaces. Target is assumed to be movable random. Without having visual sense and directional guidance to MSN and SSN, using only location guidance, MSN tracks the moving target.

The proposed system is heterogeneous, using both SSN and MSN in the same system. SSNs have the capability to send their own locations to all MSNs which are far away. If the target is in the searching range (SR) of SSN then SSN sends its own location to all MSN.

MSN are low capability sensors having a movable trolley/vehicle on which MSN's are located. The trolley moves as per the signals are given by MSN. It can move in only three directions Left (Lt), Straight (St) and Right (Rt), from the current position. Left and right rotations are exactly 45 degrees from the current position. Fig. 3 shows the

TABLE II (a). SUMMARY OF LITERATURE

Sr. No.	Year & Author	Problem Identified	Algorithm used	Outcomes
1	2015- Zhuofan Liao and others	MSN Deployment (MSD) • Target Coverage (TCOV) • Network Connectivity (NCON)	TCOV-(Heuristic Algorithm) • Basic Algorithm • TV Greedy Algorithm- Target Based Voronoi Greedy Algorithm • NCON (Approximate Algorithm) • ECST – Euclidian Minimum Spanning • ECST-H- ECST- Hungarian	• Controls on the movement of MSN • Prolong the network lifetime
2	2014- Ning Zhu and others	Mobile sensor's use in the transportation system to collect traffic information	Hybrid two-stage algorithm based on PSO and ACO	MSN's are better traffic information collector than SSN
3	2013- Dusade A. and others	Moving object tracking in WSN	MOT SVM- Moving Object Tracking using Support Vector Machine	Out performs than Aslam's work in terms of accuracy, precision, and robustness to data errors
4	2016- Jia Wei Tang and others	Real Time moving object detection	• FPGA Algorithm- Field Programmable Gate Array Algorithm • Block Matching Algorithm • RANSAC- Random Sample Consensus Algorithm	Tracking of Moving Objects
5	2016- Hamid Mahboubi and others	• Shortest path • The problem of tracking and monitoring a moving target in a field with an obstacle • Maximize network lifetime	Residual Energy based Voronoi Diagram	Prolong the lifetime of the network
6	2013- Enyang Xu and others	Track Mobile target by MSN	TOA- Time of Arrival Cubic Law	Target follower MSN with good performance
7	2013- Yifan Cai and others	Target searching in unknown environment	• HRL- Hierarchical reinforcement Learning • MAXQ	Target Searching in unknown Environment

TABLE II (b). SUMMARY OF LITERATURE (CONTINUED AS OF TABLE II (a))

Sr. No.	Proposed work compared with	SSN/MSN/ Hybrid	GPS	Stochastic Algorithm	Target	Network	Control	Simulation/Execution
1	None	MSN	No	No	Static (Multiple)	MSN	Centralized	Simulation –Matlab
2	Self's analytical results	MSN	No	PSO and ACO	No- collect traffic information	Transport System	Distributed	Analytical study (Numerical Experiment)
3	Aslam's Work	SSN	No	No	One – Movable	Binary Sensor Network	Centralized	Not Mentioned
4	NO	UAV/MSN – Unmanned Aerial Network	No	No	Movable	Real Time Network	Centralized	Experimental/ Analytic
5	No/ Simulation study compared analytically	MSN	No	No	Movable	MSN with obstacle	Centralized	Not mentioned
6	Analytical study	MSN	No	No	Movable - multiple	MSN	Centralized	Not mentioned
7	NO	Robots/ Unknown Environment	No	No	Multiple	Unknown Environment	Distributed	Not Mentioned

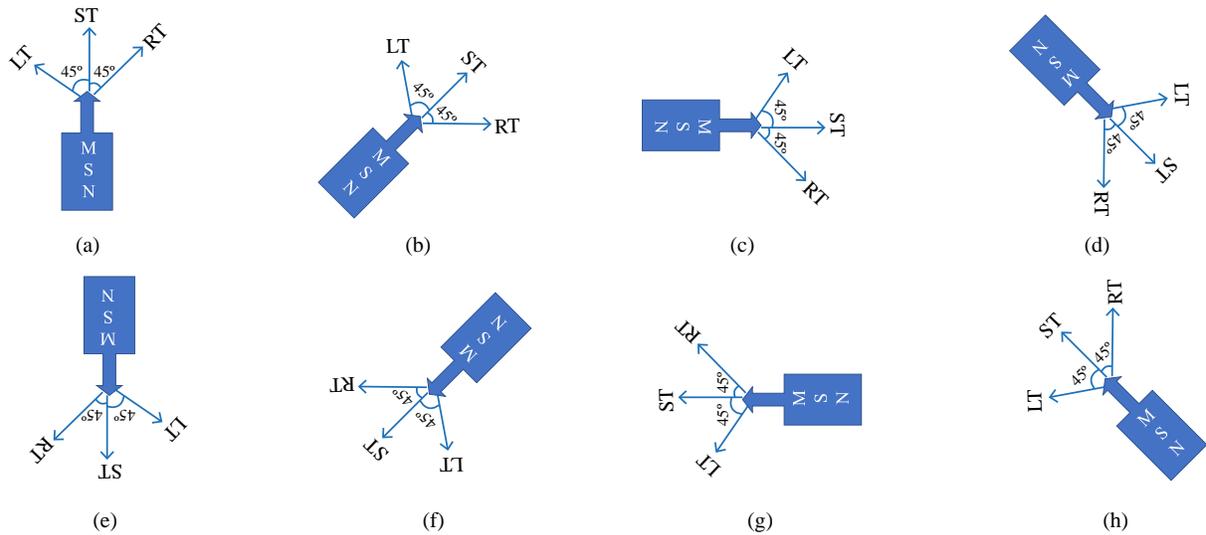


Fig. 3. The possible rotations from current position.

possible rotations from the current position. There are 8 possible initial positions of MSN and each position has its relative left, straight and right rotations for movement.

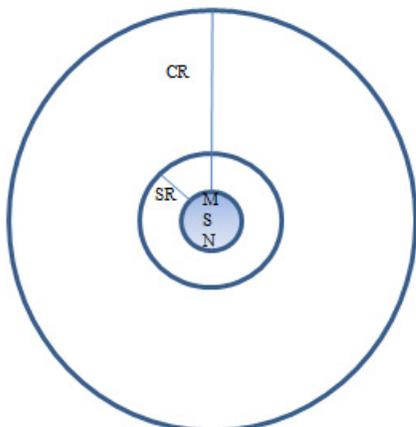


Fig. 4. SR and CR.

Searching Range (SR) is a range or radius of MSN and SSN in which it can search the target. If the target comes in the SR of any MSN, it detects the target. Communication Range (CR) / radio range is a radius of MSN in which MSNs can communicate with each other for sharing their location coordinates and sharing a message mentioning the target is in its SR. [27],[28],[29]. SR and CR are shown in Fig. 4.

A. Topology Formation by MSN

All MSNs together form a swarm. MSN moves randomly to search the target, but it must be in the range of at least one MSN's CR. Initially MSNs are located in such a way that they will not share each other's SR. Any two MSNs are a 2SR distance away from each other. If two MSNs are sharing SR, then same space will be searched by these two MSNs. By using this initial condition, the search space is increased. Fig. 5 shows the possible minimum and maximum distance between two MSN. Fig. 5 (a) shows MSNs are sharing other's SR. In such case, the same area will be searched by multiple MSNs and is a waste of time. Fig. 5 (b) shows the minimum distance between two MSN i.e. 2SR. So that it will search in the unsearched area. Fig. 5 (c) shows the maximum distance between two MSN i.e. CR to communicate with each other. If distance is increased than CR then sensors will not be able to communicate with each other.

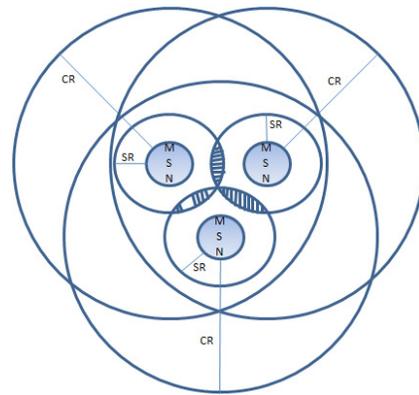


Fig. 5 (a) shows that the dashed area is searched by both sensors which are not required.

B. Basic Flow of Systems

Initially, SSNs are placed randomly in the searching space. Target (T) is also placed randomly and is movable. It moves randomly and is not GPS equipped. Swarm of MSNs is placed at any random place or at the border of searching space. MSNs move randomly by swarm technique. While moving randomly if T found, then the mission is complete, else searching process continues.

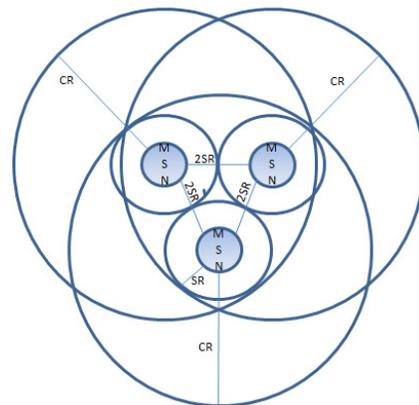


Fig. 5 (b) Shows at least 2SR should be the distance between MSN.

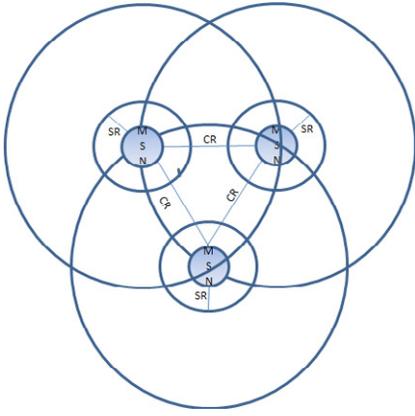


Fig. 5 (c) shows worst case scenario of sensors (maximum distance between two sensors).

Due to the movement of T, if it comes in the SR of any SSN then SSN broadcasts its own location coordinates and a message saying ‘Target is found’ to all MSNs. Once location and message received by all MSNs, all MSNs calculate its own distance $Dist_i$ from sender SSN using Euclidean distance [30], [31].

$$dist_i = d_i + n_i \quad (6)$$

Where d_i is distance between SSN and MSN and n_i is Gaussian additive noise, which has random value uniformly distributed in the range [32]:

$$d_i \neq d_i(P_n / 100) \quad (7)$$

P_n is a percentage noise. Accurate calculation of distance depends on the value of P_n

$$d_i = \sqrt{(x - x_i)^2 + (y - y_i)^2} \quad (8)$$

Where (x, y) are coordinates of SSN and (x_i, y_i) are coordinates of particles.

All MSNs share their calculated distance from Sender SSN (SSSN) to each other. The MSN which is closer to SSSN will be selected as a Leader and others will be followers. Leader MSN (LMSN) will decide its direction to reach to SSSN and move step by step towards SSN. Other MSN will not waste their energy for deciding the direction and path. They will just follow the LMSN. At every step, All MSNs calculate their current distance from SSSN and share it to all MSNs. If any Follower MSN (FMSN) is closer to SSSN than LMSN, it will be selected as a new LMSN and former will be FMSN. In such fashion, MSNs travel to SSN, without knowing directions and any manual interaction. Once all MSNs reach to SSN, they start searching again with their random search.

C. Proposed Algorithms

To search the target, four algorithms are proposed which are as follows:

1. MSN’s Movement Prediction Algorithm (MMPA) for searching (T) is shown in Fig. 6.
2. Leader Selection Algorithm (LSA) is shown in Fig. 7.
3. Leader’s Movement Prediction Algorithm (LMPA) for traveling towards SSN is shown in Fig. 8.
4. Follower’s Algorithm (FA) is shown in Fig. 9.

D. Algorithms

1. MSN’s Movement Prediction Algorithm (MMPA) for searching (T)

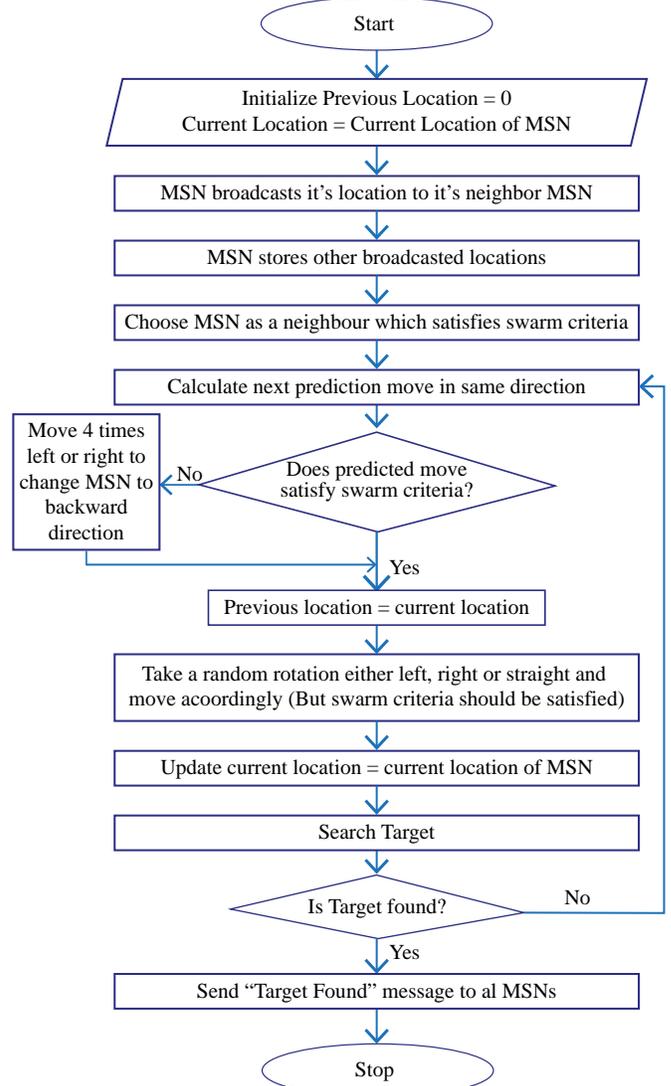


Fig. 6. MMPA Flowchart.

2. Leader Selection Algorithm (LSA)

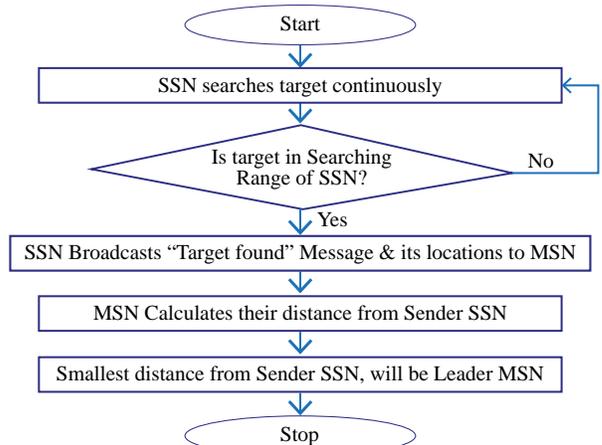


Fig. 7. LSA Flowchart.

3. Leader's Movement Predication Algorithm (LMPA) for traveling towards SSSN

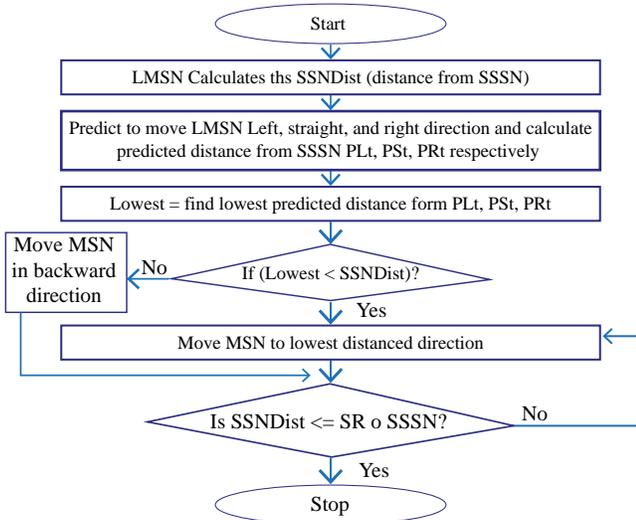


Fig. 8. LMPA Flowchart.

4. Follower's Algorithm

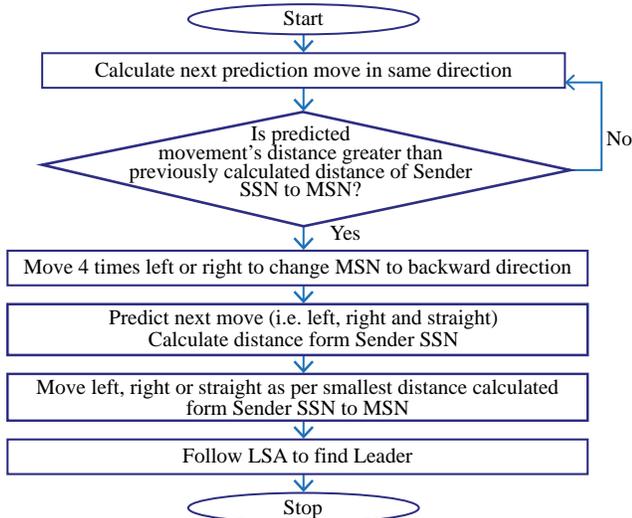


Fig. 9. FA Flowchart.

IV. RESULTS AND DISCUSSIONS

Matlab 7.0 simulator is used for the experiment. Due to the random movement of MSNs and T, the number of iterations/steps cannot be predicted. A number of times simulation is executed and tested for finding the T. The assumed parameters with respective values for the simulation are given in Table III.

SSNs are placed in searching space in such a way that whole area is covered. Searching space is divided into four quadrants and four SSNs are placed at the center of each quadrant (SSN1-(2.5, 2.5), SSN2-(2.5, 7.5), SSN3-(7.5, 2.5) and SSN4-(7.5, 7.5)) and fifth SSN is placed at the center of searching space (SSN5-(5, 5)). MSNs are placed at left-lower corner of searching space (MSN1-(0.3, 0.9) and MSN2-(0.3, 0.3)). While moving, MSNs maintain at least 2SR distance from each other and will be in the CR of other MSN. T is placed randomly in searching space. The speed of T is same as speed of MSNs. If the T is in the SR of any MSN then it is assumed that T is found.

TABLE III. PARAMETERS AND VALUES

Sr. No.	Parameter	Value
1	Searching Space:	10 X 10 (Obstacle free)
2	Total Number of MSNs:	2
3	Total Number of SSNs:	5
4	Total Number of Target (T):	1
5	Location of T:	Unknown
6	Communication Range (CR) of sensors:	1.2
7	Searching Range (SR) of sensors:	0.3
8	Targets Initial Position:	Random
9	MSNs and Target Movement:	Random
10	The speed of Target:	Same as MSN
11	GPS enabled Sensors	Yes
12	Maximum Iterations	1000

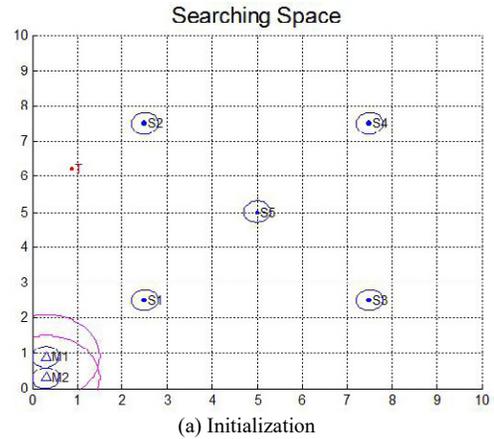
In the simulation, the movement of the target is shown in red colored dots, movement of MSN1 is shown in blue colored dots and movement of MSN2 is shown in green colored dots. Blue circles and pink circles indicate the SR and CR of the respective sensor. In SSN, CR is not visible for the sake of visibility of MSN's and T's movement but it is same as CR of SSN.

Simulations readings are observed after every 100 iterations or T comes in SR of any SSN or MSN. It is assumed that if T moves in SR of any MSN, then searching mission gets completed and it shows a total number of iterations required to find the T.

A number of times simulation is executed and some of the cases are discussed here as an example.

A. Case 1: Target found by MSNs only

Fig. 10 (a) shows the initial placement of sensors and T. T moves randomly in searching space as iteration increases. The path and area covered are shown in the red colored tail of T. At the same time MSNs also move using PSO technique to find the T. MSNs maintain the atleast 2SR distance and atmost 1 CR distance from each other. MSNs movements are shown in blue and green colored tail of MSN1 and MSN2 respectively. While moving, MSN found T in their SR at 68th iteration and is shown in Fig. 10 (b).



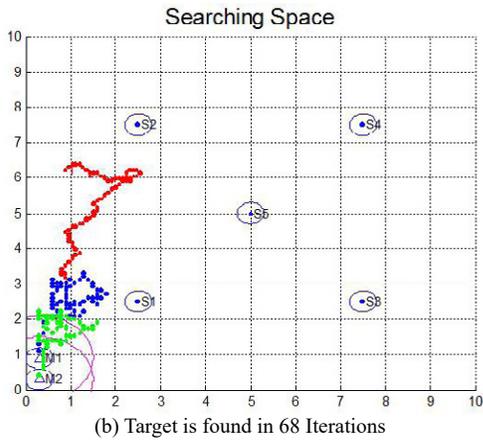


Fig. 10. Movement of MSNs and T.

B. Case 2: T is Found by MSN with the Help of SSN 4

Fig. 11 (a) shows the initial localization of sensors and T where T is located near the SSN4. After one iteration, T moves in the SR of SSN4, which is shown in Fig. 11 (b). Once T is found by SSN4, it broadcasts 'T Found' message and SSN4's location to MSNs. MSNs stops the searching and switch into the leader-follower mode and reaches to CR of SSN4. It is shown in Fig. 11 (c). While MSNs move towards SSN4, T moves randomly in another location. Once MSNs reaches in the CR of SSN4, MSNs switch their mode in searching mode and search the T as a random search strategy.

Fig. 11 (d) and 11 (e) show the movements of the sensor and T after 100th and 200th iterations respectively. T is found in 228th iteration by MSNs and is shown in Fig. 11 (f).

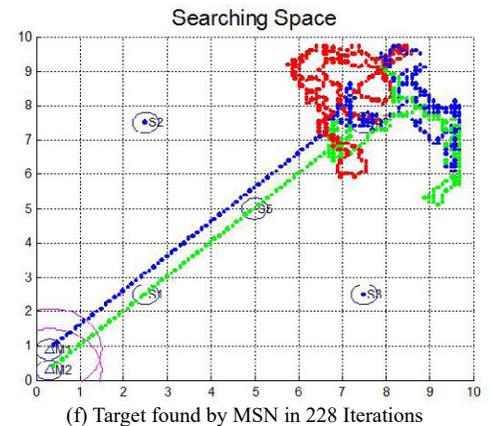
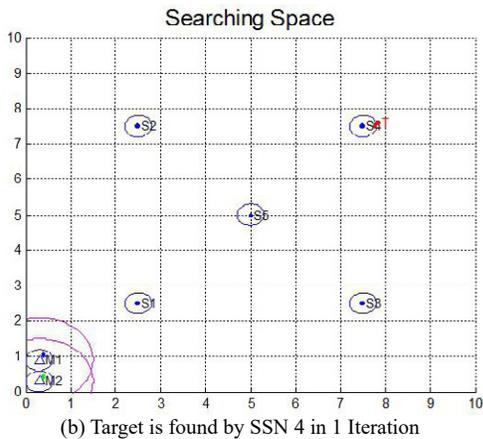
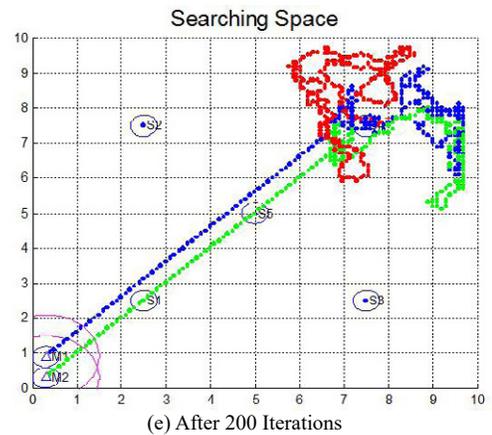
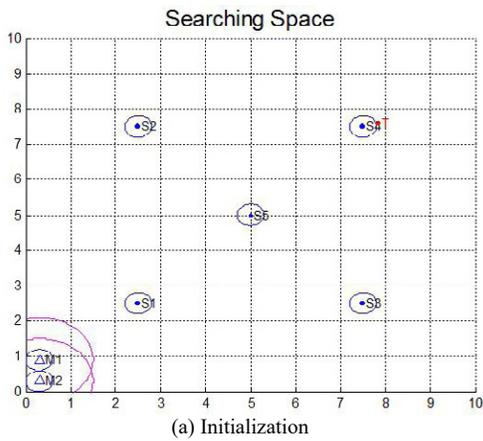
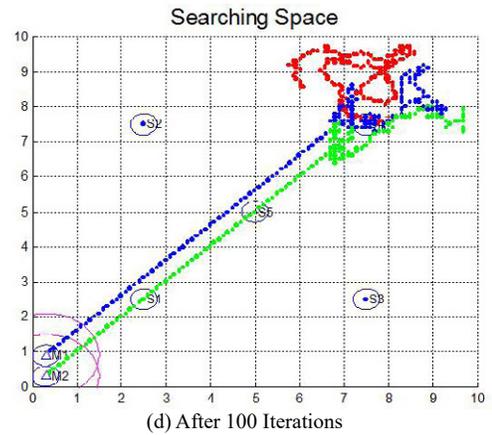
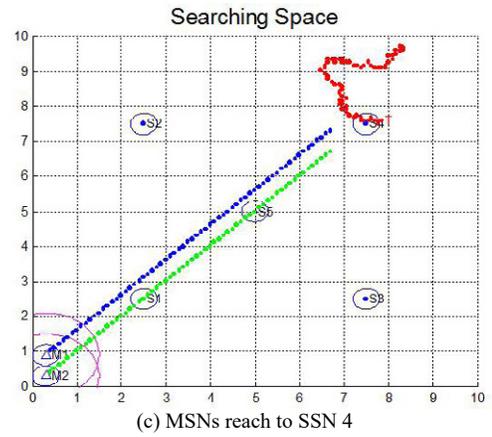


Fig. 11. Movement of MSNs and T.

C. Case 3: T is Found by MSN with the Help of SSN5

Fig. 12 (a), (b), (c) and (d) are shown as explained in case 2. Instead of SSN4, T appears in SR of SSN5. T is found in the 28th iteration of MSNs.

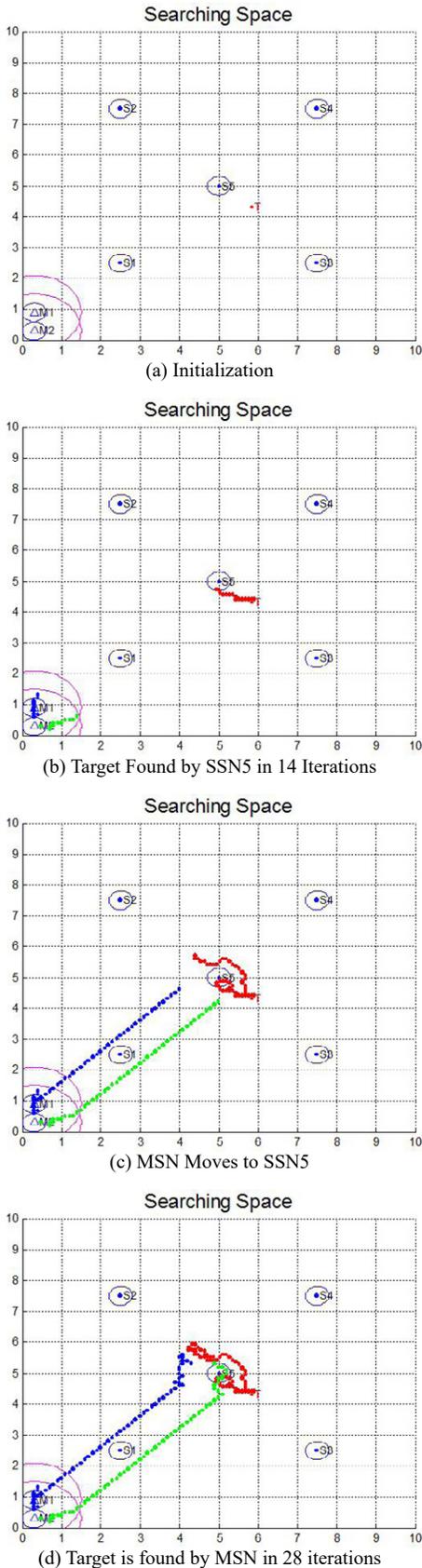
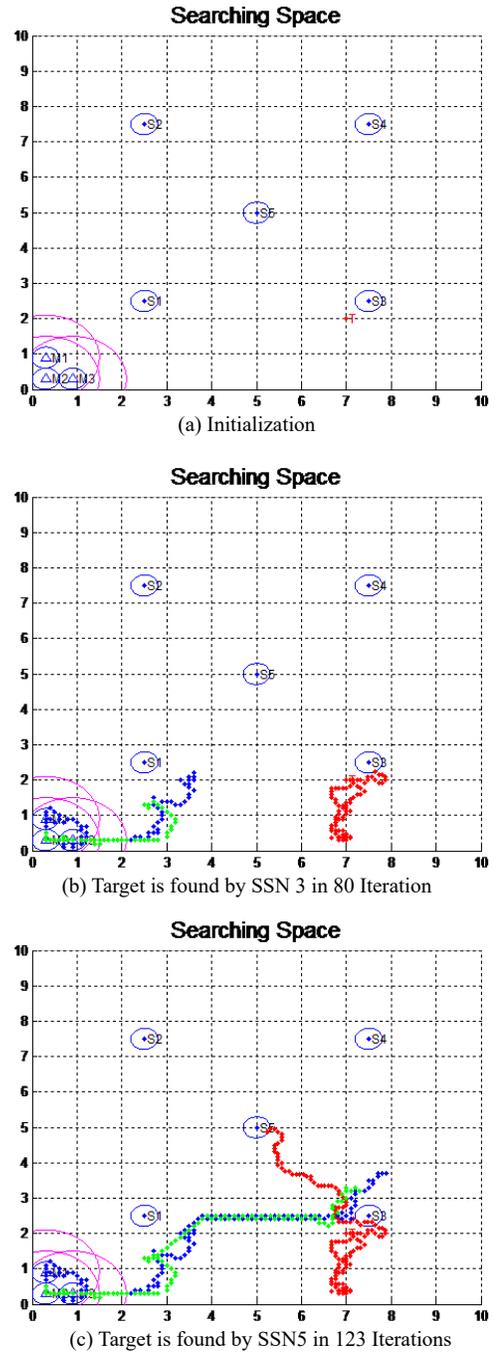


Fig. 12. Movement of MSNs and T.

D. Case 4: Involvement of Multiple SSNs

Sometimes T moves into the bigger area. Before coming all MSNs near to the SSSN, T moves far away from its found location. Once MSNs reaches to SSN, it searches T by random search. So MSNs are not able to find T near to SSN. After 123rd iteration, T is found by SSN2 so MSNs need to move towards SSN2. It is shown in Fig. 13 (a), (b), (c) and (d). Fig. 7 (e) shows T is found in 254th iterations by MSN near to SSN2. Even though T is traveling more distance in a straight way, MSNs are succeeded to find T.



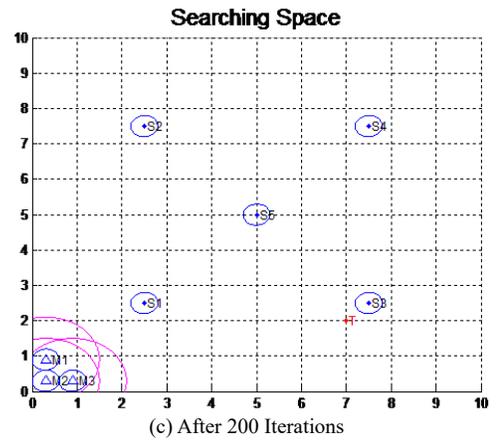
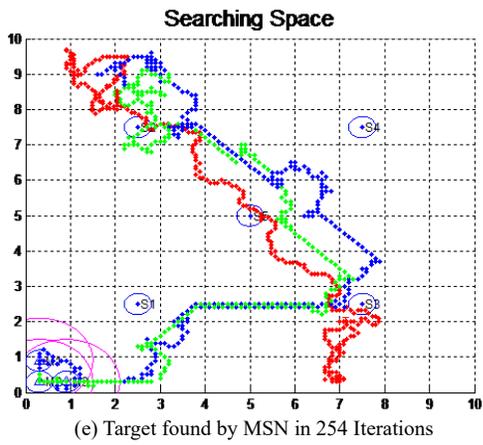
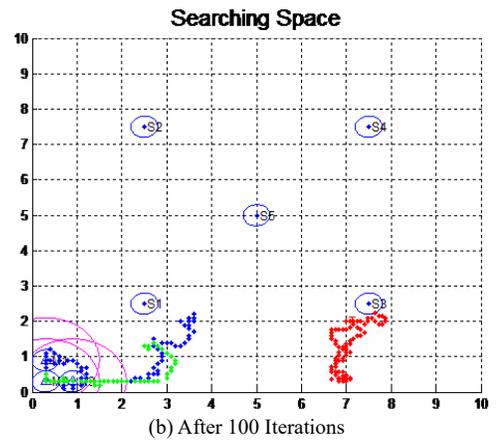
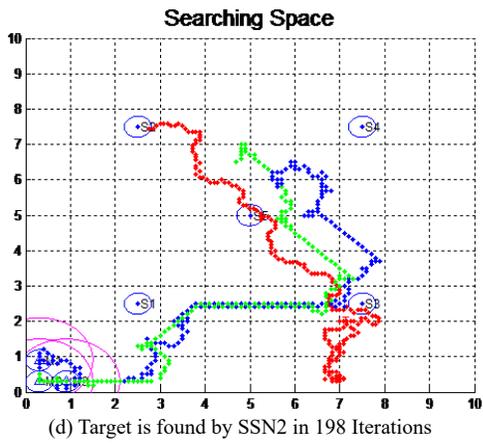
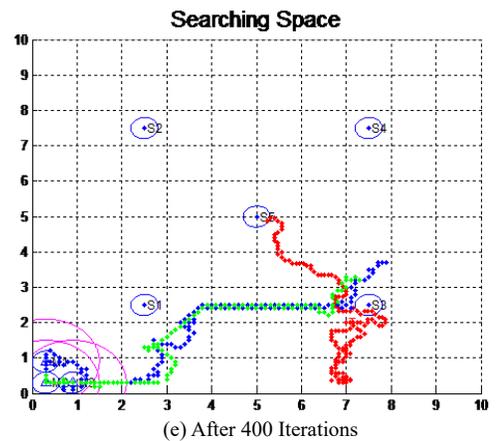
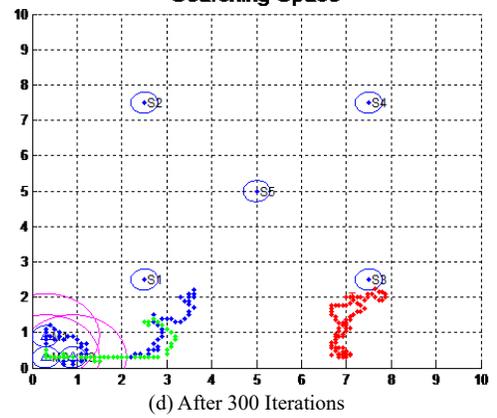
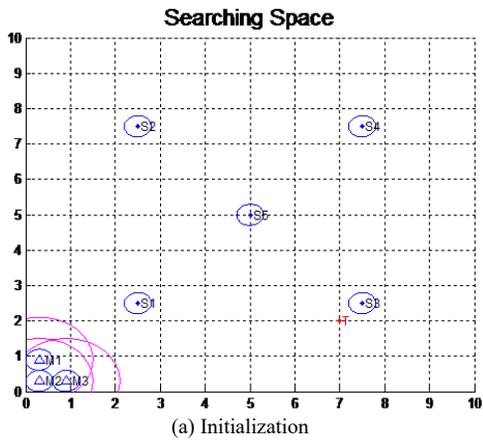


Fig. 13. Movement of MSNs and T.

E. Case 5: Worst Case

Sometimes T does not come in SR of any SSN or MSN then it requires more number of iterations. It works like only MSNs are in the network searching for T. Fig. 14 (from the figure (a) to figure (k)) shows that in 1523rd iteration T is found by MSN.



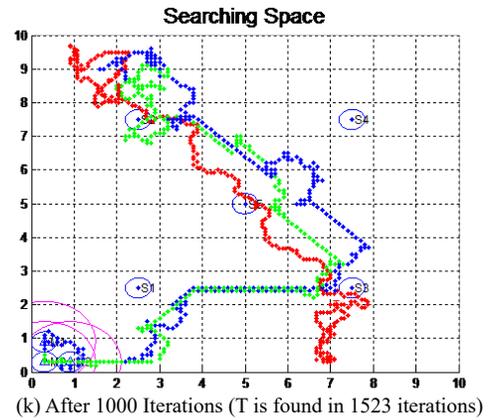
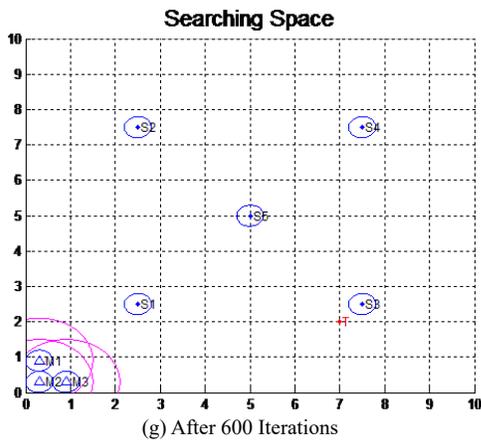
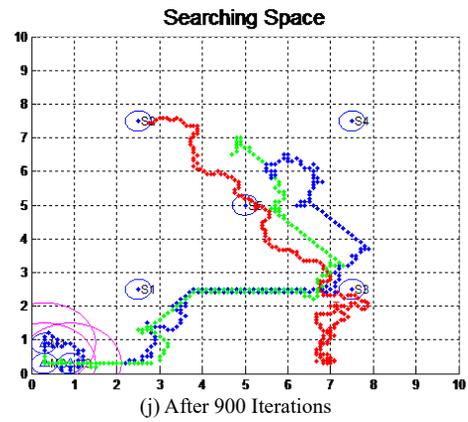
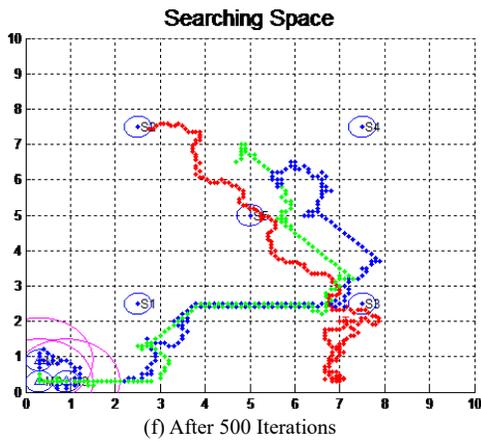
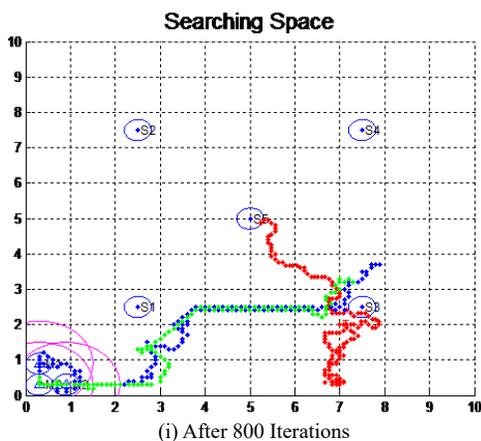
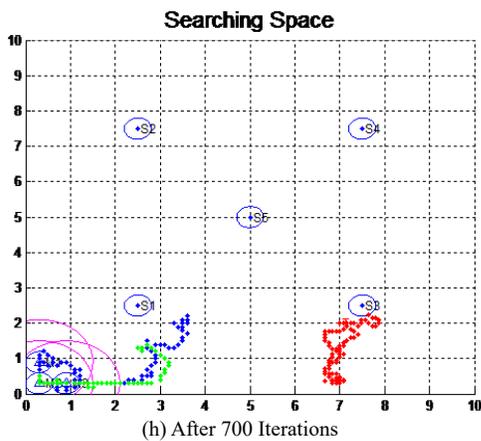


Fig. 14. Movement of MSNs and T.



F. Comparison Between the Networks with Only MSNs and Hybrid WSN

To show the effectiveness of hybrid WSN, two types of simulations are performed. The first type of simulation consists of only MSNs which are placed on the test bed. And in the second type, both SSNs and MSNs are placed.

If the size of searching space is small then both scenarios work well, but if the size of searching space is large then the simulation with only MSNs requires more number of iterations to find T as compared to hybrid WSN. The total number of iterations required for both scenarios is shown in Table IV and Fig. 15. To find the T, 10 simulations are carried out of each scenario and numbers of iterations are sorted in ascending order. Average number of iterations in SSN is only 10 but to transmit data to base station it require more energy. Only 572.9 average iterations are required for Hybrid WSN and are far less than only MSN scenario.

TABLE IV. NUMBER OF ITERATIONS TO FIND T

Sr. No.	Only SSNs (Dense Deployment)	Only MSNs (Number of Iterations)	Hybrid WSN (Number of Iterations)
1	1	52	28
2	1	168	68
3	1	1,021	228
4	1	1,021	254
5	1	1,224	452
6	1	1,335	622
7	1	1,512	784
8	1	1,627	772
9	1	1,651	998
10	1	1,724	1,523
Average	10	1133.5	572.9

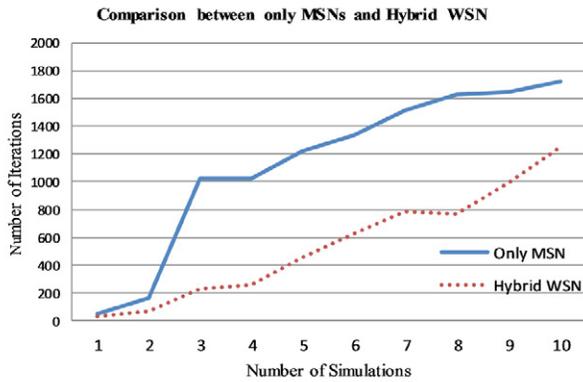


Fig. 15. Effectiveness of Hybrid WSN over the network with only MSNs (SSN is not considered due to dense nodes in network).

In Hybrid approach, if more number of SSNs are placed then the number of iterations reduces.

G. Coverage Comparison of Searching Space

To cover the complete searching space (10×10) by sensors range, more number of SSNs need to deploy. The number of SSNs needed to cover area depends upon the SR of sensors. If $SR=0.3$ in same searching space the 289 (17×17) SSNs are needed.

1) Degree of coverage

The degree of coverage can be the ratio of entire searching space and area covered by all sensors and coverage area can be area covered by all sensors through their SR [18]. In Only SSN approach when $SR=0.3$, the area coverage, and the coverage degree are 86.7 and 1.15 respectively. In only MSN approach area coverage, and the coverage degree are 0.6 and 166.67 respectively. This approach requires more time to search T.

In the proposed technique SR is 0.3 and number of sensors are only 7 (2 MSNs and 5 SSNs). The coverage area and the coverage degree of proposed work are 2.1 and 47.62 respectively, which is shown in Table V and Fig. 16. Even though proposed work has less coverage area than SSN approach and larger cover area than MSNs approach, it requires less energy than other approaches for searching T.

TABLE V. COVERAGE AREA AND COVERAGE DEGREE

Example	Searching Space	Number of Sensors	SR	Coverage Area	Coverage Degree
Only SSNs	10×10	289 (SSNs)	0.3	86.7	1.15
Only MSNs	10×10	2 (MSNs)	0.3	0.6	166.67
Proposed work	10×10	7 (2 MSNs +5 SSNs)	0.3	2.1	47.62

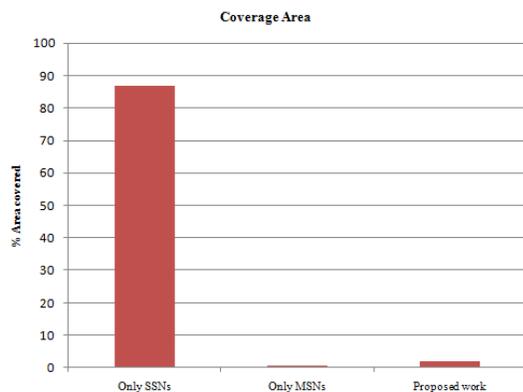


Fig. 16. Coverage Area Comparison.

In spite of time constraint, even though area covered by proposed approach is 2.1, it finds target successfully.

V. CONCLUSION AND FUTURE SCOPE

Randomly moving target is efficiently found by novel proposed technique. MSNs are allowed to move in only three random directions. Without control of third parties, MSNs move autonomously and find target successfully. A novel concept of hybrid WSN is implemented to find the target. SSNs and MSNs are utilized together to achieve objectives. If more number of SSNs & MSNs are deployed, then searching time reduces tremendously. A PSO technique in hybrid WSN is effectively implemented to find the unknown target in an unknown environment. Results show that hybrid WSN performs better than only MSNs in the network.

It is proved that using less number of sensors, target finding can be done effectively so due to less number of sensors utilized in the network, equipment cost and its maintenance cost is reduced.

Obstacle free environment is assumed for simulation. Single target is assumed and its speed is assumed same as MSNs speed. But in real case target's speed may be lower or higher than MSNs speed. In future work, it is planned to consider obstacles in searching space with multiple targets and lower or higher speed of target than MSNs.

REFERENCES

- [1] C.Y. Chong and S. Kumar, "Sensor Networks: Evolution, Opportunities and challenges," proceeding IEEE, vol. 91, no. 8, pp. 1247-1256, Aug. 2003.
- [2] Mininath Nighot and Ashok Ghatol, "GPS Based Distributed Communication Protocol for Static Sensor Network (GDGP)," Procedia Computer Science, Vol. 78, pp. 530-536, 2016.
- [3] Bahuguna, Y., D. Punetha, and P. Verma, "An analytic Study of the Key Factors Influencing the Design and Routing Techniques of a Wireless Sensor Network," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 3, pp.11-15, 2017.
- [4] Eliseo Ferrante, Ali Emre Turgut, Alessandro Stranieri, Carlo Pinciroli, Mauro Birattari, Marco Dorigo, "A self-adaptive Communication strategy for flocking in stationary and non-stationary environments," Nat. Comput (2014), vol. 13, pp. 225-245, 2014.
- [5] Hongliang Ren and M.M.Q.H. Meng, "Biologically Inspired Approaches for Wireless Sensor Networks," Proceeding IEEE International Conference, Mechatronics and Automation, 2006.
- [6] Zeng, J.C., Jie, J. and Cui, Z.H., "Particle Swarm Optimization," Science Press, Beijing (2004).
- [7] M. K. Nighot, V. H. Patil and G. S. Mani, "Multi-robot Hunting based on Swarm Intelligence," *Hybrid Intelligent Systems (HIS)*, vol. 12, pp. 203-206, 2002.
- [8] Kennedy, J. and Eberhart, R.C. "Particle swarm optimization," Proceeding IEEE Conference on Neural Networks, vol. 11, pp. 1942-1948, 1995.
- [9] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," Proc. IEEE Int. Conf. Neural Network, pp. 1042-1048, Nov.-Dec. 1995.
- [10] R.C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," Proceedings on international symposium on micro machine and human science, IEEE Service Center, USA, pp. 39-43, 1995.
- [11] J. Kennedy, "The particle swarm: social adaptation of knowledge," Proceedings of IEEE international conference on evolutionary computation, IEEE Service Center, USA, pp. 303-308, 1997.
- [12] C. Anderson and N.R. Franks, "Teams in animal societies behavior," *Ecol*, 12 (5), pp. 534-540, 2001.
- [13] P.B.S. Lissaman and C.A. Shollenberger, "Formation flight of birds Science," 168 (3934), pp. 1003-1005, 1970.
- [14] D. Manolakis, "Efficient Solution and Performance Analysis of 3D Position Estimation by Trilateration," IEEE Trans. On Aerospace and Electronic Systems, vol. 32, no. 4, pp. 1239-1248, 1996.
- [15] Xiyang He, Xiaohong Zhang, Long Tang and Wanke Liu, "Instantaneous Real-Time Kinematic Decimeter-Level Positioning with BeiDou Triple-Frequency Signals over Medium Baselines," *Sensors*, 16, 1 pp. 1-19, 2016.

- [16] Kaur, R., and S. Arora, "Nature Inspired Range Based Wireless Sensor Node Localization Algorithms," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 3, pp-1-11, 2017.
- [17] Sanjeev Wagh and Ramjee Prasad "Energy Optimization in Wireless Sensor Network through Natural Science Computing: A Survey," Journal of Green Engineering, River Publication, vol. 3, pp. 383-402, 2013.
- [18] Zhuofan Liao, Jianxin Wang, Shigeng Zhang, Jiannong Cao and Geyong Min, "Minimizing Movement For Target Coverage and Network Connectivity in Mobile Sensor Networks," IEEE Transactions on Parallel Distributed System, vol. 26, no. 7, July 2015.
- [19] Ning Zhu, Yang Liu, Shoufeng Ma, and Zhengbing He, "Mobile Traffic Sensor Routing in Dynamic Transportation Systems," IEEE Transactions on Intelligent Transportation System, vol. 15, no. 5, October 2014.
- [20] Dusadee Apicharttrisor, Kittipat Apicharttrisor and Teerasit Kasetkasem, "A moving object tracking algorithm using support vector machines in binary sensor networks," 13th International Symposium on Communications and Information Technologies (ISCIT), pp. 529-534, September 2013.
- [21] Jia Wei Tang, Nasir Shaikh-Husin, Usman Ullah Sheikh, and M. N. Marsono, "FPGA-Based Real-Time Moving Target Detection System for Unmanned Aerial Vehicle Application," International Journal of Reconfigurable Computing, Vol. 2016, Article ID 8457908, pp. 1-16, 2016.
- [22] Hamid Mahboubi, Walid Masoudimansour, Amir G. Aghdam and Kamran Sayrafian-Pour, "Maximum Lifetime Strategy for Target Monitoring With Controlled Node Mobility in Sensor Networks with Obstacles," IEEE Transactions on Automatic Control, vol. 61, no. 11, pages 3493-3508, November 2016.
- [23] H. Mahboubi, W. Masoudimansour, A. G. Aghdam, and K. Sayrafian-Pour, "Cost-efficient routing with controlled node mobility in sensor networks," in Proceedings of the IEEE Multi-Conference on Systems and Control, pp. 1238-1243, 2011.
- [24] Enyang Xu, Zhi Ding, Fellow,IEEE, and Soura Dasgupta, "Target Tracking and Mobile Sensor Navigation in Wireless Sensor Networks," IEEE Transactions on Mobile Computing, vol. 12, no. 1, pp. 177-186, January 2013.
- [25] Yifan Cai, Simon X. Yang and Xin Xu, "A combined hierarchical reinforcement learning based approach for multi-robot cooperative target searching in complex unknown environments," 2013 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL), pp. 52 - 59, 2013.
- [26] Yifan Cai, Simon X. Yang, Xin Xu and Gauri S. Mittal, "A Hierarchical Reinforcement Learning Based Approach for Multi-robot Cooperation in Unknown Environments," Proceedings of the 2011 2nd International Congress on Computer Applications and Computational Science, vol. 144 of the series Advances in Intelligent and Soft Computing, pp 69-74, 2011.
- [27] Yuanjiang Huang, José-Fernán Martínez, Juana Sendra, and Lourdes López, "The Influence of Communication Range on Connectivity for Resilient Wireless Sensor Networks Using a Probabilistic Approach," Hindawi Publishing Corporation, International Journal of Distributed Sensor Networks, vol. 2013, Article ID 482727, pp. 1-11, 2013.
- [28] Vijay Bhaskar Semwal, Vinay Bhaskar Semwal, Meenakshi Sati, and Dr. Shirshu Verma, "Accurate location estimation of moving object In Wireless Sensor network," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 1 no. 4, pp-71-75, 2011.
- [29] F. Xue and P. R. Kumar, "The number of neighbors needed for connectivity of wireless networks," Wireless Networks, vol. 10, no. 2, pp. 169-181, 2004.
- [30] L.E. Miller, "Distribution of Link Distances in a Wireless Network," J. Research of the National Institute of Standards and Technology, vol. 106, pp. 401-412, 2001.
- [31] S. De, A. Caruso, T. Chaira, and S. Chessa, "Bounds on Hop Distance in Greedy Routing Approach in Wireless Ad Hoc Networks," International Journal Wireless and Mobile Computing, vol. 1, pp. 131- 140, 2006.
- [32] Maneesha V. Ramesh, Divya P. L. et al., "Performance Enhancement in Distributed Sensor Localization using Swarm Intelligence", 2012 International Conference on Advances in Mobile Network, Communication and its Application, 2012.



Mininath K. Nighot

Mininath K. Nighot is perusing Ph.D. in computer science and engineering at Shree Sant Gadgebaba Amravati University, Amravati. He received his master degree in computer engineering from Dr. D.Y. Patil college of engineering, university of Pune, Pune in 2008. He has completed his graduation in computer science and engineering from Babasaheb Naik college of engineering, Pusad, Amravati University in 2002. His research interest includes wireless communication, protocol design for wireless sensor network and effective use of optimization techniques.



Ashok Ghatol

Ashok Ghatol is a former vice chancellor at Dr. Babasaheb Ambedkar Technological University, Lonere. He was executive committee member and vice president of Indian Society for Technical Education New Delhi. He was chairman of AIB of technical education, AICTE, New Delhi in 2009. He was awarded as best teacher by Maharashtra Government in 1998-99. He earned his Ph. D. degree in electrical engineering from IIT Powai, Mumbai. He wrote three books on different technology. He had guided more than 20 Ph.D. and 30 Master students in engineering. Optimization techniques.



Vilas M. Thakare

Vilas M. Thakare is Professor and Head in Post Graduate department of Computer Science and engg, Faculty of Engineering & Technology, SGB Amravati university, Amravati. He is also working as a coordinator on UGC sponsored scheme of e-learning and m-learning specially designed for teaching and research. He is Ph.D. in Computer Science/Engg and completed M.E. in year 1989 and graduated in 1984-85.

Novel Clustering Method Based on K-Medoids and Mobility Metric

Y. Hamzaoui¹, M. Amnai², A. Choukri³, Y. Fakhri¹

¹ LARIT, Networks and Telecommunications Team, Faculty of Sciences Kenitra (Morocco)

² Hassan 1 University, Laboratory IPOSI National School of Applied Sciences (ENSA), 26000 Settat (Morocco)

³ Cadi Ayyad University, S.A.R.S Group, ENSA Safi (Morocco)

Received 2 June 2017 | Accepted 3 October 2017 | Published 3 November 2017



ABSTRACT

The structure and constraint of MANETS influence negatively the performance of QoS, moreover the main routing protocols proposed generally operate in flat routing. Hence, this structure gives the bad results of QoS when the network becomes larger and denser. To solve this problem we use one of the most popular methods named clustering. The present paper comes within the frameworks of research to improve the QoS in MANETS. In this paper we propose a new algorithm of clustering based on the new mobility metric and K-Medoid to distribute the nodes into several clusters. Intuitively our algorithm can give good results in terms of stability of the cluster, and can also extend life time of cluster head.

KEYWORDS

MANETS, Clustering, Mobility, K-medoids.

DOI: 10.9781/ijimai.2017.11.001

I. INTRODUCTION

Ad hoc wireless networks (MANETs) are a new architecture of wireless networks, where a set of mobiles can be deployed easily in a minimum configuration, with free movement and without the existence of the centralized device. It allows a user with a terminal to access to information at any time and from any place. These mobiles cooperate with each other to overcome the Ad hoc network constraints such as dynamic topology, lack of centralized monitoring points, limited bandwidth, etc. The non-centralized management and dynamic topology in Manets require nodes (mobiles) to behave as routers in order to maintain routing information. Many routing protocols have been proposed [1] on different types of application. The research has not ceased to have efficient protocols that adapt to all mobility models [2]. The routing of the information in the Manets can be classified into two types: Flat and Hierarchical. In the first type all network nodes play the same role, this can overload the network, as well as cause other problems such as scalability and complexity, when the network becomes wider and denser. The second type of routing is used to support networks that are wide and dense. The clustering has a hierarchical structure that makes it possible to group geographically the nodes that are neighbors. This allows to each node storing all information about its group and only some information of other groups (clusters). This approach can reduce the cost of routing of information in large and dense networks. Research has not ceased to have efficient protocols that can support this type of network and structure. In order to achieve this goal, researchers propose various metrics for building

and organizing clusters in Manets. These metrics will be the basis for the clustering construct algorithms integrated in the routing protocols in order to optimize and improve them. The mobility metric allows to study the motion of the nodes, it allows to quantify the dynamic topology of the nodes in the Manets. A good metric of mobility makes it possible to better differentiate between the models of mobility [2], and to reflect the real behavior of the nodes. As well as facilitating the performance improvement of protocols.

In this article we propose a new stability metric that refines the degree of mobility, taking into account any type of motion in a coverage area of a mobile node, this metric will be the basis of a cluster building algorithm by using k-medoid to create the groups of mobile, this new approach can generate the more stable cluster and cluster head. In the rest of this paper we start with a presentation of some related work. In the second section we present the problem formulation, and in the third section we define the clustering. Then in the fourth section we present the definition of the distance, in the fifth section we explain our mobility metric, after in the sixth section we present the K-medoid algorithm before we propose the clustering algorithm in seventh section. In eighth section we present the conclusions.

II. RELATED WORK

Routing is among the most important processes in Manets, their performance is related to the density and mobility of the nodes. Several clustering algorithms have been designed and proposed, that are based on different parameters to organize the network into several groups with their cluster head. In this section we will present some interesting works based on the creation of clusters by the use of density and mobility metrics.

In [3] a new approach is described, it is implemented into OLSR (Optimized Link State Routing) protocol to build a cluster and elect a clusters head, it is based on the calculation of the less mobile node,

* Corresponding author.

E-mail addresses: Younes_jeee@hotmail.com (Y. Hamzaoui), mohamed.amnai@uhp.ac.ma (M. Amnai), choukriali@gmail.com (A. Choukri), akhri-youssef@univ-ibntofail.ac.ma (Y. Fakhri).

hence the nodes are affected by the less mobile node. According to the results, this algorithm made a great improvement of cluster stability compared with other density approaches.

A novel clustering algorithm for OLSR is proposed in [4], it gives birth to a new OLSR protocol, based on a metric that combines the two major characteristics of ad hoc networks: mobility and density. The simulation results show, that the proposed approach enhances the routing process and produces a small number of stable (less mobile) cluster heads.

The improvement of an existent study is the goal of authors in [5], it defines a new metric measuring of degree of mobility (for each node) based on incoming and outgoing nodes from a coverage area. The others use this metric to select the MPRs (Multipoint Relays) of OLSR protocol, the new OLSR protocol named Mob2-OLSR is compared with standard OLSR and Mob-OLSR, the results of simulation shows the amelioration of the performance of the Mob2-OLSR compared with others protocols.

In [6], the authors propose a novel energy aware clustering algorithm for the optimized link state routing (OLSR) protocol. This algorithm takes into account the node density and mobility and gives major improvements regarding the number of elected cluster heads and increase the network lifetime.

III. PROBLEM FORMULATION

Several routing protocols ensure a high level of quality of service in Manets networks characterized by low density and low mobility, however when the network becomes denser [7] or more agitated [8], the network generates more control message, routing table, etc. Consequently the network becomes more sensitive, and it doesn't ensure a minimum quality of service. Several studies on routing optimization adopt the clustering method to reduce the costs produced by the MANETs structures. It is based on the distribution of mobiles nodes into groups. In the literature we can have several clustering techniques; the K means clustering technique is among the best methods known [9] used in MANETs. This method is sensitive to outliers and the center of a cluster by calculating k-Means clustering can be empty. Thus, the extreme objects that have the upper values can significantly distort the distribution of clusters and their centers [10]. However these weaknesses can be avoided, if we apply a more robust method than K-Means [11] [12]. This one will allow clusters to be more stable and the center will be better chosen. In this paper we propose the use of K-medoid to produce optimal clusters. This method is based on a metric of mobility proposed by [4] and improved by our approach. In [4] the author has noticed the impact of nodes that join and / or leave the coverage area of a node on the evaluation of their degree of mobility (see Fig. 1(b)). Consequently he proposed the formula (2) which calculates the degree of mobility of a node. This brings relevant results in terms of QoS. Intuitively we added to this formula other measurement parameters that can better reflect the degree of stability of a node. These parameters are: The compute of the nodes that converge and that diverge during a determined duration of time (see Fig.1 (b), (c)). The stability of a node is not enough to elect it as a cluster head. Thus, we will add another parameter that will guarantee a longer lifetime of the cluster head. This parameter is the residual energy of a node, so that a stable node is the one that ensures the performance of these two variables (energy, stability).

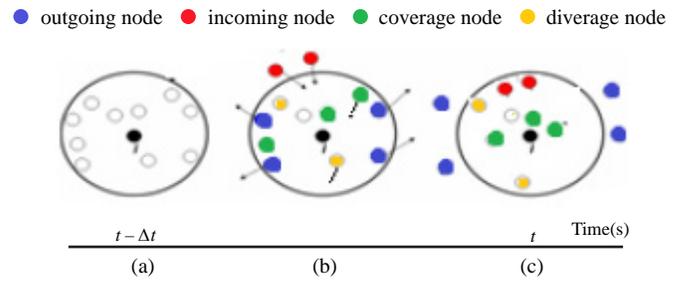


Fig 1. Illustration of movement inside the coverage area of a node.

IV. CLUSTERING

Spatial clustering algorithms can be classified into four categories. There are the partition based, the hierarchical based, the density based and the grid based [13], [14]. According to [15] clustering in ad hoc networks can be defined as a theoretical arrangement of dynamic nodes corresponding to one or more specific properties in different subsets called "Cluster". An element of a cluster is characterized by a strong similarity to components of its group, and a strong dissimilarity with respect to members of other groups [16]. Each cluster is identified by a particular node called "Cluster head". Clustering allows a node to store only part rather than all the information of the network topology. This simplifies the processing of the global topology [17]. This reduces the size of routing tables and thereafter the reduction of the control messages generated by the routing system.

The use of clustering in Manets has several advantages [3], usually a cluster structure allows the node to play one of three roles (Fig. 2):

- Cluster head: A cluster head is elected in the cluster formation process for each cluster. Each cluster should have one and only one cluster head.
- Gateway : A node is called a gateway node of a cluster if it knows that it has a bidirectional or uni-directional link to a node from another cluster.
- Members: All nodes within a cluster except the cluster head are called members of this cluster.

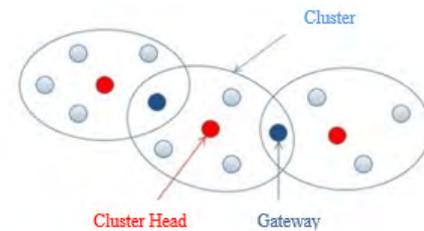


Fig 2. Node types in a cluster.

V. DISTANCE

The choice of the distance is a key issue for classification methods. To offer a relevant measure of similarity between elements, it is necessary to well use the available information at the nodes. The Minkowski distance is the most used to determine the similarity between elements [18]:

$$d(x_i, x_j) = \left(\sum_{k=1}^P |v_k(x_i) - v_k(x_j)|^l \right)^{1/l} \quad (1)$$

Where $v_k(x_i)$ is the value of the object x_i on the variable v_k . Depending on the values taken by the parameter l , we talk about:

- Euclidean distance ($l = 2$);
- Manhattan distance ($l = 1$);
- Chebychev distance ($l = \infty$).

We note that the metrics used to analyze the ad hoc performances such as density, mobility and energy can be used to express distance.

VI. NOVEL METRIC OF STABILITY

The metric that we propose helps to calculate the degree of mobility (stability) of a node; this metric can be used in the calculation of the MPR (case of OLSR protocol) in selecting stable routes (case of AODV protocol) or in clustering process. In our proposal we introduce other parameters as the number of nodes incoming and outgoing in the area coverage of node studied (see Fig. 1). These parameters are the calculation of the numbers of nodes converging and diverging in the coverage area of the node studied.

In [4], only incoming and outgoing nodes in the radio range of the studied node are used to calculate the degree of mobility in the following formula (2):

$$M_i^\lambda(t) = \lambda \frac{\text{NodesOut}(t)}{\text{Nodes}(t - \Delta t)} + (1 - \lambda) \frac{\text{NodesIn}(t)}{\text{Nodes}(t)} \quad (2)$$

In our proposed approach, we noticed that in a scenario of movement (see Fig.1 (a)), it may be that more nodes move towards the studied node. The latter will be as a gravity point (convergence point). Thus these movements towards it will give them stability over time and vice versa (in case of divergence). While we have not only limited on incoming and outgoing nodes, on the other movement being within the radio range of the studied node, we try to refine the calculation of the stability metric of a node.

We selected four types of movements that have influence on the stability or mobility of a node; we find the node that leaves the coverage area, the node that joins the zone, the node that approaches to the studied node and the node that finally moves away from the examined node and that stays in their coverage area. The first two parameters will be retained by the collected control messages. The last two will be calculated by the calculating power for two successive received messages (eg Hello message). We define the following parameters that characterize our metric:

- N_{con} defines the number of nodes that converge on the studied node.
- N_{div} defines the number of nodes that diverge towards the outside of the studied node.
- N_{in} defines the number of nodes within the area of the studied node.
- N_{out} defines the number of nodes out of the coverage area of the studied node.

Following these four types of movement we have created a metric that will rank the nodes between any of these four metrics of stability. We define each as shown in Table I:

TABLE I. CLASSIFYING OF METRICS

Metric 1: Better Stability	$\begin{cases} N_{con} > N_{div} \\ N_{in} > N_{out} \end{cases}$	$\gamma \in [0 \ 0.25[$
Metric 2: Stable	$\begin{cases} N_{con} < N_{div} \\ N_{in} > N_{out} \end{cases}$	$\gamma \in [0.25 \ 0.5]$
Metric 3: Less stable	$\begin{cases} N_{con} > N_{div} \\ N_{in} < N_{out} \end{cases}$	$\gamma \in]0.5 \ 0.75 [$
Metric 4: Poor Stability	$\begin{cases} N_{con} < N_{div} \\ N_{in} < N_{out} \end{cases}$	$\gamma \in [0.75 \ 1]$

In the four classifications, we have the first case that reflects a better stability for the node in question and the latest which represents a poor stability of the studied node, intuitively the second is better than the third, because N_{in} is greater than N_{out} and secondly, even with $N_{div} > N_{con}$ the diverging node stays in the coverage area of the studied node.

We determine subsequently the metric degree of stability that will calculate for each category the best stability node, we use in the formula (3) the coefficient of flow defined in [14], we divide the coefficient in 4 intervals (Table I) and metric degree of stability of node i will be as in (3):

$$M_i^\lambda(t) = \lambda \left(\frac{N_{div}}{N_{con} + N_{div} + N_{out}} + \frac{N_{out}}{N_{con} + N_{div} + N_{out}} \right) + (1 - \lambda) \left(\frac{N_{con}}{N_{con} + N_{div} + N_{in}} + \frac{N_{in}}{N_{con} + N_{div} + N_{in}} \right) \quad (3)$$

VII. THE DESCRIPTION OF K-MEDOID ALGORITHM

K-Medoid is a partitioning technique of clustering of object into k clusters, each cluster is presented by a Medoid, it is the most optimal located object in a cluster. The PAM (Partitioning Around Medoids) is the first K-Medoid algorithm introduced [19]. Initially, the k number of desired clusters is an input and a set of k nodes is taken randomly to be the initial representative medoid of k clusters. The final medoid (object) calculated by PAM is the most centralized position of all objects in a cluster. Thus the PAM algorithm examines in each step, all nodes (one by one) from the input dataset (nodes) that are not currently a medoid and see if they should be one. That is, the algorithm determines whether there is a node that should replace one of the existing medoids to minimize the total error (4). A node is assigned to the cluster represented by the medoid to which it is closest (minimum distance). The PAM algorithm is shown in Fig. 3. We assume that K_i is the cluster represented by medoid t_i . Suppose t_i is a current medoid and we wish to determine whether it should be exchanged with a non-medoid t_h . We wish to do this swap only if the overall impact to the cost (sum of the distances to the cluster medoid) represents an improvement. The total error by a medoid change S_{ih} is given by (4):

$$S = \sum_{h=1}^k \sum_{n_i \in C_h} dis(n_h, n_i) \quad (4)$$

S is the sum of absolute error for all objects in the data set (nodes).

Algorithm of PAM

```

Input
D={t1, t2, ..., tm} // set of nodes
K // Number of desired clusters
Arbitrarily select k medoids from D;
Repeat
  For each th not a medoid do
    For each medoid ti do
      Compute square error function Sih;
      Find i,h where Sih is smallest;
      If Sih < current Sih then;
        Replace medoid ti with th;
    Until Sih >= current Sih;
  For each ti ∈ D do
Assign ti to Kj where dist(ti, tj) is the smallest over all medoids.

```

Fig 3. Algorithm of PAM

VIII. NOVEL APPROACH OF CLUSTERING

The introduction of changes to the standard algorithm is a need, to use the k-medoid method for grouping nodes of MANETs into clusters;

this change aims to determine the K parameters of the algorithm. Firstly it is assumed that each node in Manets is an own cluster; then followed by a sequence of ascending partitions, in the end we reunite the nodes from the same neighborhood in the same cluster. Until reaching a final number of clusters. Thereafter, the K-Medoid algorithm will be used to generate more stable clusters with their cluster head. The parameters used in the medoid calculation algorithm (PAM) are the stability and the residual energy of a mobile, so each node is identified by a vector (stability, energy). This vector will be the basis for expressing the distance between nodes.

Fig. 4 presents our algorithm based on K-medoid to create the partition of the cluster nodes in Manets.

Proposed algorithm
 Input: Mobile ad hoc network of n nodes.
 Output: Network virtually partitioned into P clusters

Step 0

1. Initialization with n medoids ($M_1^0 \dots M_n^0$) each node is a cluster
2. Creation of an initial partition $P_0 = \{(C_1^0 \dots C_k^0)\}$
 - a. Initialize l to 1 (*l is iteration index*)
 - b. Assign to M_l^0 its two-hop neighbors ;
 $C_l^0 = \{x_i \in \text{network} \mid d(x_i, M_l^0)\}$;
 - c. Remove from list of medoids the $n_{C_l^0}$ nodes assigned to medoid M_l^0 ;
 - d. Move to the medoid $l + n_{C_l^0}$
 - e. Repeating steps b to d until all node are affected ;
 - f. Calculation of new medoids of k cluster obtained ($M_1^1 \dots M_k^1$): using the PAM algorithm;

Step t

3. Creation of a new partition $P_t = \{(C_1^t \dots C_k^t)\}$ by assigning to each medoid its two-hop neighbors;
4. the medoids affected by other medoids are removed from the list of medoids;
5. the isolated medoid are assigned to the list of medoids ;
6. Calculate the medoid of k clusters obtained ($M_1^0 \dots M_k^0$) : using PAM algorithm ;
7. Repeat steps 3 to 6 until that a stable partition is achieved (structure of partition P_{t+i} equals that the P_{t+i+1}) or reach n iterations.

Fig 4. Proposed Algorithm of Clustering Method.

IX. CONCLUSION

In order to reduce the routing information costs and increase the QoS in MANETs, we have used the strong method to group the nodes into several clusters. In this paper we have proposed a novel approach of clustering based in K-medoid, and using the new metric of the calculation of degree of mobility. In future work we validate our approach by their implementation in protocol routing. Consequently, this can improve the QoS in MANETs.

REFERENCES

- [1] A. Boukerche, B. Turgut, N. Aydin, M.Z. Ahmad, L. Bölöni, and D. Turgut. Routing protocols in ad hoc networks: a survey. *Computer Networks*, Vol. 55, No. 13, pp. 3032–3080, September 2011.
- [2] These de Yawut en (2009). Titre : Adaptation à la mobilité dans les réseaux ad hoc, institut de Recherche en Informatique de Toulouse.
- [3] M. Dyabi, M. Saadoune, A. Hajami, H. Allali. OLSR Clustering Algorithm Based on Nodes Mobility, *International Review on Computer and Software*, (IRECOS), Vol. 10, No. 1, pp. 36-43, January 2015.
- [4] A. Loutfi, M. Elkoutbi. Optimizing the process of OLSR clustering based on mobility and density in ad hoc networks. *International Conference on Multimedia Computing and Systems (ICMCS)*, pp. 522 – 526, May 2012.
- [5] N. Lakki, A. Ouacha, A. Habbani, M. Ajana el khadar, M. El Koutbi, J. El Abbadi. A new approach for mobility enhancement of olsr protocol. *International Journal of Wireless & Mobile Networks (IJWMN)* Vol. 4, No. 1, February 2012.
- [6] A. Loutfi, M. Elkoutbi, J. BenOthman, A. Kobbane. An energy aware algorithm for olsr clustering. *Annals of telecommunications*, Vol. 69, No. 3, pp. 201–207, April 2014.
- [7] Kwar, B.J., N.O. Song and L. Miller. On the scalability of ad hoc networks. *IEEE Communication Letters*, Vol. 8, No. 8, Aug. 2004.
- [8] C. A. Santiváñez, B. McDonald, I. Stavrakakis, and R. Ramanathan. On the scalability of ad hoc routing protocols, in *Proceedings of the IEEE Infocom*, pp. 1688–1697, June 2002.
- [9] A. M. Bagirov, J. Ugon, and D. Webb. Fast modified global k-means algorithm for incremental cluster construction, *Pattern Recognition*, Vol. 44, pp. 866-876.
- [10] Jiawei Han, Micheline Kamber. *Data Mining Concepts and Techniques*, second edition, Elsevier Science, 2007.
- [11] Norazam Arbin, Nur Suhailayani Suhaimi. Comparative Analysis between K-Means and K-Medoids for Statistical Clustering, *Third International Conference on Artificial Intelligence, Modelling and Simulation*, 2015.
- [12] Noor Kamal Kaur, Usvir Kaur, Dheerendra Singh. K-Medoid Clustering Algorithm- A Review, *International Journal of Computer Application and Technology (IJCAT)* Vol. 1, No. 1, pp. 42-45, 2014.
- [13] Y. Sun, *Radio Network Planning for 2G and 3G*, Master of Science in Communications Engineering, Munich University of Technology, 2004.
- [14] R. Whitaker, S. Hurley. Evolution of planning for wireless communication systems, *Proceedings of the 36th Annual Hawaii International Conference (HICSS'03)*, pp 296, January 2003.
- [15] Angione, G., Bellavista, P., Corradi, A. and Magistretti, E. A k-hop Clustering Protocol for Dense Mobile Ad-Hoc Networks. *26th IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW'06)*. Portugal, pp. 10, 2006.
- [16] Daniel T. Larose. *Discovering Knowledge in Data: An Introduction to Data Mining*, John Wiley & Sons, Inc., Hoboken, New Jersey. 2005.
- [17] Inn, E. and Winston, S. Mobility-Based D-Hop Clustering Algorithm for Mobile Ad Hoc Networks. *IEEE Wireless Communications and Networking Conference*, Vol. 4. Atlanta, GA, USA, Georgia, pp. 2359-2364, 2004.
- [18] A. Choukri, A. Habbani, M. Elkoutbi. Classification par similarité: OLSRKmeans, WCCS14 The 2nd World Conference on Complex Systems, Agadir Morocco, 2014.
- [19] L. Kaufman and P. J. Rousseeuw. *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons, 1990.



Younes Hamzaoui

Younes Hamzaoui obtained his DEUG degree on chemical physics in 2000 from Abdmalik Saadi University, Tanger, Morocco. Then, he received his degree on IEEA (Computers, Electronics, Electrical and Automation) in 2003. After in 2007, he obtained his DESA degree on information system and telecommunication. The author is a professor of computer science in high school. He is a PhD student of management of QoS in routing protocol, also, he is doing his research in the Laboratory in Computer Science and Telecommunications (LaRIT) in the Faculty of Sciences of Kénitra.



Mohamed Amnai

Mohamed Amnai received his bachelor's degree in 2000, in IEEA (Computers, Electronics, Electrical and Automation) from Molay Ismail's University, Errachidia city. Then, he obtained his master's degree in 2007, from Ibn Tofail University, Kenitra city. In 2011, he received his PhD on Telecommunication and computer science, from Ibn Tofail University in Kenitra city, Morocco. Since March 2014 he is a Professor, he has been an Assistant at National School of Applied Sciences Khouribga, Settat University, Morocco. The author is also an associate member of Research Laboratory in Computer Science and Telecommunications (LARIT), Team Networks and Telecommunications Faculty of Science, Kenitra, Morocco. Also, the author is an associate member of laboratory IPOS National School of Applied Sciences, Hassan 1 University, Khouribga, Morocco.



Ali Choukri

Ali Choukri is an Assistant professor at National School of Applied Sciences. He was awarded a Master of Computer Science and Telecom from the University of Ibn Tofail, Kénitra-Morocco in 2008. He got ENSET (higher Normal School of technical teaching) degree in 1992. He has a PhD from the School of Computer Science and Systems Analysis (ENSIAS). He is working within the MIS team of the Laboratory SIME, for studying ad hoc mobile intelligent communication systems, and wireless sensor networks. His research interests are in the areas of ubiquitous computing, Internet of Things, Delay/Disruption Tolerant Networks, Wireless Networks, QoS routing, Mathematical modeling and performance analysis of networks, Control and decision theory, Game theory, trust and reputation management, Distributed algorithms, Metaheuristic and optimization, Genetic algorithms.



Youssef Fakhri

Youssef Fakhri received his Bachelor's Degree (B.S) in Electronic Physics in 2001 and his Master's Degree (DESA) in Computer and Telecommunication from the Faculty of Sciences, University Mohammed V, Rabat, Morocco, in 2003 where he developed his Master's Project at the ICI Company, Morocco. He received a PhD in 2007 from the University Mohammed V - Agdal, Rabat, Morocco in collaboration with the Polytechnic University of Catalonia (UPC), Spain. He joined the Faculty of Sciences of Kénitra, Department of Computer Science and Mathematics, IbnTofail University, Morocco, as an Associate Professor in March 2009, he is the Laboratory head at LaRIT (Laboratory for Research in Computing and Telecommunications) in the Faculty of Kénitra, and Member of Pole of Competences STIC Morocco.

Spectral Restoration Based Speech Enhancement for Robust Speaker Identification

Nasir Saleem^{*1}, Tayyaba Gul Tareen²

¹ Department of Electrical Engineering, Gomal University, D.I.Khan (Pakistan)

² Department of Electrical Engineering, Iqra University, Peshawar (Pakistan)

Received 8 October 2017 | Accepted 21 December 2017 | Published 19 January 2018

unir
LA UNIVERSIDAD
EN INTERNET

ABSTRACT

Spectral restoration based speech enhancement algorithms are used to enhance quality of noise masked speech for robust speaker identification. In presence of background noise, the performance of speaker identification systems can be severely deteriorated. The present study employed and evaluated the Minimum Mean-Square-Error Short-Time Spectral Amplitude Estimators with modified *a priori* SNR estimate prior to speaker identification to improve performance of the speaker identification systems in presence of background noise. For speaker identification, Mel Frequency Cepstral coefficient and Vector Quantization is used to extract the speech features and to model the extracted features respectively. The experimental results showed significant improvement in speaker identification rates when spectral restoration based speech enhancement algorithms are used as a pre-processing step. The identification rates are found to be higher after employing the speech enhancement algorithms.

KEYWORDS

A Priori SNR, Spectral Restoration, Speech Enhancement, Speaker Identification, Mel Frequency Cepstral Coefficients, Vector Quantization.

DOI: 10.9781/ijimai.2018.01.002

I. INTRODUCTION

SPEECH enhancement aspires to improve quality by employing a variety of speech processing algorithms. The intention of the enhancement is to improve the speech intelligibility and/or overall perceptual quality of speech noise masked speech. Enhancement of speech degraded by background noise, called noise reduction is a significant area of speech enhancement and is considered for diverse applications for example, mobile phones, speech/speaker recognition/identification [1] and hearing aids. The speech signals are frequently contaminated by the background noise, which affects the performance of speaker identification (SID) systems. The SID systems are used in online banking, voice mail, remote computer access etc. Therefore, for effective use of such systems, a speech enhancement system must be positioned in front-end to improve identification accuracy. Fig.1 shows the procedural block diagram of speech enhancement and speaker identification system. The algorithms for speech enhancement are categorized into three fundamental classes, (i) filtering techniques including spectral subtraction [2-5] Wiener filtering [6-8] and signal subspace techniques [9-10], (ii) Spectral restoration algorithms including Mean-Square-Error Short-Time Spectral Amplitude Estimators [11-12] and (iii) speech-model based algorithms. The systems presented in [6-8, 11-13] principally depend on accurate estimates of signal-to-noise ratio (SNR) in all frequency bands, because gain is computed as function of spectral SNR. A conventional and recognized technique for SNR estimation is decision-directed (DD) method suggested in [11] The DD technique tails the shape of instantaneous SNR for a priori SNR

estimate and brings one-frame delay. Therefore, to avoid one-frame delay, momentum terms are incorporated to get better tracking speed of system and avoid the frame delay problem. All the mentioned systems in [11-13] can significantly improve speech quality. Binary masking [14-18] is another class that increases speech quality and intelligibility simultaneously. This paper presents Mean-Square-Error Short-Time Spectral Amplitude Estimators with modified *a priori* SNR estimation to reduce background noise and to improve identification rates of speaker identification systems in presence of background noises. The paper is prepared as follows. Section 2 presents the overview of speech enhancement system; section 3 gives speaker identification system; section 4 presents the experimental setup, results and discussions, and section 5 presents the summary and concluding remarks. The Matlab R2015b is used to construct the algorithms and simulations.

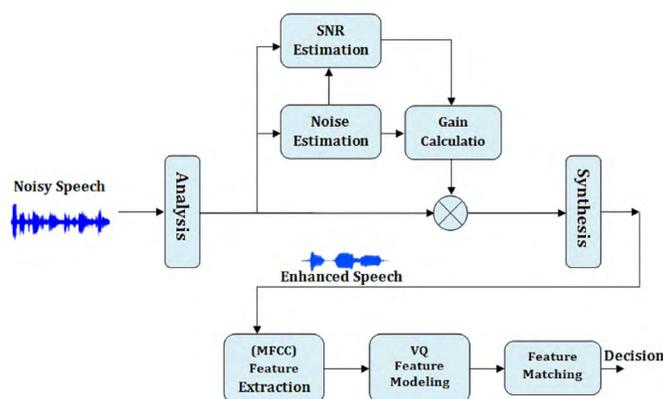


Fig. 1. Procedural block diagram of Speech enhancement and speaker identification system.

* Corresponding author.

E-mail address: nasirsaleem@gu.edu.pk

II. SPECTRAL RESTORATION BASED SPEECH ENHANCEMENT SYSTEM

In classical spectral restoration based speech enhancement system, the noisy speech is given as; $y(t) = s(t) + n(t)$, where $s(t)$ and $n(t)$ specify clean speech and noise signal respectively. Let $Y(k, \omega_k)$, $S(k, \omega_k)$ and $N(k, \omega_k)$ show $y(t)$, $s(t)$ and $n(t)$ respectively with spectral element ω_k and time frame k . The quasi-stationary nature of speech is considered in frame analysis since noise and speech signals both reveal non-stationary behavior. A speech enhancement algorithm involves in multiplication of a spectral gain $G(k, \omega_k)$ to short-time spectrum $Y(k, \omega_k)$ and the computation of spectral gain follows two key parameters, a *posteriori* SNR and the *a priori* SNR estimation:

$$\gamma(k, \omega_k) = \frac{|Y(k, \omega_k)|^2}{E\{|N(k, \omega_k)|^2\}} = \frac{|Y(k, \omega_k)|^2}{\sigma_n^2(k, \omega_k)} \quad (1)$$

$$\xi(k, \omega_k) = \frac{E\{|S(k, \omega_k)|^2\}}{E\{|N(k, \omega_k)|^2\}} = \frac{\sigma_s^2(k, \omega_k)}{\sigma_n^2(k, \omega_k)} \quad (2)$$

Where $E\{\cdot\}$ shows expectation operator, $\gamma(k, \omega_k)$ and $\xi(k, \omega_k)$ presents a *posteriori* SNR estimation and a *a priori* SNR estimation. In practical implementations of a speech enhancement system, squared power spectrum density of clean speech $|X(k, \omega_k)|^2$ and noise $|D(k, \omega_k)|^2$ are unrevealed as only noisy speech is available. Therefore; both instantaneous and a *a priori* SNR need to be estimated. The noise power spectral density is estimated during speech gaps exploiting standard recursive relation, given as:

$$\hat{\sigma}_n^2(k, \omega_k) = \beta \hat{\sigma}_n^2(k-1, \omega_k) + (1-\beta) \sigma_v^2(k-1, \omega_k) \quad (3)$$

Where, β is the smoothing factor and $\hat{\sigma}_v^2(k-1, \omega_k)$ is estimation in previous frame. The SNR can be calculated as:

$$\text{SNR}_{\text{INST}}(k, \omega_k) = \frac{|S(k, \omega_k)|^2}{|N(k, \omega_k)|^2} \quad (4)$$

$$\xi_{\text{DD}}(k, \omega_k) = \alpha \frac{|G(k-1, \omega_k) * Y(k, \omega_k)|^2}{\hat{\sigma}_n^2(k, \omega_k - 1)} + (1-\alpha) F\{\gamma(k, \omega_k) - 1\} \quad (5)$$

Where α is smoothing factor and has a constant value 0.98, $\xi_{\text{DD}}(k, \omega_k)$ is a *a priori* noise estimate via decision-direct (DD) method whereas $F\{\cdot\}$ is half-wave rectification. By setting α as a fixed value near to 1, the DD approach introduces less residual noise. However, it may lead to delay in estimation since a fixed value cannot track the rapid change of speech. The DD is an efficient method and achieves well in speech enhancement applications however; the *a priori* SNR follows the shape of instantaneous SNR and brings single-frame delay. To overcome single-frame delay, a modified form of DD method is used to estimate a *a priori* SNR. The modified *a priori* SNR is written as:

$$\mu_{\text{MDD}}(k, \omega_k) = \frac{|G(k-1, \omega_k) * Y(k, \omega_k)|^2}{\hat{\sigma}_n^2(k, \omega_k - 1)} + \mu(k, \omega_k) + (1-\alpha) F\{\gamma(k, \omega_k) - 1\} \quad (6)$$

$$\mu(k, \omega_k) = \zeta [\xi_{\text{PRIOR}}(k-1, \omega_k) - \xi_{\text{PRIOR}}(k-2, \omega_k)] \quad (7)$$

Equation (6) shows the modified DD (MDD) version used in the speech enhancement system, α is smoothing parameter ($\alpha=0.98$), ζ is

momentum parameter ($\zeta=0.998$), $\mu(m, \omega_k)$ shows momentum terms and $\lambda_D(m, \omega_k)$ is the estimation of background noise variance. The $\xi_{\text{MDD}}(k, \omega_k)$ shows a *a priori* SNR estimation after modification. The estimated power spectrum of the clean speech magnitude $S_{\text{EST}}(k, \omega_k)$ is attained by multiplying gain function with noisy speech $Y(k, \omega_k)$ as:

$$|S_{\text{EST}}(k, \omega_k)| = |Y(k, \omega_k)| * G(k, \omega_k) \quad (8)$$

The gain function $G(k, \omega_k)$ is given as:

$$G(k, \omega_k) = \min \left\{ \zeta, \frac{\xi(k, \omega_k)}{1 + \xi(k, \omega_k)} \left[\frac{1}{2} \int_{v(k, \omega_k)}^{\infty} \frac{e^{-t}}{t} dt \right] \right\}$$

$$v(k, \omega_k) = \frac{\xi(k, \omega_k)}{1 + \xi(k, \omega_k)} \gamma(k, \omega_k) \quad (9)$$

Where, ζ is used to avoid large gain values at low a *posteriori* SNR and $\zeta=10$ is chosen here.

III. SPEAKER IDENTIFICATION SYSTEM

The intention of a Speaker identification system is to identify information regarding any speaker which is categorized into two sub-categories called as Speaker identification (SID) and speaker Verification (SVR). For SID, the Mel Frequency Cepstral coefficient (MFCC) and Vector Quantization (VQ) is used to extract the speech features and to model the extracted features respectively. The speaker identification system drives in two stages, the training and testing stages. In training mode the system is allowed to create the database of speech signals and formulate a feature model of speech utterances. In testing mode, the system uses information provided in database and attempts to segregate and identify the speakers. Here, the Mel frequency Cepstral Coefficients (MFCCs) features are used for constructing a SID system. The extracted features of speakers are quantized to a number of centroids employing vector quantization (VQ) K-means algorithm. MFCCs are computed in training as well as in testing stage. The Euclidean distance among MFCCs of all speakers in training stage to centroids of isolated speaker in testing stage is calculated and a particular speaker is identified according to minimum Euclidean distance.

A. Feature Extraction

The MFCCs are acquired by pre-emphasis of speech initially to emphasize high frequencies and eliminate glottal and lip radiations. The resulting speech is fragmented, windowed, and FFT is computed to attain spectra. To estimate human auditory system, triangular band-pass filters bank is utilized. A linear scale is used to compute center frequencies which are lower than 1 kHz, while logarithmic scale is considered for center frequencies higher than 1 kHz. The filter bank response is given in Fig. 2. The Mel-spaced filter bank response is given as:

$$\text{Mel}(f) = 2595 \log\left(1 + \frac{f}{700}\right) \quad (10)$$

The DFT is computed on log of Mel spectrum to figure Cepstrum as:

$$M_k = \sqrt{\frac{2}{N_f}} \sum_{n=1}^{N_f} \log(\hat{S}(n)) \cos\left(\frac{g\pi}{N_f} (n-0.5)\right) \quad (11)$$

Where M_g shows MFCCs, \hat{S} is n^{th} Mel filter output, K is number of

MFCCs chosen between 5 to 26, and N_f is the number of Mel filters. Initially few coefficients are considered since most of the specific information about speakers is present in them.

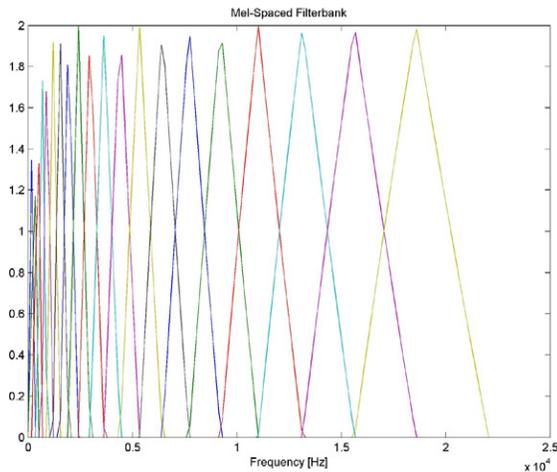


Fig. 2. Mel-Spaced Filter bank Response.

B. Vector Quantization

Vector quantization (VQ) is a lossy compression method based on the block coding theory [20]. The purpose of VQ in speaker recognition systems is to create a classification system for every speaker and a large set of acoustic vectors are converted to lesser set that signifies centroids of distribution shown in Fig. 3. The VQ is employed since all MFCC generated feature vector cannot be stored and extracted acoustic vectors are clustered into a set of codewords (referred to as codebook) and this clustering is achieved by using the K-Means Algorithm which separates the M feature vectors into K centroids. Initially K cluster-centroids are chosen randomly within M feature vectors and then all feature vectors are allocated to nearby centroid, and the creating the centroids, all other new clusters follow the same pattern. The process keeps on until a certain condition for stopping is reached, i.e., the mean square error (MSE) among acoustic vector and cluster centroid is lower than a certain predefined threshold or there are no additional variations in cluster-center task [21].

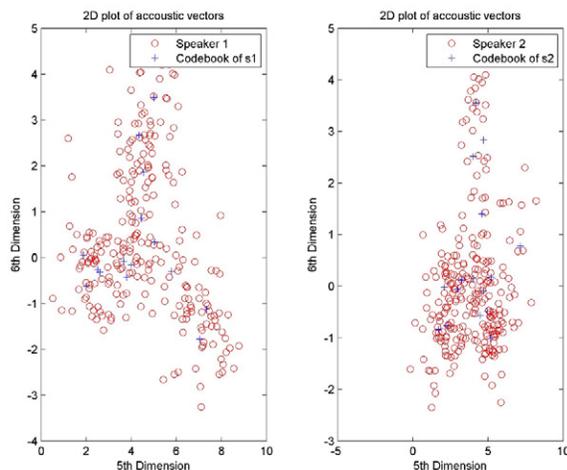


Fig. 3. 2D acoustic Vector analysis for speakers.

C. Speaker Identification

The speaker recognition phase is characterized by a set of acoustic feature vectors $\{M1, M2, \dots, Mt\}$, and is judged against codebooks in list. For all codebooks a distortion is calculated, and a speaker with the

lowest distortion is selected, and this distortion is the sum of squared Euclidean distances among vectors and their centroids. As a result, all feature vectors in M sequence are compared with codebooks, and the codebooks with the minimum average distance are selected. The Euclidean distance between two points, $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$ and $\eta = (\eta_1, \eta_2, \dots, \eta_n)$ is given by [21-22]:

$$\sqrt{\left[(\lambda_1 - \eta_1)^2 + (\lambda_2 - \eta_2)^2 + \dots + (\lambda_n - \eta_n)^2 \right]} = \sqrt{\sum_{i=1}^n (\lambda_i - \eta_i)^2} \quad (12)$$

IV. RESULTS AND DISCUSSION

Six different speakers, three male and three female, were selected from Noizeus [23] and TIMIT database, respectively, while 50 speech sentences uttered by the speakers are considered during training stage for speaker identification. In testing stage, speech utterances are selected at random to access the identification rates. To evaluate performance of system, four signal-to-noise ratio levels, including 0dB, 5dB, 10dB and 15dB are used. Also three noisy situations including car, street and white noise are used to degrade the clean speech. The Perceptual evaluation of speech quality (PESQ) [23] and Segmental SNR (SNRSeg) [24] is used to predict the speech quality after speech enhancement. Three sets of experiments are conducted to measure the speaker identification rates including, clean speech with no background noise, speech degraded by background noise and speech processed by the spectral restoration enhancing algorithms. The presented system is compared to various baseline state-of-art speech enhancement algorithms. The baseline algorithms include MMSE, Spectral subtraction (SS), and signal subspace (Sig_Sp). Table I shows the PESQ scores obtained with the spectral restoration based algorithm and baseline algorithms. The proposed algorithm performed very well in noisy environments and at all SNR levels against baseline speech enhancement algorithms. A considerable improvement in PESQ scores is evident which shows that the proposed speech enhancement algorithm effectively reduced various background noise sources from target speech. Similarly, Fig. 4 shows PESQ scores obtained after applying Minimum Mean-Square-Error Short-Time Spectral Amplitude Estimators with modified *a priori* SNR estimate (MMSE-MDD). The modified version offers the best results consistently in all SNR levels and noisy conditions when compared to noisy and speech processed by traditional MMSE-STSA speech enhancement algorithm. Table II shows the SNRSeg results obtained with the spectral restoration based algorithm and baseline algorithms. Again in terms of SNRSeg, the proposed speech enhancement algorithm outperformed against baseline algorithms. Significant SNRSeg improvements are evident from the obtained results. Fig. 5 shows the speech quality in terms of segmental SNR (SNRSeg) where highest SNRSeg scores are obtained with MMSE-MDD. The enhanced speech associated with six speakers is tested for speaker identification. Table III offers the percentage identification rates achieved with proposed speech enhancement algorithm against baseline algorithms. The speaker identification rates are remarkably improved with the proposed algorithm in various noise environments at all SNR levels as compared to baseline algorithms and unprocessed noisy speech. At low SNR (0dB) a significant increase in identification rates is observed in all noise environments which clearly showed that the noise is effectively eliminated. Fig. 6 shows the identification rates, the lowest identification rates are observed in presence of background noise (Babble, car and street) however, employment of the speech enhancement before speaker identification has tremendously increased the identification rates which are evident in Fig.5. The identification rates for MMSE-MDD are higher in all SNR conditions and levels.

TABLE I. PESQ ANALYSIS AGAINST BASELINE SPEECH ENHANCEMENT ALGORITHMS AND NOISY SPEECH

Noise Type	SNR (in dB)	Noisy Speech	Spectral Subtraction	Signal Subspace	MMSE	Proposed
Babble Noise	0	1.72	1.89	1.91	1.89	1.97
	5	2.11	2.19	2.29	2.23	2.35
	10	2.43	2.53	2.61	2.55	2.69
	15	2.66	2.71	2.76	2.71	2.83
Car Noise	0	1.79	1.91	2.01	1.87	2.07
	5	1.97	2.23	2.31	2.21	2.45
	10	2.31	2.42	2.62	2.61	2.72
	15	2.45	2.56	2.76	2.78	2.91
Street Noise	0	1.77	1.93	1.96	1.88	2.13
	5	2.05	2.21	2.31	2.12	2.43
	10	2.41	2.57	2.59	2.55	2.69
	15	2.54	2.65	2.69	2.61	2.86

TABLE II. SEGMENTAL SNR (SNRSEG) ANALYSIS AGAINST BASELINE SPEECH ENHANCEMENT ALGORITHMS AND NOISY SPEECH

Noise Type	SNR (in dB)	Noisy Speech	Spectral Subtraction	Signal Subspace	MMSE	Proposed
Babble	0	0.11	1.21	1.55	1.12	1.66
	5	1.13	1.77	1.89	1.83	2.01
	10	1.45	2.11	2.17	1.99	2.37
	15	1.64	2.34	2.38	2.28	2.44
Car	0	0.10	1.32	1.28	1.13	1.63
	5	1.23	1.89	1.93	1.78	1.98
	10	1.56	2.14	2.21	1.97	2.41
	15	1.66	2.29	2.33	2.37	2.57
Street	0	0.18	1.29	1.41	1.16	1.59
	5	1.43	1.88	1.92	1.72	1.99
	10	1.53	2.21	2.23	2.01	2.39
	15	1.67	2.35	2.39	2.21	2.51

TABLE III. SPEAKER IDENTIFICATION RATES OF SPEECH ENHANCEMENT ALGORITHMS (IN PERCENTAGE)

Noise Type	SNR (in dB)	Noisy Speech	Spectral Subtraction	Signal Subspace	MMSE	Proposed
Babble	0	41	52	55	56	62
	5	58	64	67	69	71
	10	77	81	83	84	79
	15	85	88	89	88	91
Car	0	40	51	53	55	58
	5	56	66	69	71	73
	10	76	81	85	87	88
	15	82	89	89	90	91
Street	0	38	49	54	57	59
	5	46	67	69	71	73
	10	71	80	82	86	88
	15	80	85	87	90	92

V. SUMMARY AND CONCLUSIONS

This paper presents the Mean-Square-Error Short-Time Spectral Amplitude Estimators with modified *a priori* SNR estimation to reduce the background noise and to improve identification rates of speaker identification systems in presence of background noises. The lowest identification rates are reported when background noises such as babble, car and street are present. By implementing the proposed speech enhancement algorithm as pre-processing step, the identification rates are increased about 40%, 38% and 35% at low SNR level (0dB) in all noise environments. The proposed speech enhancement algorithm

offered significant improvements in terms of PESQ and SNRSeg scores. The speaker identification rates are higher than baseline algorithms in all noise environments and at all SNR levels consistently. In presence of noise, it is difficult to identify specific speaker, however; the use of a speech enhancement system prior to speaker identification remarkably increased the identification rates. On the basis of experimental results, it is concluded that the use of the proposed speech enhancement algorithm as preprocessor can remarkably increase the speaker identification in many noisy environments as compared to many other speech enhancement algorithms.

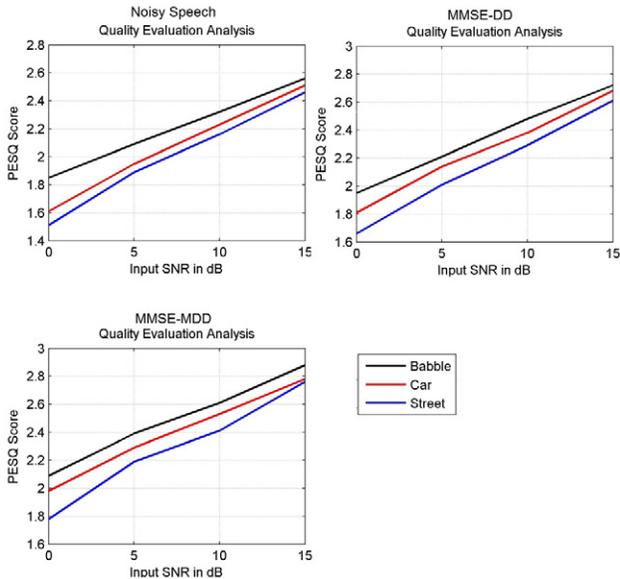


Fig. 4. PESQ Analysis.

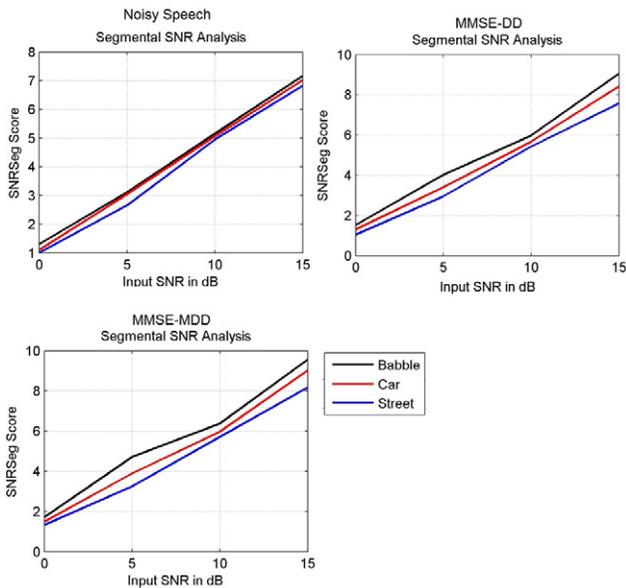


Fig. 5. SNRSeg Analysis.

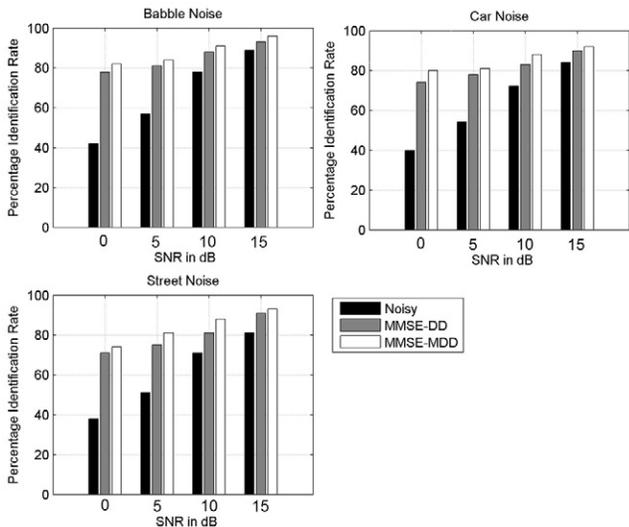


Fig. 6. Speaker identification rate analysis.

- [1] Hicham, E.M., Akram, H., Khalid, S. (2016) Using features of local densities, statistics and HMM toolkit (HTK) for offline Arabic handwriting text recognition. *J. Electr. Syst. Inform. Technol.*, 4(3), 387-396. <http://dx.doi.org/10.1016/j.jesit.2016.07.005>
- [2] Berouti, M., Schwartz, M., and Makhoul, J. (1979). Enhancement of speech corrupted by acoustic noise. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp: 208-211.
- [3] Kamath, S. and Loizou, P. (2002). A multi-band spectral subtraction method for enhancing speech corrupted by colored noise. *IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 4, pp. 44164-44164.
- [4] Gustafsson, H., Nordholm, S., and Claesson, I. (2001). Spectral subtraction using reduced delay convolution and adaptive averaging. *IEEE Trans. on Speech and Audio Processing*, 9(8), 799-807.
- [5] Saleem N, Ali S, Khan U and Ullah F, (2013). Speech Enhancement with Geometric Advent of Spectral Subtraction using Connected Time-Frequency Regions Noise Estimation. *Research Journal of Applied Sciences, Engineering and Technology*, 6(06), 1081-1087.
- [6] Lim, J. and Oppenheim, A. V. (1978). All-pole modeling of degraded speech. *IEEE Trans. Acoust., Speech, Signal Proc.*, ASSP-26(3), 197-210.
- [7] Scalart, P. and Filho, J. (1996). Speech enhancement based on a priori signal to noise estimation. *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 629-632.
- [8] Hu, Y. and Loizou, P. (2004). Speech enhancement based on wavelet thresholding the multitaper spectrum. *IEEE Trans. on Speech and Audio Processing*, 12(1), 59-67.
- [9] Hu, Y. and Loizou, P. (2003). A generalized subspace approach for enhancing speech corrupted by colored noise. *IEEE Trans. on Speech and Audio Processing*, 11, 334-341.
- [10] Jabloun, F. and Champagne, B. (2003). Incorporating the human hearing properties in the signal subspace approach for speech enhancement. *IEEE Trans. on Speech and Audio Processing*, 11(6), 700-708.
- [11] Ephraim, Y. and Malah, D. (1984). Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator. *IEEE Trans. Acoust., Speech, Signal Process.*, ASSP-32(6), 1109-1121.
- [12] Ephraim, Y. and Malah, D. (1985). Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. *IEEE Trans. Acoust., Speech, Signal Process.*, ASSP-23(2), 443-445.
- [13] Cohen, I. (2002). Optimal speech enhancement under signal presence uncertainty using log-spectra amplitude estimator. *IEEE Signal Processing Letters*, 9(4), 113-116.
- [14] Saleem, N. (2016), Single channel noise reduction system in low SNR. *International Journal of Speech Technology*, 20(1), 89-98. doi: 10.1007/s10772-016-9391-z
- [15] Saleem, N., Mustafa, E., Nawaz, A., & Khan, A. (2015). Ideal binary masking for reducing convolutive noise. *International Journal of Speech Technology*, 18(4), 547-554. doi:10.1007/s10772-015-9298-0
- [16] Saleem, N., Shafi, M., Mustafa, E., & Nawaz, A. (2015). A novel binary mask estimation based on spectral subtraction gain induced distortions for improved speech intelligibility and quality. *Technical Journal, UET, Taxila*, 20(4), 35-42.
- [17] Boldt, J. B., Kjems, U., Pedersen, M. S., Lunner, T., & Wang, D. (2008). Estimation of the ideal binary mask using directional systems. In *Proc. Int. Workshop Acoust. Echo and Noise Control*, pp. 1-4.
- [18] Wang, D. (2008). Time-frequency masking for speech separation and its potential for hearing aid design. *Trends in Amplification*, 12(4), 332-353. doi:10.1177/1084713808326455
- [19] Wang, D. (2005). On ideal binary mask as the computational goal of auditory scene analysis. In *Speech separation by humans and machines*, pp: 181-197. doi:10.1007/0-387-22794-6_12
- [20] Gray R.M. (2013). Vector Quantization. *IEEE ASSP Magazine*, 1(2), 4-29.
- [21] Likas A., Vlassis and Verbeek J. J., (2003). The global k-means clustering algorithm. *Pattern Recognition*, 36(2), 451-461.
- [22] Khan S. S and Ahmed A. (2004). Cluster center initialization for Kmeans algorithm. *Pattern Recognition Letters*, 25(11), 1293-1302.
- [23] Hu Y. and Loizou P. (2007). Subjective evaluation and comparison of speech enhancement algorithms. *Speech Commun.*, 49(7-8), 588-601. doi:10.1016/j.specom.2006.12.006

- [24] Rix A.W., Beerends J. G., Hollier M. P., Hekstra A.P. (2001). Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs. In Acoustics, Speech, and Signal Processing (ICASSP), 749–752. doi: 10.1109/ICASSP.2001.941023



Nasir Saleem

Engr. Nasir Saleem received the B.S degree in Telecommunication Engineering from University of Engineering and Technology, Peshawar-25000, Pakistan in 2008 and M.S degree in Electrical Engineering from CECOS University, Peshawar, Pakistan in 2012. He was a senior Lecturer at the Institute of Engineering and Technology, Gomal University, D.I.Khan-29050, Pakistan.

He is now Assistant Professor in Department of Electrical Engineering, Gomal University, Pakistan. Currently, he is pursuing Ph.D. degree in electrical Engineering from University of Engineering and Technology, Peshawar-25000, Pakistan. His research interests are in the area of digital signal processing, speech processing and speech enhancement.



Tayyaba Gul Tareen

Engr. Tayyaba Gul Tareen received the B.S degree in Electrical Engineering from University of Engineering and Technology, Peshawar-25000, Pakistan in 2008 and M.S degree in Electrical Engineering from CECOS University, Peshawar, Pakistan in 2012. Currently, She is pursuing Ph.D. degree in electrical Engineering from Iqra University, Peshawar-25000, Pakistan. Her research interests are in the

area of digital signal processing.

Hybrid Model for Passive Locomotion Control of a Biped Humanoid: The Artificial Neural Network Approach

Manish Raj^{1*}, Vijay Bhaskar Semwal^{2*}, G.C.Nandi¹

¹ Department of AI and Robotics, Indian Institute Information Technology, Allahabad (India)

² Department of Computer Science, Indian Institute of Information Technology, Dharwad (India)

Received 20 January 2017 | Accepted 25 May 20176 | Published 3 October 2017



ABSTRACT

Developing a correct model for a biped robot locomotion is extremely challenging due to its inherently unstable structure because of the passive joint located at the unilateral foot-ground contact and varying configurations throughout the gait cycle, resulting variation of dynamic descriptions and control laws from phase to phase. The present research describes the development of a hybrid biped model using an Open Dynamics Engine (ODE) based analytical three link leg model as a base model and, on top of it, an Artificial Neural Network based learning model which ensures better adaptability, better limits cycle behaviors and better generalization while negotiating along a down slope. The base model has been configured according to the individual subjects and data have been collected using a novel technique through an android app from those subjects while walking down a slope. The pattern between the deviation of the actual trajectories and the base model generated trajectories has been found using a back propagation based artificial neural network architecture. It has been observed that this base model with learning based compensation enables the biped to better adapt in a real walking environment, showing better limit cycle behaviors. We also observed the bounded nature of deviation which led us to conclude that the strategy for biped locomotion control is generic in nature and largely dominated by learning.

KEYWORDS

Legged Locomotion, Passive Walking, Hybrid Model, Back Propagation Algorithms, Artificial Neural Networks, Error Analysis.

DOI: 10.9781/ijimai.2017.10.001

I. INTRODUCTION

BIPED locomotion has inherent complexity due to high degrees of non-linearity, high dimensionality, under actuation (in swing phase), over actuation (in stance phase). Consequently, correct analytical model and intuitive solutions are difficult to conceive and hard to find [1]. Researchers have tried to make kinematics based biped locomotion models and formulate control strategy corresponding to several sub-phases of walking gait and tried to develop a controller for entire gait cycle by superimposing the control strategy of the control sub-phases [2]. However, due to non-linearity, the condition of superposition is not appropriate and hence the overall control strategy based on the model failed. To mitigate this, a number of heuristics were incorporated in the control strategy [3], [14], [15]. Appreciating this inherent complexity, here we propose to develop a computational model based on the captured data of human bipeds. The analytical model has been customized with the corresponding humanoid morphological parameters and its output has been stored [4], [5], [16]. Now the deviation of trajectories between the actual gait pattern and the output of the analytical base model has been calculated [6], [7]. To capture and subsequently predict this error

pattern a neural network has been trained by using back-propagation algorithm. With this trained neural network, we are in a position to correct the analytical base model by compensating the difference. Through this compensation, we are incorporating the non-linearity of the actual gait in the output of the analytical model. We also observed the bounded nature of the difference which led us to conclude that the strategy for human locomotion control is generic in nature [8], [9], [13].

This paper has been arranged in the following manner, in the next section different types of biped models have been described followed by their critical analysis. In the subsequent section, we have developed an ODE based base model, which will serve the purpose of the analytical part of our hybrid model. Subsequently, we have described the details of our experimental data and the consequent error analysis. For the paper, a model was developed and section fifth explains its main features. Then the last section describes our results and discussions.

II. THE ANALYTICAL BASE MODEL

Due to computational simplicity, we prefer the analytical model. Here we have developed three link ODE based analytical model. The Analytical model provides the real world trajectories. These trajectories will be the input to our neural net which will then give error according to the training set provided to neural net. The generated error will be added to the real world trajectories to give stable limit

* Corresponding author.

E-mail addresses: rajmanish.03@gmail.com (Manish Raj), vsemwal@gmail.com (Vijay Bhaskar Semwal).

cycle. Real world trajectories are parameterized according to various morphological parameters like mass of hip, shank, and thigh, length of the shank and thigh part of the human subject.

Analytical model provides a stable limit cycle for almost all parameters, but starts falling for angles less than 4 degrees of inclination. So for those low degrees of inclination, we have to provide stable trajectories using artificial neural network. Use of analytical model in our project can be seen in the Fig 1. Since analytical model plays such a vital role in our project we need to do the critical analysis of analytical model that includes (a) description of model, (b) input and output to analytical model, (c) error control method used in the model, and (d) description of mathematical equations used in the model. Due to high requirements and increasing density of robots in our daily lives, to make them move more efficiently through their environments is very important. Robots with more than two legs are more stable than the ones with two legs. Even then, due to their deficiency of incorporating in real world we have to move towards two legged robots or bipedal locomotion [10]. This is the reason there have been many models regarding that and ours is one of them.

A. The Analytical Model Development

We have considered a three-link biped as shown in Fig. 1. In our model, the stance phase leg is considered to be a single link, whereas, the swing leg has a knee joint. The system is represented by 3 variable angles,

1. q_1 : Angle between the normal to the ground the fixedleg (corresponding to stance phase)
2. q_2 : Angle between the two legs
3. q_3 : Knee angle of the swing leg

B. Static Properties of Model

Model consists of five parts that are for each leg two parts, shank the lower leg and thigh the upper leg, and the additional mass m_H i.e., the mass of hip which is considered as point mass. All the joints are considered friction less but ground friction is not zero, in fact it is high enough so that no sliding can occur. All the three links have their mass and their corresponding moment of inertia which is located at the centre of these respective links. The leg that stands firmly on ground is called stance leg and the other leg that will move freely will be called swing leg. Gamma g is the slope angle which is considered positive on downward slopes. Hip will be the intersection of both the legs, which we consider as one link. All these static properties will be adjusted according to the individual subject parameters in our model shown in Fig.1. All these properties or terms or parameters related to our model are shown in Table 1.

C. Dynamic Properties

The dynamic properties in our model will be the angular velocities and the joint angles of the knee, hip and the ankle joints because all the joints specified in the static properties are rotational joints. The degree of freedom (d.o.f) for our model will be three. One leg will always be in stance position so, its knee d.o.f will be omitted. We will consider the knee d.o.f for only the swing leg whereas the third is for the angle between the knees and the hip. So we need to consider only three angles and their corresponding velocities. The various angles can be described in two ways based on their reference to world reference frame or relative to their parent frame. We have selected the second option that is relative to their parent frame. By applying the forward dynamic calculations on these dynamic properties, we fined out our limit cycle that will provide the criteria for stability.

TABLE 1. STATIC PROPERTIES OF THE BASE MODEL

Variable	Explanation of the variable used	How is the variable given its value
Thigh	Length of the thigh	Input
Shank	Length of the shank	Input
L	Total length of the leg	$L = \text{thigh} + \text{shank}$
lst_1	Stance shank, lower length	Input
lst_2	Stance thigh, lower length	Input
lsw_1	Swing thigh, lower length	Input
lsw_2	Swing shank, lower length	Input
ust_1	Stance shank, upper length	$ust_1 = \text{shank} - lst_1$
ust_2	Stance thigh, upper length	$ust_2 = \text{thigh} - lst_2$
usw_1	Swing thigh, upper length	$usw_1 = \text{thigh} - lsw_1$
usw_2	Swing shank, upper length	$usw_2 = \text{shank} - lsw_2$
m_1	Stance shank mass	Input
m_2	Stance leg mass	Input
m_3	Swing leg mass	Input
m_4	Swing shank mass	Input
M_H	Hip mass	Input
g	Gravitational constant	Input
Gamma (γ)	Angle of inclination	Input

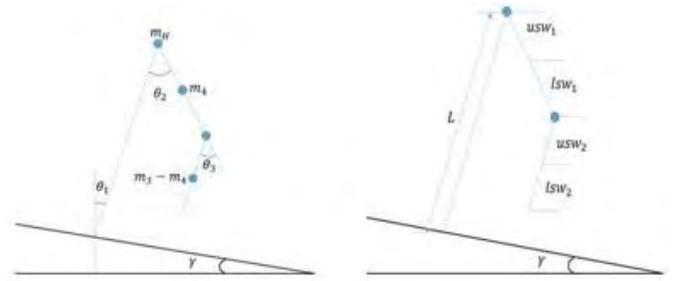


Fig. 1. Static Properties of the model.

D. The Mathematical Model

The analytical model models the bipedal walking by dividing the entire gait cycle into the following phases and computes the dynamics of each phase separately:

- Three link phase: It is the most important phase in which the swing leg moves freely and knee is bent.
- Knee strike: It is the time when the knee of swing leg gets straight and the knee gets locked. At this point of time the system loses one of its degree of freedom because the swing leg becomes one single link instead of two separate parts (shank, thigh). And so shank part loses its angular momentum but conserving the angular momentum of whole system. This phase is considered to be the discrete phase of our step cycle. Now the three link model has converted into the compass gait model which is the predecessor of our three link model.
- Heel Strike: It is the phase which marks the end of one step cycle i.e. one gait. At this point of time the heel just strikes the ground after which the stance leg will become the swing leg and previous swing leg will become the stance leg and it again comes in continuous three link phase. The heel strike is assumed to be the inelastic collision.

The dynamic equations of the model while negotiating with the down slope under the force of gravity can be written in the coupled form as shown in (1).

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = \tau \quad (1)$$

Where $M(q)$ denotes the inertia matrix, $C(q, \dot{q})$ denotes Coriolis and Centrifugal vector and $G(q)$ will be the gravity vector, and all these matrices use avariant which is the angle with respect to the parent reference frame.

III. SIMULATION OF THE MODEL

Since the equations of motion in the three link model are nonlinear, we need to solve it by integrating the equations numerically. Dynamics of this model includes four phases that are categorized in two domains continuous dynamics that can be defined by the set of differential equations and discrete phases. Continuous dynamics include three link phase and two link phase. Whereas discrete phase includes knee strike and heel strike. Continuous phase are the passive elements of analytical model which means that during these phases analytical model performs passive walking. All these four phases are shown in Fig 2. Output of this model will be the expected trajectories that are based on the various parameters provided to analytical model like all static properties: mass of hip, shank and thigh and length of shank ,thigh and the slope angle which is one of the most important factor of input. On the basis of output provided by model we will find our error or the deviation of the trajectories from the analytical model to that of the real human subject, that will then be used to train the neural net which will then predict the error pattern for the test subjects, provided all parameters are same for both the connectionist and analytical model, so that after compensating that error we can get stable limit cycle. One example of output can be seen in the set of Fig. 3. We have shown the limit cycle analysis of knee and hip joints as well as the output of the walker for 7 degree. Fig. 4 is the Limit Cycle for analysis of knee and hip joints for 4 degree walker.

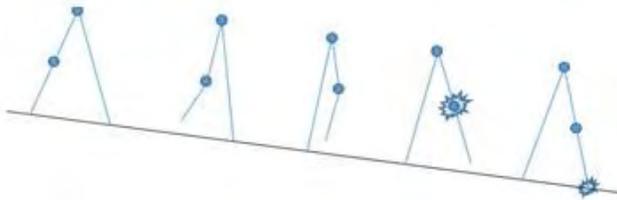
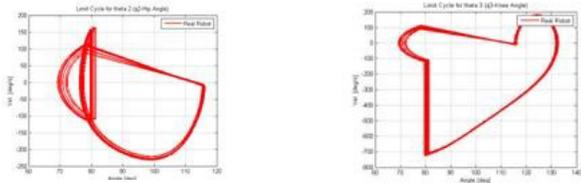
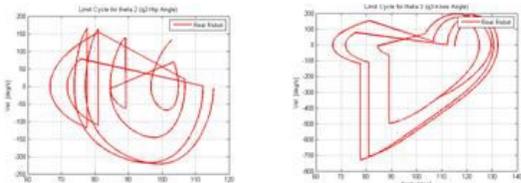


Fig. 2. Various phases of the passive gait.



(a) Limit Limit Cycle (LC) for Hip at 7 degree (b) Limit Cycle (LC) for Knee at 7 degree

Fig. 3. Limit cycles (LC at 7 degree).



(a) LC for Hip at 4 degree (b) LC for Knee at 4 degree

Fig. 4. Limit cycles at 4 degree.

However, the model has several limitations such as reduced degrees of freedom, simplified contact constraint, which are not realistic. That inspired us to improve the model using actual walking data [17], [18].

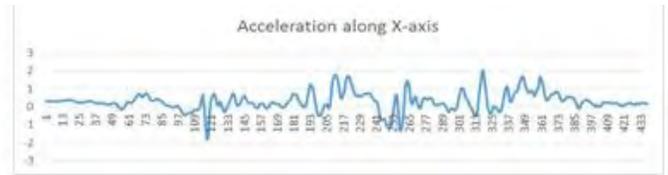
IV. EXPERIMENTAL DATA COLLECTION

We have collected data of 24 students for the purpose of training the connectionist part of our hybrid model. The data is basically again composed of static and dynamic parameters. The static parameters consist of height, weight, thigh length and shank length of each person. The rest of the parameters can be calculated from the above 4 by simple arithmetic and some heuristics [8] regarding the human morphological structure. The dynamic parameters consisted of the hip and knee angles of the subject captured during walking down a slope. The passive gait was captured by using a plank and making it inclined at 4 degree angle and making the subject to walk on it down the inclined. We used an android application named Physics Toolbox Accelerometer developed by Vieyra Software, which will give us the acceleration along x and y axis direction [11], [12]. The smartphone was attached to the subject's hip and knees and the data was captured shown in Fig. 5.

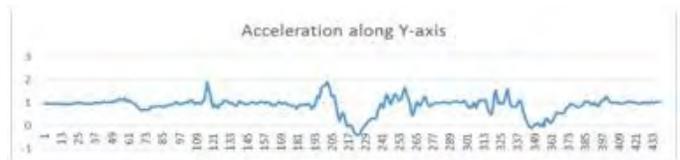


Fig. 5. Physics Toolbox Accelerometer, Vieyra Software used for gait capture.

The output of the accelerometer is the acceleration along x and y axis direction so we transformed it to q_1 and q_2 by inverse kinematic equation. Fig.6 (a), (b) is the output for acceleration in x and y plane respectively.



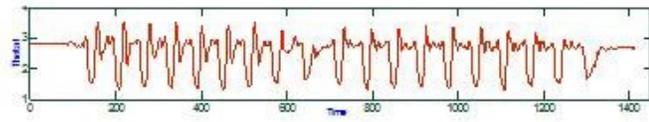
(a) Output of the application in X-plane



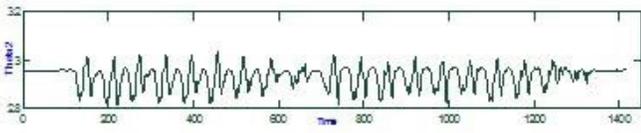
(b) Output of the application in Y-plane

Fig. 6. Output of the application in X-Y plane.

Next, we have applied convolution based filters on the data to smoothen it and have it zero-corrected and also the frequency of the accelerometer of the mobile phone was noted for the necessary sampling which was required for finding out the deviation from the base model shown in Fig. 7. The morphological parameters of the humansubject have been passed on to the analytical model developed in Section II and the output has been stored. This stored data have been compared with the experimental data so collected. To improve upon our analytical base model in the subsequent section, we would like to find an error pattern through detailed error analysis.



(a) Trajectory of the Knee joint after transformation



(b) Trajectory of the hip joint after transformation

Fig. 7. Trajectories after transformation.

V. ERROR ANALYSIS

We have the output of Analytical Base Model (unstable for 4 degree) and the Human data (stable for 4 degree). We are now going to calculate the error in Analytical model with respect to Human data. But before Error analysis, we do the following thing:

Zero correction: We have to take the corresponding deviation of the analytical model's output with the actual gait captured of the human subject. So, we have done the zero correction of the captured data, since we are dealing with time-series data.

Filtering: We have filtered the curve, which we are getting by the Human data using convolution based filters.

Time sampling: Since we are comparing the corresponding data from both the sources, we had to match the frequency of the data. As we could not change the frequency of accelerometer (207 Hz), because it was the inherent property of the smart phone that we were using, so we changed the time step in the Analytical model to be 0.00483s as the time gap between two consecutive data.

Normalization: We made the output of both the data sources to be set between [0, 1] by using the following formula:

$$x_i^{New} = \frac{x_i - x_{min}}{(x_{max} - x_{min})} \quad (2)$$

After this we calculate the errors. We calculate the following 4 types of error and then compared them to select one of them.

RMS (root mean square) Error: It represents the standard deviation of the difference between the data. The RMS serve to combine the magnitude of the errors in predictions for various times into a single measure of predictive power.

$$RMS_{error} = \sqrt{\frac{1}{N} \sum_{i=1}^{i=N} (D_i - A_i)^2} \quad (3)$$

Where D_i = Desired data at a particular time, A_i = Actual data at a particular time and N = number of data.

Mean Percentage Error: It is the mean of the percentage error by which desired value differ by actual value.

$$Meanpercentage_{error} = \frac{1}{N} \sum_{i=1}^{i=N} \frac{mod(D_i - A_i)}{D_i} \quad (4)$$

Absolute Error: It is the average of the absolute difference between the desired data and the actual one.

$$Absolute_{error} = \frac{1}{N} \sum_{i=1}^{i=N} \frac{(D_i - A_i)}{N} \quad (5)$$

Relative Error: It is the mean of the relative difference between the data.

$$Relative_{error} = \frac{1}{N} \sum_{i=1}^{i=N} \frac{(D_i - A_i)}{D_i} \quad (6)$$

After calculating the 4 types of errors we have checked which type of error is more closely related to our dataset by calculating the Spearman's rank correlation coefficient of our data. This correlation gives us the idea about the rank of similarity between the gait captured from the actual human subject and the output of the analytical model. It can be calculated as:

$$\rho = 1 - 6 \frac{\sum_{i=1}^{i=N} d_i^2}{N(N^2 - 1)} \quad (7)$$

Where d_i = Difference between the corresponding datapoints, N = number of data, r = Spearman's rank of correlation. From this, we calculate the correlation of each type of error with the Spearman's rank as:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\left(\sqrt{(n(\sum x^2) - (\sum x)^2)}\right) \left(\sqrt{(n(\sum y^2) - (\sum y)^2)}\right)} \quad (8)$$

The error with the highest correlation with the Spearman's rank is chosen from all the 4 errors, because it has strong affinity to the signal strength of similarity between the data from the two sources. We have then plotted this output in Fig.8. From the figure, it is clearly visible that the mean percentage error has the maximum correlation for both knee and hip. Hence, this error is suitable for training the neural network for finding the error pattern recognition.

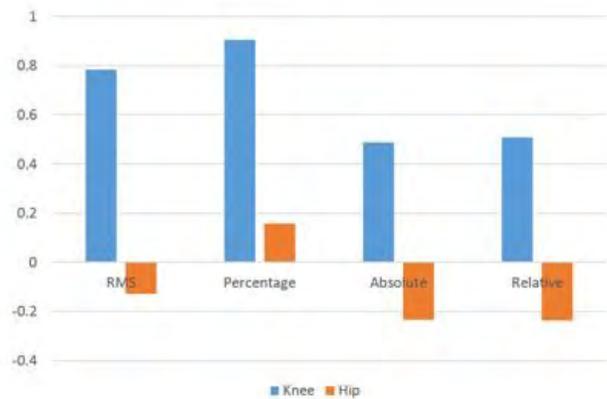


Fig. 8. Correlation between Spearman's rank and the various errors.

VI. THE ARTIFICIAL NEURAL NETWORK APPROACH

With the mean percentage error identified as the suitable output, we are now in a position to create a back-propagation based neural network architecture which can be used to compensate for the instability in the analytical model. This learning ability of our hybrid model can enhance the adaptability of the model and will also extend the operating zone of the entire model [19], [20], [21], [22], [23]. The most well-known methodologies to machine learning are simulated neural systems, also known as hereditary calculations.

In Artificial Neural Network, a Perceptron is a classification of the input of a data into one of a few conceivable outputs i.e., a characterization calculation that makes its expectations focused around a direct indicator capacity consolidating a set of weights with the feature vector. The perceptron algorithm has been used since a long time and has its roots in the 1950s. A multilayer perceptron is a feed forward neural system with one or more concealed layers. Regularly, the system comprises of an Input layer of source neurons, no less than one or more Hidden or shrouded layers of computational neurons, and an Output layer of computational neurons. The data signs are engendered in a forward bearing on a layer-by-layer premise. The implementation of multilayer artificial neural network in the project for the biped locomotive simulation is done via 4 layers: Input layer, Two hidden layers, and Output layer.

A. Use of Neural Network in our Model

In this proposed model, we have used a fully connected neural network with 2 hidden layers, 1 input layer and 1 output layer as shown in Fig.9. We are training the neural net using back-propagation algorithm [1]. The following shows the various input and output parameters and algorithm of the neural network.

Algorithm 1: Algorithm for Multiple Training Encoding Strategy of BAM

Input : thigh length (l_{thigh}), shank length (l_{shank}), swing thigh lower length (l_{sl}), swing shank lower length (l_{ssl}), swing leg mass (msl), swing shank mass (m_{ss}), hip mass(m_h)

Output :qH (i); qK (i); qA(i)

loop: until convergence criteria is mate

Activation

$$Y_j(P) = \text{sigmoid}[\sum_{i=1}^n \hat{a}X_{ij}(P)W_{ij}]$$

$$Y_k(P) = \text{sigmoid}[\sum_{i=1}^m \hat{a}X_{ik}(P)W_{ik}]$$

Weight Training : Output layer

$$j_k(p) = Y_k(p) : [1 - Y_k(p)] : \epsilon_k(p)$$

$$Dw_{jk}(p) = a : Y_j(p) : j_k(p)$$

$$w_{jk}(p + 1) = w_{jk}(p) + Dw_{jk}(p)$$

Hidden Layer

$$j_{ij}(p) = Y_j(p) [1 - Y_j(p)] [\hat{a}X_{jk}(p) w_{jk}(p)]$$

$$Dw_{ij}(p) = a : x_i(p) : j_{ij}(p)$$

$$w_{ij}(p + 1) = w_{ij}(p) + Dw_{ij}(p)$$

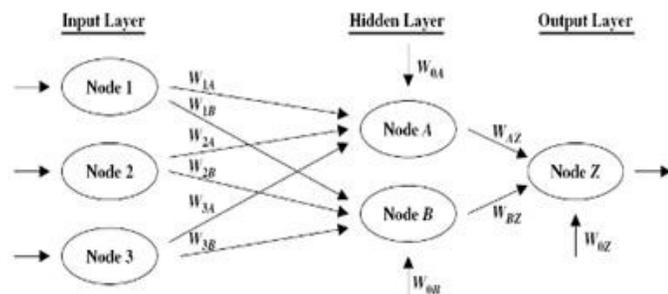


Fig. 9. Architecture of the neural network used in the model.

The neural network has been used to capture the pattern of the

deviation of the joint angles from the analytical base model and that of the real captured data. The error convergence of the neural network is shown in Fig.10.

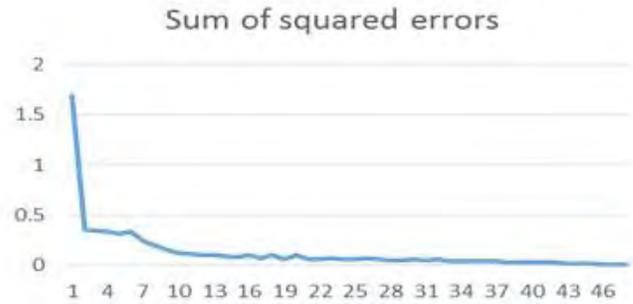


Fig. 10. Convergence of the sum of squared errors.

VII. DEVELOPMENT OF HYBRID MODEL

After finding the error pattern, we have coupled the base model with a neural net based extension so as to generate the passive walker’s stable joint trajectories as accurate as possible in a computationally faster manner as base trajectories and then it is compensated with the error pattern as developed in the earlier section. This integrated system is being called as a hybrid model.

Using the neural network based extension, the error patterns corresponding to the single individual has been obtained (after normalization) as shown in the following chart described in Fig.11.

name	total_l	total_mass	thigh_l	shank_l	thigh_lower_l	shank_lower_l	leg_mass	shank_mass	hip_mass	error_knee	error_hip
vaibhav	1.730	58.000	0.401	0.427	0.228	0.242	9.674	2.755	48.326	0.453	0.356
arvind	1.790	55.000	0.415	0.442	0.235	0.250	9.174	2.613	45.826	0.514	0.332
amk	1.750	63.000	0.406	0.432	0.230	0.245	10.508	2.993	52.492	0.703	0.365
prashant	1.790	67.000	0.415	0.442	0.235	0.250	11.176	3.183	55.824	0.516	0.335
sandeep	1.800	60.000	0.418	0.445	0.237	0.252	10.008	2.850	49.992	2.608	0.324
arjun	1.700	65.000	0.394	0.420	0.224	0.238	10.842	3.088	54.158	0.437	0.358
pratyush	1.680	64.000	0.390	0.415	0.221	0.235	10.675	3.040	53.325	0.519	0.364
manish	1.700	62.000	0.394	0.420	0.224	0.238	10.342	2.945	51.658	0.872	0.365
Bharat	1.650	58.000	0.383	0.408	0.217	0.231	9.674	2.755	48.326	1.931	0.365
yash	1.960	72.000	0.455	0.484	0.258	0.274	12.010	3.420	59.990	0.572	0.241
pratap	1.630	57.000	0.378	0.403	0.214	0.228	9.508	2.708	47.492	0.495	0.366
naveen	1.670	51.000	0.387	0.412	0.220	0.233	8.507	2.423	42.493	0.475	0.336
sunny	1.730	67.000	0.401	0.427	0.228	0.242	11.176	3.183	55.824	0.663	0.339
mohit	1.800	65.000	0.418	0.445	0.237	0.252	10.842	3.088	54.158	0.542	0.356
sidhart	1.980	65.000	0.459	0.489	0.260	0.277	10.842	3.088	54.158	0.536	0.291
pankaj	1.650	48.000	0.383	0.408	0.217	0.231	8.006	2.280	39.994	0.435	0.334
aman	1.630	48.000	0.378	0.403	0.214	0.228	8.006	2.280	39.994	0.455	0.345
subhash	1.900	66.000	0.441	0.469	0.250	0.266	11.009	3.135	54.991	0.470	0.325

Fig. 11. Data used for training the ANN for error pattern matching.

VIII. RESULTS

In this section we have compared the limit cycle behavior between base model alone and the hybrid model (base model together with the neural net) for the passive walking along different angles of inclination as shown in Fig.12 and Fig. 13. We have identified that the base model is unstable below 4 degrees and above 10 degrees of inclination. For comparison, we have used walking along 4 degrees of inclination. As mentioned in the earlier section, the limit cycle behaviour clearly shows that the system is unstable at 4 degrees of inclination. Now as seen from Fig. 12(a) and 13(a), the limit cycle behavior shows that the same system is stable at 4 degrees angle of inclination. This comparison clearly shows while the base model fails to justify stable passive walking by a real subject under certain conditions, the hybrid model fully supports the same. The confusion matrix obtained by the experiment is also given in Fig.14.

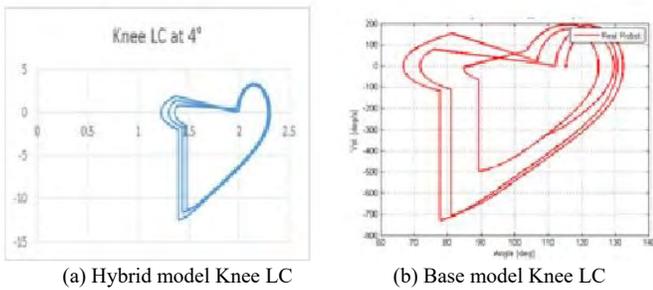


Fig. 12: Comparison of LC

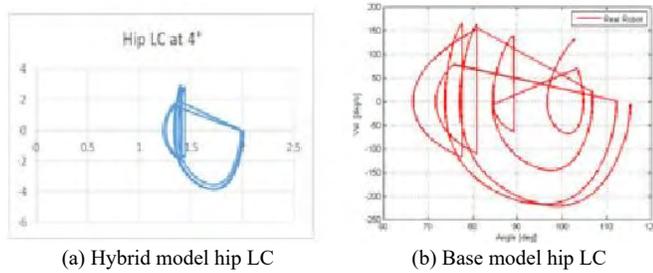


Fig. 13. Comparison of LC.

Training Confusion Matrix					
Output Class	1	2	3	4	
1	79 20.5%	6 1.6%	0 0.0%	4 1.0%	88.9%
2	4 1.0%	63 16.3%	1 0.3%	1 0.3%	81.3%
3	2 0.5%	9 2.3%	87 22.5%	8 2.1%	82.1%
4	11 2.8%	13 3.4%	9 2.3%	89 23.1%	73.0%
	82.3%	86.2%	89.7%	87.3%	82.4%
	17.7%	30.8%	10.3%	12.7%	17.6%
Target Class	1	2	3	4	

Validation Confusion Matrix					
Output Class	1	2	3	4	
1	12 14.5%	2 2.4%	0 0.0%	1 1.2%	80.0%
2	1 1.2%	16 19.3%	0 0.0%	0 0.0%	94.1%
3	2 2.4%	1 1.2%	17 20.5%	3 3.6%	73.3%
4	6 7.2%	3 3.6%	3 3.6%	16 19.3%	42.9%
	57.1%	72.7%	85.0%	80.0%	73.5%
	42.9%	27.3%	15.0%	20.0%	26.5%
Target Class	1	2	3	4	

Test Confusion Matrix					
Output Class	1	2	3	4	
1	15 18.1%	1 1.2%	0 0.0%	0 0.0%	83.4%
2	2 2.4%	15 18.1%	1 1.2%	0 0.0%	83.2%
3	1 1.2%	5 6.0%	20 24.1%	1 1.2%	74.1%
4	3 3.6%	4 4.8%	0 0.0%	15 18.1%	68.2%
	71.4%	80.0%	86.2%	93.8%	78.3%
	28.6%	40.0%	4.8%	6.3%	21.7%
Target Class	1	2	3	4	

All Confusion Matrix					
Output Class	1	2	3	4	
1	106 19.2%	9 1.6%	0 0.0%	5 0.9%	86.3%
2	7 1.3%	94 17.0%	2 0.4%	1 0.2%	93.4%
3	5 0.9%	15 2.7%	124 22.5%	12 2.2%	79.3%
4	20 3.6%	20 3.6%	12 2.2%	120 21.7%	69.8%
	76.8%	88.1%	88.9%	87.0%	80.4%
	23.2%	31.9%	10.1%	13.0%	19.6%
Target Class	1	2	3	4	

Fig. 14. Confusing matrix of Experiment.

IX. CONCLUSIONS

Modelling passive biped walking along a slope analytically is extremely difficult due to high degrees of non-linearity, high dimensionality, under actuation (in swing phase), overactuation (in stance phase), etc. That is why we hardly see any biped humanoid robot capable of exploiting passive walking, which we human being do it naturally through learning. However passive walking capability is extremely needed so as to reduce the energy spent for the robot for walking. In this research we have shown how to exploit the limited capability of analytical model and

extend it through a fusion between the analytical model and the neural network based connectionist model, which may help a biped walker to walk like human exploiting passivity thereby, reducing the energy spent from walking. We have shown that the trajectories obtained as an output of the hybrid model can be effectively used as a reference trajectory for execution level control.

REFERENCES

- [1] G.C.Nandi, Auke Ijspeert, Anirban Nandi, "Biologically Inspired CPG Based Above Knee Active Prosthesis," 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, France, Sept, 22-26, 2008. pp. 2368-2373.
- [2] Semwal, Vijay Bhaskar, et al. "Biologically-inspired push recovery capable bipedal locomotion modeling through hybrid automata," Robotics and Autonomous Systems, vol. 70, pp.181-190, 2015.
- [3] A. Balluchi, et al., "Automotive engine control and hybrid systems: Challenges and opportunities," Proc. IEEE, vol. 7, no. 7, pp. 888-912, July 2000.
- [4] B. Lennartsson, M. Tittus, B. Egardt, and S. Pettersson, "Hybrid systems in process control," Control Syst. Mag., vol. 16, no. 5, pp. 45-56, 1996.
- [5] Lygeros, J., Johansson, K.H., Simic, S.N., Jun Zhang, Sastry, S.S., "Dynamical properties of hybrid automata," IEEE Transactions on Automatic Control, vol.48, no.1, pp.2-17, Jan 2003.
- [6] McGeer, T. "Passive dynamic walking," The International Journal of Robotics Research, vol. 9, no. 2, pp. 62-82.
- [7] Sinnet, R. W, Powell, M. J., Shah, R. P., Ames, A. D. "A human-inspired hybrid control approach to bipedal robotic walking." In 18th IFAC World Congress (pp. 6904-11), 2011.
- [8] Semwal, Vijay Bhaskar, Manish Raj, and G. C. Nandi. "Biometric gait identification based on a multilayer perceptron." Robotics and Autonomous Systems, vol. 65, pp. 65-75, 2015.
- [9] S. J. Hogan, "On the dynamics of a rigid-block motion under harmonic forcing," Proc. Royal Society, A, vol. 425, pp. 441-476, 1989.
- [10] Jeong, Man-Yong, and In-Young Yang. "Characterization on the rocking vibration of rigid blocks under horizontal harmonic excitations," International Journal of Precision Engineering and Manufacturing, vol. 13, no. 2, pp. 229-236, 2012.
- [11] Semwal, V.B., Nandi, G.C., "Toward Developing a Computational Model for Bipedal Push Recovery: A Brief," Sensors Journal, IEEE, vol.15, no.4, pp.2021-2022, 2015.
- [12] Kajita, Shuuji, et al. "Biped walking pattern generation by using preview control of zero-moment point," Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on. Vol. 2. IEEE, 2003.
- [13] Parashar A, Parashar A, Goyal S. "Push Recovery for Humanoid Robot in Dynamic Environment and Classifying the Data Using K-Mean," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 2, pp. 29-34, 2016.
- [14] Raj, Manish, Vijay Bhaskar Semwal and G. C. Nandi. "Bidirectional association of joint angle trajectories for humanoid locomotion: the restricted Boltzmann machine approach," Neural Computing and Applications, pp. 1-9, 2016.
- [15] Semwal, Vijay Bhaskar, and Gora Chand Nandi. "Generation of Joint Trajectories Using Hybrid Automate-Based Model: A Rocking Block-Based Approach," IEEE Sensors Journal, vol. 16, no. 14, pp. 5805-5816, 2016.
- [16] Semwal, Vijay Bhaskar, et al. "Design of Vector Field for Different Subphases of Gait and Regeneration of Gait Pattern." IEEE Transactions on Automation Science and Engineering, vol. PP, no. 99, pp. 1-7, 2016.
- [17] Raj, Manish, Vijay Bhaskar Semwal and G. C. Nandi. "Multiobjective optimized bipedal locomotion," International Journal of Machine Learning and Cybernetics, pp. 1-17, 2017.
- [18] Nandi, Gora Chand, et al. "Modeling bipedal locomotion trajectories using hybrid automata," Region 10 Conference (TENCON), 2016 IEEE. IEEE, 2016.
- [19] Semwal, Vijay Bhaskar, et al. "An optimized feature selection technique based on incremental feature analysis for bio-metric gait data classification," Multimedia Tools and Applications, pp. 1-19, 2016.

- [20] Vijay BhaskarSemwal, Pavan Chakraborty and G.C. Nandi, "Less computationally intensive fuzzy logic (type-1)-based controller for humanoid push recovery," Robotics and Autonomous Systems, vol. 63, Part 1, pp. 122-135, 2015.
- [21] Vijay BhaskarSemwal and G.C. Nandi, "Robust and more accurate feature and classification using deep neural network," Neural Computing and Application, vol. 28, no. 3, pp. 565-574.
- [22] Vijay BhaskarSemwal and G.C. Nandi., "Study of humanoid Push recovery based on experiments," IEEE International Conference on Control, Automation, Robotics and Embedded Systems (CARE), 2013.
- [23] Vijay BhaskarSemwal, Pavan Chakraborty and G.C. Nandi., "Biped model based on human Gait pattern parameters for sagittal plane movement," IEEE International Conference on Control, Automation, Robotics and Embedded Systems (CARE), 2013, pp. 1-5.



Manish Rajt

Graduated in Electronics and Communication Engineering. He obtained his M.Tech. Degree from the Indian Institute of Information Technology, Allahabad in second place in Robotics & AI department. Currently he is on the Ph.D. program at IIIT Allahabad. His research interests are biped locomotion, control systems, nonlinear dynamics, humanoid robotics, chaos and fractal analysis, artificial

intelligence, soft computing, artificial life simulation, mathematical foundations of robotics, and hybrid systems.



Dr. Vijay Bhaskar Semwal

Dr. Vijay Bhaskar Semwal obtained his B.Tech. from the College of Engineering Roorkee, Roorkee, in 2008. He received his M.Tech. from IIIT Allahabad in 2010. Currently, he is serving as Assistant professor in Department of CSE at IIIT Dharwad and before joining here he was serving as faculty at NIT Jamshedpur. He has submitted his Ph.D. at IIIT Allahabad, the final evaluation is yet to

haapen. Before becoming a research scholar at IIIT Allahabad, he worked as a Senior System Engineer (R&D) with Siemens Gurgaon and Bangalore. He has worked for various major organizations, such as Siemens AG and Newgen. Currently he is serving as chair for IEEE student branch of IIIT-Allahabad. For 2013-14 he served as publicity committee mentor for IEEE student branch IIIT Allahabad, session 2015-16 and he has contributed as vice chair for IEEE student branch IIIT Allahabad, session 2014-15. His research interests are machine learning, evolutionary algorithms, analysis of biped locomotion and humanoid push recovery, artificial intelligence, design & analysis of algorithms, biometric identification, Brain wave based authentication.



Prof. G. C. Nandi

Graduated from Indian Institute of Engineering, Science & Technology (Formerly Bengal Engineering College, Shibpur, Calcutta University), in 1984 and post graduated from Jadavpur University, Calcutta in 1986. He obtained his PhD degree from Russian Academy of Sciences, Moscow in 1992. He was awarded National Scholarship by Ministry of Human Resource Development (MHRD),

Govt of India in 1977 and Doctoral Fellowship by External Scholarship Division, MHRD, Govt. of India in 1988. During 1997 he was visiting research scientist at the Chinese University at Hong Kong and he was also visiting Faculty with Institute for Software Research, School of Computer Sciences, Carnegie Mellon University, USA, (2010-2011). Currently, he is serving as the senior most Professor and Dean of Academic Affairs of Indian Institute of Information Technology, Allahabad. From January to July 2014, he served as the Director-in-Charge of Indian Institute of Information Technology, Allahabad. Professor Nandi is the Senior Member of ACM, Senior member of IEEE, Chairman, ACM-IIIT-Allahabad Professional Chapter, (2009-2010), Chartered Member of Institute of Engineers (India), Member of DST (Department of Science and Technology, Govt. of India) Program Advisory Committee member of Robotics, Mechanical and Manufacturing Engineering. He has published more than 100 papers in the various refereed journals and international conferences. His research interest includes robotics specially biped locomotion control & humanoid push recovery, artificial intelligence, soft computing and computer controlled systems.

Users Integrity Constraints in SOLAP Systems. Application in Agroforestry

Abdallah Bensalloua Charef*, Hamdadou Djamila

Computer Science Laboratory of Oran (LIO) - University of Oran 1 Ahmed Benbella, Oran (Algeria)

Received 6 August 2017 | Accepted 16 December 2017 | Published 26 January 2018



ABSTRACT

SpatialData Warehouse and Spatial On-Line Analytical Processing are decision support technologies which offer the spatial and multidimensional analysis of data stored in multidimensional structure. They are aimed also at supporting geographic knowledge discovery to help decision-maker in his job related to make the appropriate decision. However, if we don't consider data quality in the spatial hypercubes and how it is explored, it may provide unreliable results. In this paper, we propose a system for the implementation of user integrity constraints in SOLAP namely "UIC-SOLAP". It corresponds to a methodology for guaranteeing results quality in an analytical process effectuated by different users exploiting several facts tables within the same hypercube. We integrate users Integrity Constraints (IC) by specifying visualization ICs according to their preferences and we define inter-facts ICs in this case. In order to validate our proposition, we propose the multidimensional modeling by UML profile to support constellation schema of a hypercube with several fact tables related to subjects of analysis in forestry management. Then, we propose implementation of some ICs related to users of such a system.

KEYWORDS

Agroforestry, Integrity Constraints, SOLAP, Spatial Data Warehouse, Spatial Decision Support System, Spatial Multidimensional Analysis, UML Multidimensional Modeling.

DOI: 10.9781/ijimai.2018.01.003

I. INTRODUCTION

Spatial On-Line Analytical Processing (SOLAP) and Spatial Data Warehouses (SDWs) are decision support technologies that allow spatial and multidimensional analysis of multisource data. Basing on the multidimensional model, these technologies allow a multidimensional representation of data in hypercubes (datacubes). The hypercubes are essentially based on facts and dimensions concepts. The facts tables are analysis subjects which are described by numerical, textual or spatial measures. The dimensions define the analyzing axis. They can be organized in hierarchies of several levels. This allows the visualization of the measures at different levels of detail.

The data quality in the spatial hypercubes is important since this data is used as a basis for decision-making. In fact, poor data quality could lead to poor decision making. Without considering data quality in spatial hypercubes, it may provide unreliable results. As a result, integrity constraints (ICs) become very important to improve the logical consistency of any database, which increase data quality.

In transactional spatial databases, spatial integrity constraints are defined along the database conceptual models to preserve spatial data quality [1], [2]. However, in analytical process other kind of integrity constraints must be considered for maintaining consistent analysis using the hypercube. It is related to the manner of navigating and querying data with consideration of aggregation rules, hierarchies and other concepts of the multidimensional models.

In the database community, some research works study integrity

constraints for non-spatial datacubes [3], [4]. However, for studying integrity constraints for spatial datacube, its specific characteristic related especially to spatial data features should be considered. Therefore, the objective of this paper is to propose a system for the implementation of user integrity constraints in SOLAP namely "UIC-SOLAP". It takes into account fundamental considerations about users querying and aggregating in spatial OLAP. In addition, other data integrity constraints are implemented in order to obtain efficient analysis within the decision-making process.

Section 2 reviews a state of the art related to quality management in SDW and SOLAP in one hand. In the other hand, some researches in the field of forestry management are cited. In Section 3 our contribution is stated. Section 4 discusses the integrity constraints from spatial databases to spatial datacubes. In Section 5 the proposed system namely "UIC-SOLAP" for user integrity constraints in SOLAP is depicted while the process adopted by it, is presented in Section 6. Section 7 explains the case study related to an implementation of Spatial Data Warehouse for forest management. In Section 8, obtained results are discussed related to the application of the "UIC-SOLAP" system with SOLAP prototype developed for forestry management. Finally, Section 9 concludes and draws more research directions for the implementation of the integrity constraints and data quality in spatial hypercubes.

II. STATE OF THE ART

Quality management in SDW and SOLAP systems is an important research issue. Indeed, precision in SDW was addressed in [5] and [6] which provide logic models and an indexing technique for storing and querying vague spatial data.

* Corresponding author.

E-mail address: charef_bensalloua@yahoo.fr

In [7] authors have studied the problem of completeness in classical DW. They proposed solution taking into account missing values in hierarchies. Inconsistencies which refer to the existence of logical contradictions in SDW can be also controlled by ICs. As mentioned in [8], ICs expression on conceptual models is essential for taking into account all quality rules. For this aim, [9] proposes ad-hoc multidimensional conceptual models allowing the expression of some data ICs using logic predicates. In [10], authors propose an extension of the Entity-Relationship (ER) model for the design of spatio-temporal data warehouses. They define a set of ad hoc pictograms to express ICs on spatial data (topological relations between spatial members). Authors in [11] propose a UML profile for DWs, but they consider only a very small number of data ICs. Aggregation problems are presented in [12] by defining simple schema constraints with UML multiplicities. In [13], the author concludes that several languages for the specification of ICs in spatial databases have been proposed but they are not efficient for defining ICs in spatial hypercubes. Indeed, the author presents a formal language for the specification of integrity constraints in the conceptual model of spatial hypercubes. This language is based on controlled natural languages as well as natural hybrid languages with pictograms. In [14], authors study complex structural aggregation constraints. Basing on a UML(Unified Modeling Language) model, [8] shows that the Spatial OCL (Object Constraints Language) which is an OCL extension for spatial data, allows the definition of a large number of ICs on spatial data. Data modeling languages can express only a very limited number of integrity constraints as for example the cardinalities of the relations between classes of objects. For the expression of IC, Many works use non-standard languages, namely logical, natural, visual or hybrid. Others prefer OCL because it integrates easily with UML. Therefore, it is necessary to use languages dedicated for the ICs specification that are interoperable with the language of definition of the data structures. In [15], the author presents a method for the conceptual modeling of the SOLAP ICs after having introduced the main concepts on which this method is based namely the UML profiles and Spatial OCL.

In the field of natural resources specifically in the forest management and preserving, researches published are relevant for consideration in forest environments as well as in social and economic anxieties. In [16], the author implements a participatory decision making in forest planning. Also, authors in [17], propose a Decision Support for forest management. Authors in [18] treat the problem of forest land use policies. In [19], authors integrate multiple criteria decision analysis in participatory forest planning. Authors in [20] propose a participatory multi-criteria assessment of forest planning policies in conflicting situations. An empirical study on voting power in participatory forest planning is proposed in [21]. For forest protection, authors in [22], propose software for forest fire detection and extinction.

III. CONTRIBUTION

In forestry, we notice the complexity of effective management of natural areas for ecological, economic and social purposes. Indeed, the use of SOLAP application for decision-making allows analyzing this domain in various axes of reflection. However, different kinds of heterogeneity may exist between the data sources (remote sensing, GIS applications ... etc.). As a result, several problems of inconsistency may emerge. In most studies, the discussed problems are related to the heterogeneity of manipulated data. So, few studies have addressed issues related to the end users when decision-making is related to many sub-domains.

We notice also, that, the same territory is usually, under the control of several actors with different anxieties namely territorial

management, exploiting of forest products and even local collectivities with social anxieties. However, the use of SOLAP tool by different actors independently allows them to extract decisional indicators which are heterogeneous. In fact, different user's ICs will generate unsteadiness in the whole decision system.

To the best of our knowledge, no study has treated the formulation of users ICs when using a hypercube with multiple facts tables. Indeed, defining multiple user-defined ICs may generate semantic inconsistencies which affect the analytical quality of such a system.

The main objective of the current study is to propose a system for the implementation of user integrity constraints in SOLAP namely "UIC-SOLAP".

Our proposal integrates all defined Integrity Constraints which are related to the users of SOLAP system along different stages. Indeed, properties of objects in spatial data sources may be adjusted according to every kind of final user. Also, we must verify user queries and aggregations when requesting the SOLAP by the user.

In addition, Integrity Constraints must be defined to respect particular information of a user, so that only concerned data is accessible. These are concerned by the respect of Integrity Constraints and they are all integrated in our proposal namely "UIC-SOLAP".

Secondary objectives are also achieved in this paper. Indeed, we propose a classification of ICs related to the users of SOLAP systems. After that, we propose a multidimensional model for forest spatial data warehouse (SDW). For this aim, we propose a UML profile which allows considering the main properties of spatial and Multi Dimensional (MD) modeling. We integrate also, users ICs definition with formalization of users profiles for specifying visualization ICs according to their preferences. Also, we define inter-facts ICs within the same hypercube. Indeed, unlike the traditional OLAP world, Unified Dimensional Model (UDM), the model for data in Microsoft SQL Server Analysis Services (SSAS), allows to have multiple fact tables within a cube [23].

Then, we propose the exploitation of the UML profile to support "constellation schema" of the forestry Spatial Data Warehouse (SDW). This will help to build a hypercube with several fact tables related to subjects of analysis in forestry management.

Finally, we test our proposal using data concerning forest areas of Mostaganem in Algeria to apply our methodology for verifying the user's integrity constraints in such a system.

IV. INTEGRITY CONSTRAINTS FROM SPATIAL DATABASES TO SPATIAL DATACUBES

Spatial Data Warehouses (SDWs) integrate, organize in a multidimensional way and store very large volumes of spatial and non-spatial data from multiple sources to support the decision-making process within an organization [24].

Spatial OLAP (SOLAP) systems are a category of software tools that allow interactive exploration based on a spatio-multidimensional approach at several levels of detail of the SDW [25]. They extend and enrich the OLAP systems by new concepts and operators which allow effectuating spatial analyzes. Consequently, they allow the visualization of the results of analytical queries in the form of tables, charts and maps.

SOLAP systems are based on data stored in SDWs. As a result, SOLAP analysis quality includes spatial data quality which is the characteristic that distinguishes spatial datacubes from traditional ones.

Here, we must take into account the internal quality defined by its geometric and semantic accuracy; and the external quality which is related to the user needs or quality in the context of use.

In the Spatial Databases, [26] classifies spatial ICs into three categories:

1. Topological ICs which include all geometric properties and relationships on data;
2. Semantic ICs which concern the meaning of geographical features and
3. User-defined ICs which are business rules in spatial databases.

However, in [13], the author notes some limitations when trying to apply these ICs to real applications.

In the Temporal Databases, a formal classification for temporal ICs is introduced in[27]. Indeed, ICs are categorized according to the transaction time (when a fact is stored in the database), valid time (when a fact was true in the reality), and both.

In the Spatiotemporal Databases static and transition ICs is integrated in[26] with the categories of spatial ICs. Here, other temporal ICs are involved. They restrict the possible lifecycle of objects like the IC that specifies the period of applicability of a rule.

The quality of SOLAP decision-making depends on the quality of the data stored and on the way in which the data is explored. Integrity constraints (ICs) are defined in [28] as assertions, typically defined in the conceptual model of an application, that are aimed at preventing the appearance of incorrect data in a database. In SOLAP system context, IC can be used for:

1. identifying erroneous data stored,
2. identifying incorrect analytical queries and
3. defining correct aggregation rules.

In analytical processing, a multidimensional model is based on concepts like dimension, hierarchy, level, member, fact, and measure. As a result, an integrity constraint specification must hold specific semantics supporting these concepts. Hence, it is more efficient to express spatial multidimensional integrity constraints referring to multidimensional elements. At the conceptual level, SDW ICs define conditions that data, metadata or analytical expressions are expected to satisfy. In [8], authors see that the definition of ICs at a conceptual level makes it possible to consider quality problems in the early stages of development. Indeed, they propose two new classifications of SDW ICs: (1) a classification oriented SDW concepts (Fact, Dimension, Aggregation, etc.) and (2) an implementation-oriented classification. The first category categorizes ICs according to the multidimensional elements that these ICs imply. The second group categorizes ICs following the implementation levels in the SOLAP architecture. Indeed, the conceptual expression of any IC depends on the concerned SDW concepts; when its physical translation depends on the SOLAP tier where it is implemented. In fact, he reveals that ICs can be specified using OCL and Spatial OCL at the conceptual level. However, Spatial OCL and UML do not allow specifying all IC proposed in that classification. This is related especially to metadata IC, the aggregation ICs specific to an application field and also visualization ICs. The author in [15], proposes an extension of the precedent classification by introducing a new class related to “Query ICs”. This aims to avoid misinterpretations of the results as shown in Fig. 1.

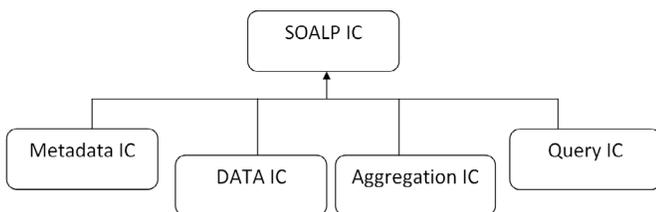


Fig. 1. Classification of SOLAP ICs.

However, the solution can be by adding in the UML profile:

1. more tagged values representing the main metadata,
2. more stereotypes allowing users to define their own aggregation ICs and also,
3. formalization of user profile for specifying visualization ICs according to their preferences.

A. New facts ICs in Constellation Schema

According to [13], a hyper-cell consists of a pair (L, MS) where L is a finite set of dimension levels and MS is a finite set of measures of a hypercube schema. The author classifies fact ICs into three categories. Indeed, a fact can be defined for one fact of a hyper-cell (f-Inter0), for several facts of the same hyper-cell (f-Inter1), and for several facts of several hyper-cells (f-Inter2) as shown in Fig. 2.

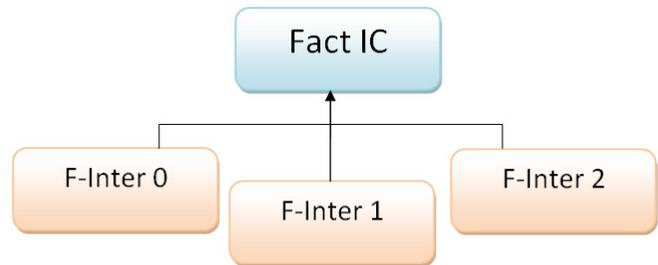


Fig. 2. Categories of fact ICs.

In this classification, the author considers ICs facts of one or several hyper-cells corresponding to the same facts table.

In our proposal, we consider new ICs for measures in several facts tables within the same hypercube as shown in Fig. 3. This situation can occur when considering a constellation schema of a hypercube with several fact tables. In our case study, these fact tables are related to subjects of analysis in forestry management.

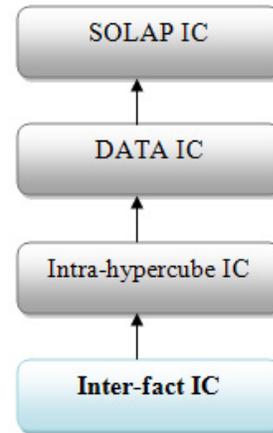


Fig. 3. Inter-fact ICs within a hypercube.

B. Users ICs in SOLAP

ICs related to the users of SOLAP system are numerous. In our proposal, when we use a hypercube with a constellation schema with several users, we classify these ICs into four categories as shown in Fig. 4. This classification may not to be exhaustive but we have focused on these parts that we consider directly linked to the user of such a system.

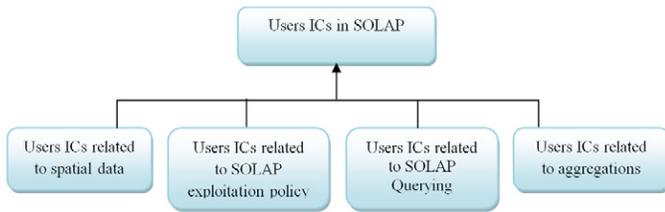


Fig. 4. Users ICs in SOLAP.

1) Users ICs related to Spatial Data

Integrity constraints apply to the database, a set of states that are valid by virtue of properties of objects that need to be stored according to every kind of final user of SOLAP system.

Geometric properties relationships of spatial data are related to «topological», «metric» and «ordering» relationships [30]. Here, semantic ICs concern the meaning of geographical features. The user can define integrity constraints which allow database consistency to be maintained according to his need [26].

After definition of the integrity constraints, we need to check if they are consistent. Indeed, before checking if a database satisfies a set of spatial semantic constraints, we need to make sure that the constraints themselves must not to be in conflict [31].

For example, a user rule must be activated to locate *oil station* at a given distance from the *forest*. That must not be in contradictory with another user-defined IC.

2) Users ICs related to SOLAP Exploitation Policy

At the modeling stage, the UML profile may be enriched with a stereotype called “UserProfile”, which allows specifying constraints depending on particular information of a user or a group of users [32]. As a result, we customize information depending on the characteristics of the user that is requesting that information.

The Fig. 5. illustrates user ICs for an example of secure multidimensional modeling adapted from [32]. It shows a part of the Secure Model, named ‘Hospital’ which is based on a typical health-care system. In this example authors notice the following considerations related to the user of the system:

- Patients may access their own information as patients.
- deny access to data of patients who have been treated before the date of initial contract.
- deny access to admission information to users whose working area is different than the area of a particular admission instance.

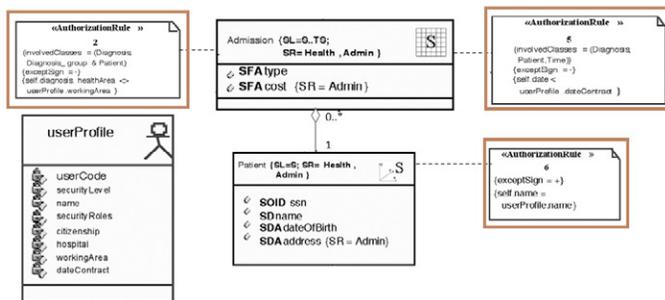


Fig. 5. Example of user ICs adapted from [32].

3) Users ICs related to SOLAP Querying

A SOLAP query is a combination of measures and members of different dimensions. A querying IC verifies the validity of this combination that may be effectuated by any user.

For example, it does not make sense to request geographic distribution of silvicultural effectuated by “*Mostaganem*” forest administration “*conservation*”, in 1985 in “*Zemoura*” forest.

Here, according to the multidimensional model represented with a constellation schema proposed in Fig. 10, we have an analytical query which combines:

- Fact table: [*“silvicultural”*].
- SpatialMeasure: [*“localisation”*].
- Spatial aggregation level: [*forest= “Zemoura”*].
- Spatialaggregation level: [*conservation= “Mostaganem”*].
- Timedimension: [*year=“1985”*].

The result will demonstrate that there is no silvicultural activities; while, according to the administrative division after February, 1984, “*Zemoura*” forest became under the direction of another conservation namely “*Relizane*” (western of Algeria).This can produce unreliable decision. So a query integrity constraint is able to prohibit this combination of parameters.

4) Users ICs related to Aggregations

The author in [15] has implemented aggregation ICs in the UML profile to avoid the implementation of semantically and structurally incorrect models of SDW.

For example, in order to force the user not to aggregate non-additive measures using the aggregate function “sum”, the author defines the OCL statement which forbids this as shown in Fig. 6.

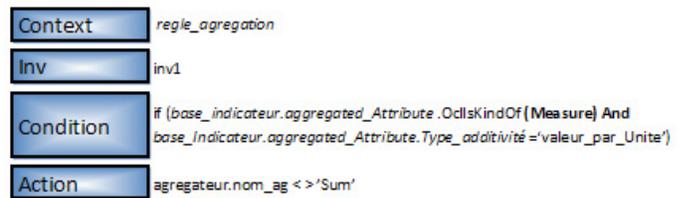


Fig. 6. Example of IC related to aggregations.

Generally, this kind of aggregation rules is valuable for every user. However, any user need to use a limit number of measures related to a specific facts table with group of dimension levels. Hence, we can attach the aggregation rules to the “*UserProfile*” stereotype defined in Case Study section.

V. THE PROPOSED SYSTEM FOR USER INTEGRITY CONSTRAINTS IN SOLAP “UIC-SOLAP”

In the current study we propose a system namely “*UIC-SOLAP*” for checking integrity constraints which regulate the interactions between the user and the SOLAP process as shown in Fig. 7.

This system integrates all defined Integrity Constraints which are related to the users of SOLAP system along different stages of the latter. Indeed, in data sources which can be spatially referenced data, properties of objects may need to be stored according to every kind of final user.

In the interaction with SOLAP Server, the user uses combination of measures and members in analytical queries and he uses aggregation rules. These are concerned by the respect of Integrity Constraints.

Also, when using the user-interface of a SOLAP application, Integrity Constraints must be defined to respect particular information of a user, so that only concerned data is accessible.

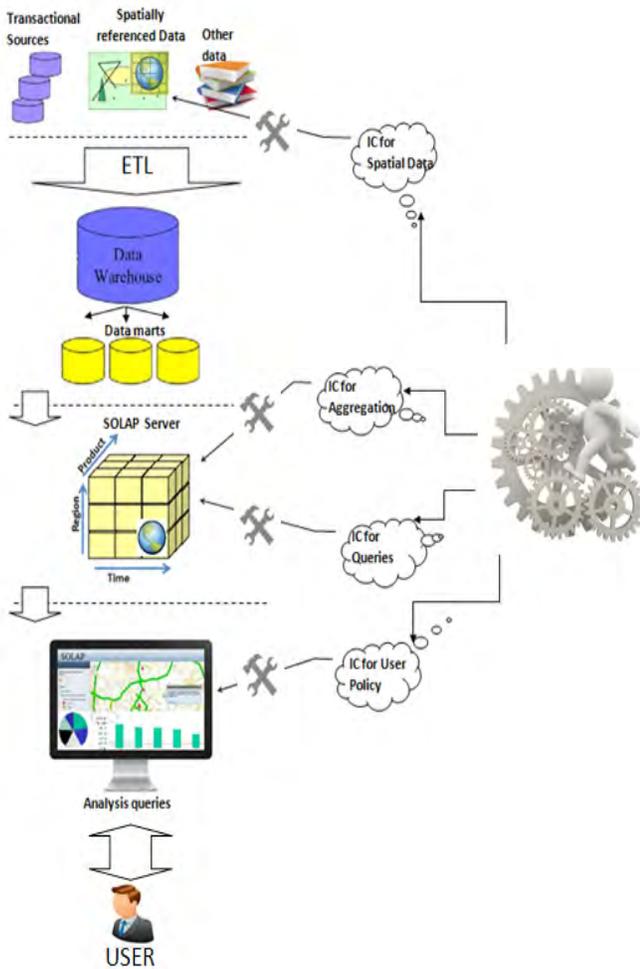


Fig. 7. “UIC-SOLAP” User Integrity Constraints in SOLAP.

VI. THE PROCESS ADOPTED BY “UIC-SOLAP”

Throughout the SOLAP building and using, the process adopted by the proposed system namely “UIC-SOLAP”, can be described as following:

- Firstly, the users can make different assertions in the spatial database according to their needs. Along this task, we must take care that these assertions are not in conflict between them.
- Secondly, to identify incorrect analytical queries, we must verify the validity of the combination of measures and members.
- Then, for defining correct aggregation rules, we must define in the modeling stage, stereotypes allowing users to define their own aggregation ICs.
- And finally, for the respect of user policy, formalization of user_profile concept is assured allowing specifying constraints depending on particular information of a user.

The operating algorithm for applying UIC_SOLAP can be as follows:

Algorithm

```

// check data user_integrity_constraints when analyzing
data, before applying “ETL” process
S ← ∅; // set of spatial data properties
P → ∅; // final list of properties initially empty
For (every kind of final_user of SOLAP)
  Begin
    S ← Read (database_properties); //user will enter
    //spatial properties according to him
    Verify_contradictory (P,S); //verify if entered properties
    are contradictory with the others
    While (not ok) // While S is in contradictory with at
    least one of database_properties
      Begin
        redefine_properties(S) // S will be redefined
        Verify_contradictory (P,S);
      End while
    P ← P+S; // add S to final list of properties
  end_for;
// (checking exploitation policy user integrity constraints)
When connecting to SOLAP
  Read (user_id)
  Apply_profile (user_id) //applying a set of parameters
  allowing user to see only information related to his
  speciality
  // ... When querying the SOLAP
  While (user is connected)
    Begin
      //checking querying user integrity constraints
      Read(query)
      Verify_combin ({dimension_members}, measure)
      If (not OK) then show_message(“inconsistent query will
      generate inconsistent results”);
      //checking querying user integrity constraints
      Read (aggregation)
      Verify_aggregation(query, aggregation_model)
      //to Verify if requested aggregation function is
      allowed to be applied
      If (not OK) then show_message(“not allowed aggregation
      will generate inconsistent results”)
    End while
  End Algorithm.

```

For the modeling, in our approach we use Unified Modeling Language (UML) since it becomes a standard using object-oriented concept and is widely used for designing various systems and software. UML can be also easily extended with profile to adapt it to specific domain namely multidimensional modeling.

The activity diagram of the proposed system namely “UIC-SOLAP” is shown in Fig. 8.

VII. CASE STUDY

The case study used to illustrate our proposal concerns the multidimensional (MD) analysis in a Spatial Decision Support System (SDSS) for forest management.

We use an extension of UML for MD modeling with three mechanisms: stereotypes, tagged values and constraints in the UML profile [33]. We use also Object Constraint language (OCL) to formalize IC constraints which refine the definitions of stereotypes and tagged values [34].

The forest SDW metamodel we propose is adapted from [35]. It

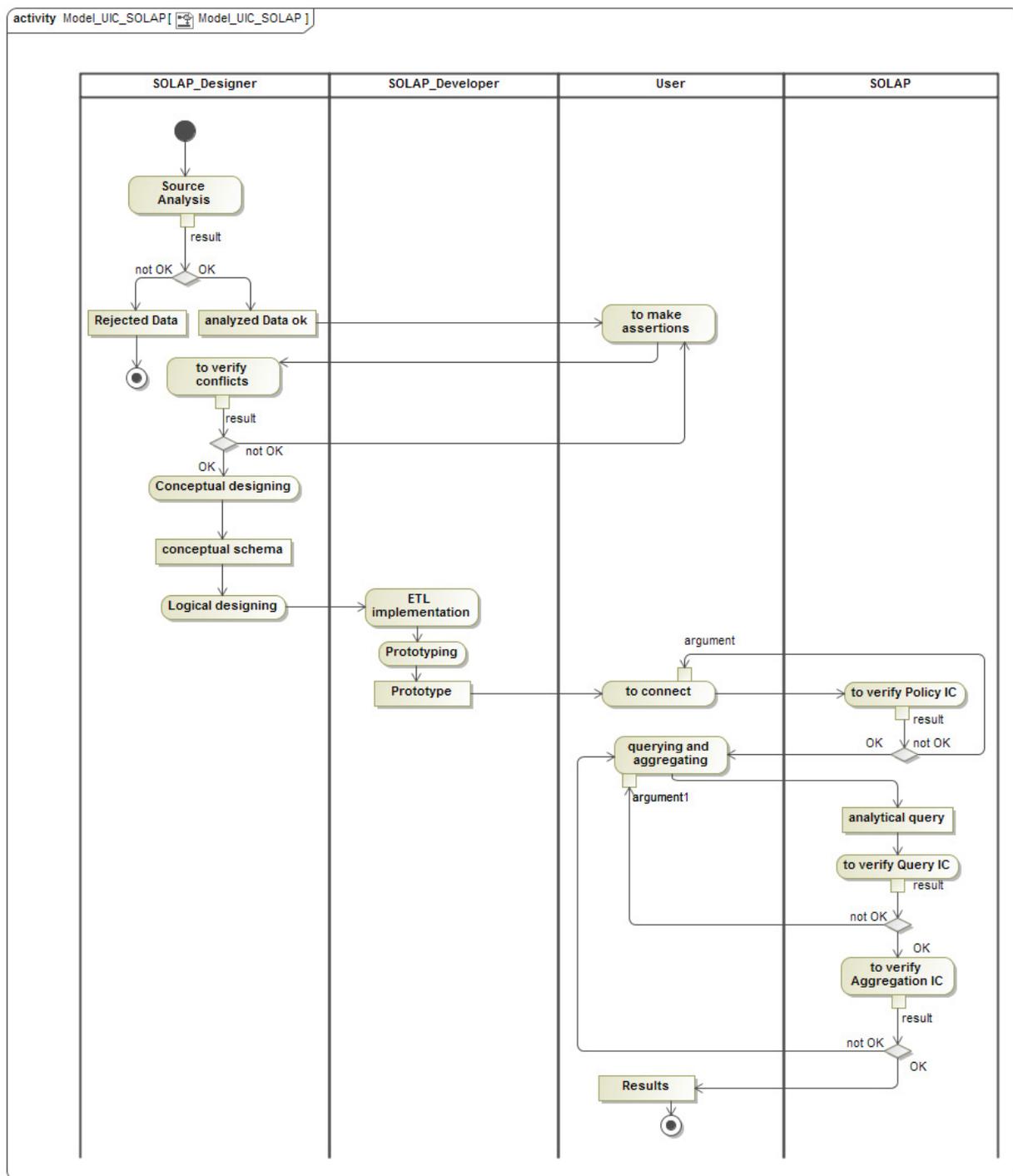


Fig. 8. Activity diagram for “UIC-SOLAP”.

allows representing the main static concepts of the SDW. We used it thanks to its completeness and clarity. However, this metamodel does not allow creating a hypercube with several fact tables. So we introduce some modifications as depicted in Fig. 9. To do this we propose one-to-many cardinality in the composition relationship between hypercube package stereotype and the one of fact class. This allows having multiple fact tables within a hypercube [23]. Some stereotypes were removed and others were added where necessary.

We must ensure the modeling of all the measures, dimensions and hierarchies used for data aggregations required in the analysis process.

The class diagram for the forestry SDW is shown in Fig. 10. It is considered as spatial because it has at least one spatial measure and at

least one spatial dimension.

In our case study, users are interested in the analysis of three aspects of the forestry domain, defined using a hypercube with three facts tables sharing some dimension tables in a constellation schema, which are:

1. «Silvicultural» facts table related to forestry activities such as afforestation, spacing, growth and yield, etc. It is described using three measures where one is spatial (localisation) and two are numerical (volume_fin and superficie);
2. «Event» facts table related to fires, illegal cut, etc. It is described using three measures where one is spatial (localisation) and two numerical which are volume_fin for the financial cost and superficie, the area concerned;

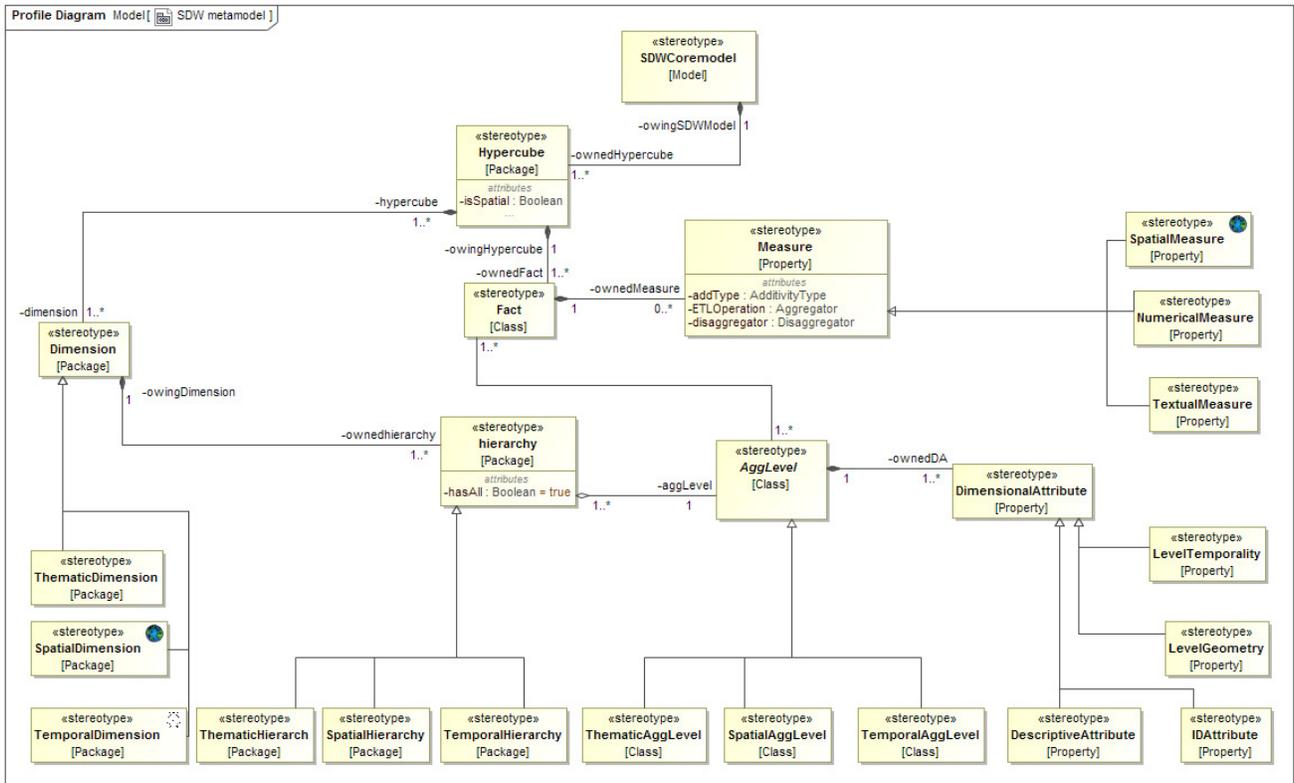


Fig. 9. Forest SDW metamodel.

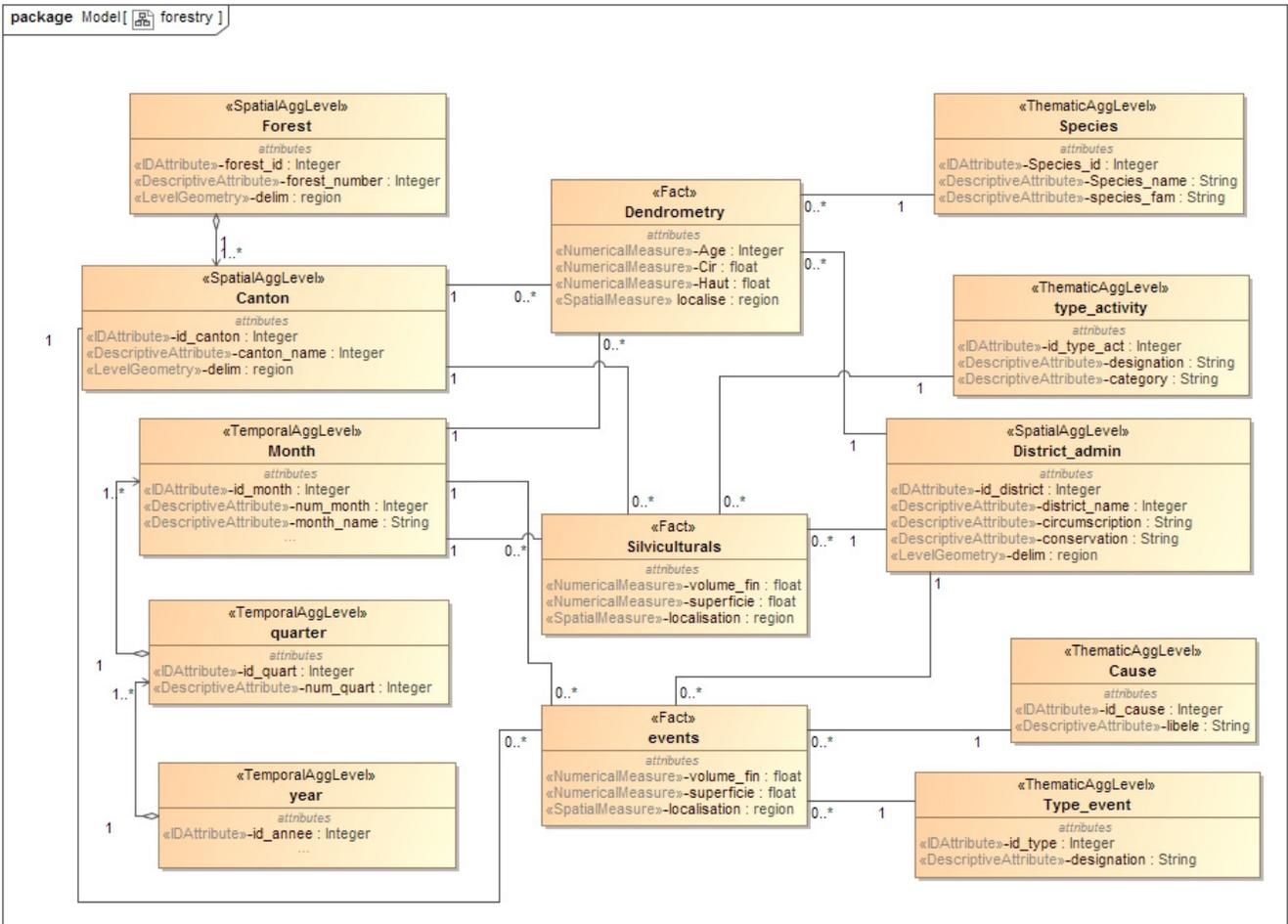


Fig. 10. Constellation schema of forest SDW.

3. «Dendrometry» facts table related to the measurements of various characteristics of trees, such as their circumference, *high*, *age*, etc. It is described using three numerical measures (*cir*, *haut*, and *age*), and one spatial measure (*localise*).

Dimensions can be:

1. Temporal, to effectuate business analysis at three levels: month, quarter, or year;
2. Spatial, to effectuate business analysis at geo-referenced areas. We have two spatial dimensions: administrative with levels (district, circumscription and conservation); and ecologic with levels (canton and forest);
3. Thematic, which is related to the business processes. We have four thematic dimensions: *Species*, *Type_activity*, *Cause* and *Type_event*.

We define the *UserProfile* stereotype for specifying IC depending on particular information of the user. In the class diagram, we have just one class of this stereotype; that has no association to other classes. It is used only for formulating IC as shown in Fig. 11.



Fig. 11. UML Stereotype and Class for UserProfile.

For the new type of inter-facts IC within one hypercube cited in section 4.1, in the MD schema of forest SDW we introduce the assertion that verifies prospective conditions between two or more facts tables.

Fig. 12 shows an inter-facts IC between two facts tables *Silviculturals* and *Events* of the same hypercube. It specifies that superficies of reforestation must be greater than superficies destroyed by fires. We note that specification languages of ICs are explained in [13]. However, they go beyond the objective of the current study.



Fig. 12. Inter-facts IC.

VIII. DISCUSSION OF RESULTS

We have developed a SOLAP application with JAVA Eclipse, the object-oriented programming language. With the integration of ArcGIS runtime SDK, the solution provides cartographic synchronization with tabular and diagram displays as shown in Fig. 13. Different users may pass their analytical queries using buttons and combo boxes available in the interface. They can also formulate their queries directly in MultiDimensional eXpression (MDX) language to communicate with the “Silvi” hypercube. The prototype is tested with a dataset concerning the Mediterranean forests of Mostaganem department situated in north western of Algeria.

Different users with different aspects of interest can manipulate the system with different manners. This can generate inconsistent analytical results. Hence, the formulation of ICs at different stages can avoid integrity problems.

In the following examples, we will focus on some ICs related to users of SOLAP application.

- Firstly, for the Users ICs related to spatial data, we take into consideration “user-defined” ICs. They are business rules which control the integrity of data stored and results. Here we have three examples related to users of each one of facts tables of the “Silvi” hypercube:

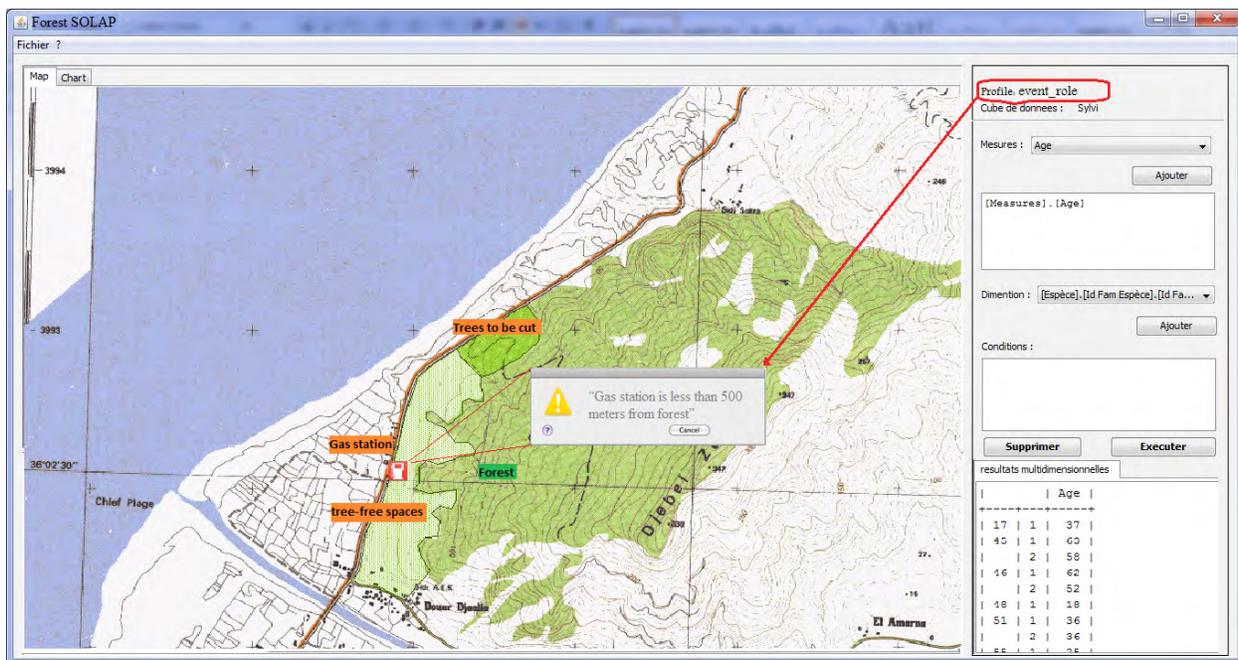


Fig. 13. Data User ICs in Forest SOLAP.

1. "The distance between a forest and a gas station shall be more than 500 meters".
2. "Afforestation can be effectuated at forest spaces tree-free".
3. "it is authorized to cut down a tree when its circumference >= 4m; height >= 20m; age >= 50 and it is near a road".

These ICs may be contradictory since the users are heterogeneous and data will be used in deferent facts tables within the same hypercube.

In our proposal we try to remove these contradictory issues. Since our prototype may be used by any kind of users, they may query any one of facts tables. Their interests and rules may be different; hence, one user may not accept afforestation near a gas station even it is forest space tree-free; It can be dangerous for other spaces.

- For the Users ICs related to exploitation Policy, by the use of "UserProfile" concept, a "userrule" will be activated only if the concerned user is logging on. We create three profiles: *Event_role*, *Silvicultural_role* and *dendrometry_role* related to the users of Events, *Silviculturals* and *dendrometry* facts tables consecutive.

In this case, if the role is "Event_role" then only the related IC message appears "Gas station is less than 500meters from the forest"; despite "existing of very old trees near the road to be cut" and "existing of tree-free spaces within the forest to be reforested".

We note that the decisional problem can be solved by the application of collaborative group decision support system. However, it goes beyond the objective of the current study which is focused on user ICs in SOLAP application.

- For the Users ICs related to querying, considering the new administrative division, the system we propose verify the correctness of measures and dimension members combination in user queries. For the example above concerning combination of measures of *Silviculturals* facts table and ecological spatial dimension "forest=Zemoura" with administrative spatial dimension "conservation=Mostaganem" is false when time dimension member is after February 1984. Hence, the message in Fig. 14 appears.

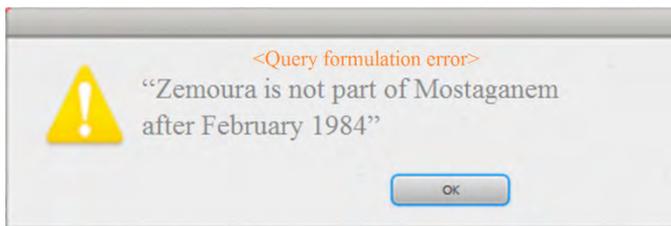


Fig. 14. Notification for non respected query IC.

- And finally, for the Users IC related to aggregation, it defines how user wants to aggregate measures for calculating different indicators. Here, we indicate that every user has a set of indicators for the multidimensional analysis related to his specialty. So we propose that in the aggregation model to be defined in the UML profile, for each indicator we must specify which function to be used for aggregating the measure. ICs are defined in the metamodel and can be controlled when validating the conceptual model.

In this case, since the IC is related to the indicator and the latter is associated to the user, hence we can associate an IC including the 'UserProfile specialty' to the facts table's name. Indeed, this association can enrich the analysis consistency, since every user is able to effectuate correct queries relating to his specialty.

We note that querying a data warehouse is the job of the analyst or decision maker. However, in the current study we focus on the case

when we use a hypercube with multiple facts tables. For example, in group decision systems, they may exist different actors (users). Each of them queries facts of his interest, and has no idea how to use other measures. This seems to be facultative, however it can be considered as a short cut and earlier control for measures aggregation. For example, concerning the *dendrometry* facts table, only users for whom UserProfile.specialty = "dendrometry", will be able to aggregating its measures.

IX. CONCLUSION AND FUTURE WORK

Spatial Data Warehouses (SDW) and SOLAP systems are used for discovering business information for Spatial Decision Support System (SDSS). In forestry management, the decision-making process requires the use of multidimensional model including multiple facts tables within the constellation schema.

As stated in studies presented in this paper, SOLAP analysis goodness depends on three quality aspects which concern data, aggregation and queries formulated. For addressing these issues, Integrity Constraints (ICs) have been acknowledged to be an excellent approach. Indeed, they define conditions that must be satisfied to improve the accuracy, consistency, and completeness of databases. ICs are often defined in conceptual models to allow handling these issues at the early stages of development.

The main objective of the current study is to highlight the ICs essentially related the users of SOLAP solution for the effective management of forestry areas and its activities. Indeed, we propose a UML modeling of the forest SDW profile. Next, we use it for the constellation schema modeling with three facts tables sharing spatial, temporal and thematic dimensions. Then, we propose a classification of ICs related the users of SOLAP systems. We propose also a new type of inter-facts ICs related to the use of hypercube with multiple facts tables. Finally, results of our proposal are shown using data concerning forest areas of Mostaganem in Algeria.

In our future works, we plan to extend our results by analyzing other interesting cases in which the SOLAP users ICs can be posed. This can be achieved by enriching our classification presented in this paper. We will address also, other aspects related to the quality in spatial data warehouses and verification of logical consistency in SOLAP Systems.

ACKNOWLEDGMENT

A part of this study was carried out in collaboration with experts from the Forestry Administration of Mostaganem (Algeria). Thereby we would like to acknowledge them with gratitude, in particular Mrs. "Touafek Narimann", the forest Conservator who left us forever and will remain in our hearts and our prayers as long as we live.

REFERENCES

- [1] M.A.Mostafavi, G. Edwards and R. Jeansoulin, "An Ontology-based Method for Quality Assessment of Spatial Data Bases", Proceedings of the 4th International Symposium on Spatial Data Quality, Vienna, Austria. 2004.
- [2] S. Vallières, J. Brodeur and D. Pilon, "Spatial Integrity Constraints: A Tool for Improving the Internal Quality of Spatial Data. Fundamentals of Spatial Data Quality, R. Devillers and R. Jeansoulin (Eds.), ISTE, pp. 161-177. 2006.
- [3] C. A. Hurtado, C. Gutierrez and A. O. Mendelzon, "Capturing summarizability with integrity constraints in OLAP", ACM Transactions on Database Systems, 30(3), pp. 854-886, 2005.
- [4] F. Ghazzi, F. Ravat, O. Teste and G. Zurfluh, "Contraintes pour modèle et langage multidimensionnels". Ingénierie des Systèmes d'Information 9(1), pp. 9-34, 2004.

- [5] T. L. L. Siqueira, R. C. Mateus, R. R. Ciferri, V. C. Times, and C. D. A. Ciferri, "Querying vague spatial information in geographic data warehouses advancing geoinformation science for a changing world" Volume 1 of Lecture Notes in Geoinformation and Cartography, pp. 379–397. Springer, 2011.
- [6] D. Perez, M. J. Somodevilla, and I. H. Pineda, "Fuzzy spatial data warehouse : A multidimensional model", Decision Support Systems Advances, pp. 3–9. IEEE Computer Society, 2007.
- [7] C. E. Dyreson, T. B. Pedersen, and C. S. Jensen, "Incomplete information in multidimensional databases", Multidimensional databases, pp. 282–309, IGI Publishing, 2003.
- [8] K. Boulil, S. Bimonte and F. Pinet, "Un modèle UML et des contraintes OCL pour les entrepôts de données spatiales. De la représentation conceptuelle à l'implémentation", Ingénierie des Systèmes d'Information, 16(6): 11-39, 2011.
- [9] F. Ghozzi, F. Ravat, O. Teste, and G. Zurfluh, "Modèle multidimensionnel à contraintes", Revue d'Intelligence Artificielle 17(1-3), 43–55, 2003.
- [10] E. Malinowski and E. Zimanyi, "Advanced data warehouse design. Advanced Data Warehouse Design: From Conventional to Spatial and Temporal Applications", Data-Centric Systems and Applications, Volume. ISBN 978-3-540-74404-7. Springer, 2008.
- [11] O. Glorio, and J. Trujillo, "An MDA Approach for the Development of Spatial Data Warehouses", In I.-Y. Song, J. Eder, et T. Nguyen (Eds.), Data Warehousing and Knowledge Discovery, Volume 5182 of Lecture Notes in Computer Science, pp. 23–32. Springer, 2008.
- [12] J.-N. Mazon, J. Lechtenborger, and J. Trujillo, "A survey on summarizability issues in multidimensional modeling", Data & Knowledge Engineering, 68(12), 1452–1469. 2009.
- [13] M. Salehi, "Developing a model and a language to Identify and specify the integrity Constraints in spatial datacubes", PhD Thesis presented at the Faculty of high studies of LAVAL University, Quebec, 2009.
- [14] F. Pinet and M. Schneider, "A unified object constraint model for designing and implementing multidimensional systems". In Journal on Data Semantics XIII, Volume 5530 of Lecture Notes in Computer Science, pp. 37–71. Springer, 2009.
- [15] K. Boulil, "Une Approche Automatisée basée sur des Contraintes d'Intégrité définies en UML et OCL pour la Vérification de la Cohérence Logique dans les Systèmes SOLAP - Applications dans le domaine agri-environnemental", PHD thesis, University Blaise Pascal Clermont-Ferrand II, France, 2012.
- [16] J. Ananda, "Implementing participatory decision making in forest planning", Environment Managing, 39, 534–544, 2007.
- [17] A. Kangas, M. Kurttila, J. Kangas, T. Hujala and K. Eyvindson, "Decision Support for Forest Management", Springer: New York, NY, USA, 2015.
- [18] C. Prell, K. Hubacek and M. Reed, "Stakeholder analysis and social network analysis in natural resource management", Internet Society, 22(6), 501–518, 2009.
- [19] E. Nordström, L. O. Eriksson and K. Öhman, "Integrating multiple criteria decision analysis in participatory forest planning: Experience from a case study in northern Sweden. Forest Policy Econ. 2010, 12, 562–574.
- [20] M. Acosta and S. Corral, "Participatory Multi-Criteria Assessment of Forest Planning Policies in Conflicting Situations: The Case of Tenerife", Forests, 6, 3946–3969, 2015.
- [21] N. Vainikainen, A. Kangas and J. Kangas, "Empirical study on voting power in participatory forest planning". Journal of Environmental Management, 88, 173–180, 2008.
- [22] R. G. Crespo, V. H. M. García, J. J. R. Escibano, E. T. Franco and J. M. C. Lovelle, "Forest fire detection and extinction software", In Proceedings of the 2009 International Conference on Artificial Intelligence, ICAI 2009 (Vol. 2, pp. 885-890).
- [23] H. Sivakumar, C. Matt, M. Sethu, Z. Robert and G. Y. L. Denny, "Professional Microsoft SQL Server Analysis Services 2008 with MDX", Indianapolis, Indiana, USA. Wiley Publishing Inc Press, 2009.
- [24] N. Stefanovic, J. Han, and K. Koperski, "Object-based selective materialization for efficient implementation of spatial data cubes", IEEE Transactions on Knowledge and Data Engineering, 12(6), 938–958, 2000.
- [25] Y. Bédard, S. Rivest and M. J. Proulx, "Spatial On-Line Analytical Processing (SOLAP): Concepts, Architectures and Solutions from a Geomatics Engineering Perspective. Data Warehouses and OLAP: Concepts, Architectures and Solutions", R. Wrembel and C. Koncilia (Eds.), Idea Group Publishing, p. 298-319, 2007.
- [26] S. Cockcroft, "A Taxonomy of Spatial Data Integrity Constraints", Geoinformatica, 1(4), pp. 327-343, 1997.
- [27] M. H. Bohlen, "Valid Time Integrity Constraints", Technical Report 94-30, Department of Computer Science, University of Arizona, Arizona, USA, 22 pages, 1994.
- [28] K. Boulil, S. Bimonte and F. Pinet, "Un cadre conceptuel basé sur UML et Spatial OCL pour la définition des contraintes d'intégrité dans les systèmes SOLAP". Actes des 8èmes journées francophones sur les Entrepôts de Données et l'Analyse en ligne (EDA 2012). Bordeaux, France, 2012.
- [29] M. Duboisset, "Un système de contraintes d'intégrité OCL pour les bases de données Spatiales – Application à un Système d'Information pour l'Épandage Agricole", PHD thesis, école doctorale pour l'Ingénieur (SPI), Clermont-Ferrand, France, 2007.
- [30] M. J. Egenhofer, "A Formal Definition of Binary Topological Relationships", Proceedings of the 3rd International Conference on Foundation of Data Organization and Algorithms, LNCS 716, Springer-Verlag, Paris, France, p. 457-472, 1989.
- [31] L. Bravo, and M. A. Rodriguez, "Semantic Integrity Constraints for Spatial Databases", Proceedings of the 3rd Alberto Mendelzon International Workshop on Foundations of Data Management, Arequipa, Peru, May 12-15, 2009.
- [32] R. Villarreal, E. Fernández-Medina, M. Piattini and J. Trujillo, "A UML 2.0/OCL Extension for Designing Secure Data Warehouses", Journal of Research and Practice in Information Technology, 38(1), 2006.
- [33] J. Trujillo, E. Soler, E. Fernández-Medina and M. Piattini, "A UML 2.0 profile to define security requirements for Data Warehouses", Computer Standards and Interfaces, 31(5), pp. 969–983, 2009.
- [34] K. Boulil, F. Le Ber, S. Bimonte, C. Grag and F. Cernesson, "Multidimensional modeling and analysis of large and complex watercourse data: an OLAP-based solution", Ecological Informatics, 24, 90–106, 2014.
- [35] K. Boulil, S. Bimonte, F. Pinet, "Conceptual model for spatial data cubes: A UML profile and its automatic implementation". Computer Standards & Interfaces, 38, pp 113–132, 2015.



Abdallah Bensalloua Charef

Specialised in software engineering, Spatial Techniques, Geographical Information Systems (GIS), Remote Sensing and Geo-decisional Systems. He is computer science professor at the University of Mostaganem in Algeria.



Hamdadou Djamil

Specialised in Decision Support Systems, Multicriteria Analysis and Collaborative and spatial decisional Systems. She is professor and leads a research team at the laboratory of computer science of Oran (LIO) at the University of Oran I in Algeria.

A Study on Persuasive Technologies: The Relationship between User Emotions, Trust and Persuasion

Wan Noorashya Wan Ahmad ^{1,2}, Nazlena Mohamad Ali *

¹ Institute of Visual Informatics, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor (Malaysia)

² Faculty of Computing & Informatics, Universiti Malaysia Sabah, Labuan International Campus, F.T. Labuan (Malaysia)

Received 12 November 2017 | Accepted 14 February 2018 | Published 23 February 2018



ABSTRACT

A successful persuasive technology is able to persuade people to change from one state to a more well known state. Therefore, to allow for a change, persuasive technology must be able to affect users' emotion and make the user trust the technology so that they will adopt the persuasive technology into their daily life routine, as well as continue to use the technology for long period. This paper is aimed to study the relation between users' emotion with trust and persuasion and how they may contribute to the success of changing a person attitude or behavior towards a certain context or issue. Twenty five participants have completed the study in 6 weeks by using two types of persuasive technology that were assessed at three different interaction stages: pre, during and post. Result shows that emotions have a significant effect on trust, whereas the effect of emotions on persuasion using the persuasive technology was mediated by trust.

KEYWORDS

Emotion, Persuasion, Trust, Persuasive Technology.

DOI: 10.9781/ijimai.2018.02.010

I. INTRODUCTION

NOWADAYS, computing technologies are not only designed to help human performing daily tasks such as doing the administration work, or teaching in the classroom, but also to persuade and motivate people to change an attitude or behavior toward issues or objects. This type of computing technology known as persuasive technology is defined as "a technology that is designed to change attitudes or behavior through persuasion and social influence, but not through coercion" [1]. To persuade simply means to convince and therefore, trust is one of the key aspects of persuasion. It is the key to underpin confidence in user in using persuasive technology and also gives credence on the provided information or advices delivered through it which usually manifest themselves in the form of a change in attitude or behavior [2]. Since persuasion used specific strategies to elicit emotions in persuadees, that differ it from conviction where the strategies are primarily based upon reasoning [3]. Hence, as a consequence, emotion could play an important role in supporting behavior to build trust.

Particularly, there has been an increasing study on the influence of emotion on trust [4][5][6]. Those studies proved that emotion has brought impact and may alter the decision to trust with different emotion having different impacts. However, similar study that is related to persuasive technology is yet to found except [7][8]. Clearly, persuasion is a positive way compared to the negative meaning associated with coercion, thus emotion is consider to be the ideal means

to promote trust since it plays an important role in transmitting the induced emotion from a computer or computer software to the user [2]. This study is indeed motivated by the notion that (i) user's emotions is another important component towards the establishment of trust [9] and (ii) emotion could have a powerful influence over cognition and decision-making that could lead to cognitive actions to change someone attitude, as suggested by a number of research studies and theoretical models [10]. Thus, the objective of this study is to examine the relationship between emotion, trust and persuasion and how they affect each other. The next section of this paper will further describe the proposition of the relationship and the methodology used to study the proposition. Next, the result and discussion is presented to explain the discovered relationship. Lastly, the paper ends with conclusion.

II. BACKGROUND STUDY

Emotion has been defined in several ways by researchers in scientific community. According to Johnson-Laird et al [11], emotion is a feeling of emotional states, whereas Ortony et al [12] stated that emotions involved positive or negative valence, for example happy is a positive valence emotion while anger is under negative valence dimension. Emotions are acute, intentional states, which exist in a relatively short period of time, are related to a particular event, object, or action [13].

There were three reasons that make emotion a primary aspect of the experience of trust [14]; (i) experience of trust embodies affect whether in terms of intense feeling or subtle, (ii) different affective state may affect a person's experience of trust in making judgment towards others trustworthiness and (iii) trust is part of emotional that built on expectations. The lack of trust problem was still prominent and become the main concern in using persuasive technology

* Corresponding author.

E-mail addresses: wanaishya@gmail.com (Wan Noorashya Wan Ahmad), nazlena.ali@ukm.edu.my (Nazlena Mohamad Ali).

[15]. Therefore, emotion is seen as the importance component that contributes to trust a persuasive technology. Not to mention, a number of researchers have put emotions as an important consideration in understanding the development of trust and its changes, for example, Williams [16] and Andersen and Kumar [17] suggested that emotions influence how people judge on others' trust, Tomlinson and Mayer [18] agreed that emotions have impact on trust repair, and a study showed that people experiencing positive valence (i.e. happy and gratitude) are more receptive to increase their trust level [6]. This could somehow summarize that expectation that we expect upon interaction is constitute by positive or negative feeling that we felt which lead to judgment or decision-making about something or someone.

In this study, emotions are classified into two aspects of condition, namely the emotional states and emotional experience. Mental representation is the idea of something that could be reported [19] which in this case, it is related to the emotional state. Emotional states is referred to as conscious experience of emotional states that can be reported in terms of emotional words such as angry, happy, scared and sad. In appraisal theory, feeling is a mental representation of emotional experience, a state of conscious experience of emotion [13]. Emotional experience is the emotional response resulting from individual experience in using technology [20]. Emotional experience that user experienced is depending on the appropriateness of the interaction events with user's goal and values, and how user control the interaction event as well as their reaction towards the event [20]. Therefore, emotional experience can be used to show user 's emotional response that classified as pleasant (positive) or unpleasant (negative) resulting from interaction with technology.

Positive emotional response can build a sense of trust and engagement with users. People will forgive shortcomings occurred in an application or technology if the application or technology reward them with positive emotion. Therefore, a deep understanding of user emotions is an important aspect to establish trust and strengthen the persuasion process. Thus, we hypothesized that the emotion construct (i.e. emotional states and emotional experience) will predict trust in persuasive technology.

H1: *Emotion will positively affect trust in persuasive technology.*

A scientific definition of trust that is well accepted across disciplines is unclear. We illustrated several researchers' definition on trust that suited to our study. Mc Allister [21] defined trust as "the extent to which a person is confident in, and willing to act on the basis of, the words, actions, and decisions of another". In addition, Rousseau et al. [22] established trust definition as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behaviour of another". However, trust in persuasive technology is defined as the expectation that users have on the technology to perform as it should be without harming the user [23]. Based on the overall definition of trust, we view trust in this study as the confidence towards the persuasive technology in carry out things as it supposed to do.

Consistent with the literature study on human perception in making trust evaluation [21][24], users' trust consist of cognitive and affective trust. Cognitive trust is referred to as elements that can increased trustworthiness towards a technology whereas affective trust is emotional response towards a technology [25][26]. As both perceptual beliefs are interconnected [21], therefore, trust in persuasion technology should be studied in both aspects to study the overall users' trust. Since trust is one of the key elements in the process of persuasion [1], one need to be persuaded repetitively in repetitive interactions which will only possible to happen if users trust the systems. Since [6] have discovered that emotion has influence on trust, we expect that its effect on persuasion using persuasive technology would be positively mediated.

H2: *Trust in persuasive technology will positively mediated the relationship between emotion and persuasion using persuasive technology.*

H3: *Trust will positively affect persuasion using persuasive technology.*

In addition, persuasion is defined as an individual's subjective evaluation of persuasive technology used and its impact on the self [27]. It is a form of influence to change the way a person's believe, behave, or feel [23]. Crano and Prislin [28] argue that the main aspect to be taken into account when it comes to persuasion involves the construction of basic attitudes. Persuasive technology is a human-computer persuasion since it promotes a method of persuasion using a computer technology [1]. Although computer is a non-living object, the intentionality to persuade is actually comes from the creator, distributor and the people who adopt the technology. The fact that trust is an important attributes in human-computer interaction (HCI) where trust has relation with computer designer's responsibilities to ensure that people who adopt the technology to change their attitude or behavior should achieve the desired intend of what the designers want to accomplish [23]. In the beginning, creating and providing people with persuasive technology may work primarily, however, after sometime it may create persuasion disorder and weaken trust [9]. Thus, emotion is seen as a component that can build confidence and strengthen the power of persuasion in persuasion technology. Thus, the following hypothesis is built.

H4: *There is a relationship between emotion, trust and persuasion in the use of persuasive technology.*

III. METHODOLOGY

To investigate the hypothesized role of emotion in this study, a total of 30 participants volunteered university students and staff from public and private university were employed for the study, in which 25 of them managed to complete the study in 6 weeks to use and evaluate two persuasive technology, fitness application and environmental game each represent different types of persuasive technology; tool and medium. The tool category PT consist of *MyFitnessPal*¹, *Fitocracy*² and *MapMyFitness*³ that allows for fitness monitoring including physical activity or/and food consumption. The medium category PT, *Fitocracy* and *MapMyFitness* have same goal to create awareness on issues of environment.

Different experimental design approaches were used to measure different aspect of evaluation. User emotions were measured at three stages of interaction; pre, during and post, while user's trust were measured in pre-post interaction stages, whereas, persuasion were measured in post-interaction stages. Quantitative measurement scales for each aspects of evaluation were adopted from several studies. Each measurement for aspects of evaluation was using a 5-point Likert scales. Two types of measures were used to evaluate construct of emotion; Geneva Emotion Wheel (GEW) [13][29] that consist of 20 items is used to measure emotional states while Positive and Negative Affect System (PANAS)[30] which also consist of 20 items (10 items of positive valence, 10 items of negative valence) is used to measure emotional user experience. Meanwhile, measurements of trust in PT consist of 10 items. Four of them measured cognitive trust [31] whereas six items measured affective trust [25]. User evaluations towards the PT and its effect on user after 6 weeks of usage were measure using three items from [27].

The hypotheses testing was conducted using regression and correlation analyses to investigate the effect of variables (i.e. emotion, trust) towards persuasion using persuasive technology

1 <http://www.myfitnesspal.com>

2 <http://www.fitocracy.com>

3 <http://www.mapmyfitness.com>

from pre-interaction to post-interaction as suggested by the theory of prominence-interpretation [1]. For that reason, a dummy (D) variable is created to clean the time factors in emotion and trust variable since each variables were measured at three (pre-during-post) and two (pre-post) interaction stages, but not for persuasion variable as it was measured in post-interaction stage only. The value of “0” is given to pre-interaction and value “1” is given to post-interaction.

IV. FINDINGS AND DISCUSSIONS

The background profile of participants is presented in Table I. The study which involved 25 participants was dominated by fifteen female participants compared to ten male participants in whom eighteen of them are Malays. The participants come from five different age groups; each seven of them from group below 25 years, between age 26 to 30 years, and between age 31 to 35 years, while three participants were from age group 36 to 40 years and only one participant was above 40 years old. These participants consist of twenty two university students from UKM, UiTM and UPM, as well as eight university employees from UKM, Nottingham University and UniKL. Out of the twenty five participants, 32 percent of them were from the IT background.

The results analyses of the hypotheses testing are presented in Table I and Table II.

From the results in Table I, we can see that the correlation coefficient suggests that the user’s emotions and user’s trust in persuasive technology have a moderate linear relationship, and found to be .195 indicates 19.5% of the variance in trust, $F(2, 47) = 5.69, p < 0.05$ shows that user’s trust changes significantly with respect to differences in user’s emotion. The equation on the influence of user’s emotions towards user’s trust in using persuasive technology is $2.305 + 0.074\text{Emotion}$. The relationship between emotion and trust in persuasive technology was positive ($\beta = .353$) with trivial effect size ($B = 0.07$). As predicted, effect on user’s emotions influenced user’s trust in using persuasive technology, thus supporting H1. This mean, for each increment or decrement happens in user’s emotions will affect the increment or decrement of user’s trust towards persuasive technology. Result from this finding demonstrates that positive emotions (increment in valence) will increase user’s trust level, whereas negative emotions (decrement

in valence) will led to the decrement of user’s trust level [4][6][32].

Our model of emotion which constitute of emotional states and emotional experience explains 13.5% of the variance in trust. Of these two variables, emotional experience makes the largest unique contribution ($\beta = .345, p < 0.05$) compared to emotional states that makes insignificant contribution ($\beta = .044, p > 0.1$) to user to trust persuasive technology. This finding was unexpected and suggests that the mixed emotional states (i.e. consist of different valence and control) that users experienced and rated in evaluating the persuasive technologies have lead to this result in which different emotional states have different effect on trust. This finding corroborates the findings of [33], who found that different emotions caused different effects on people’s judgment. Moreover, based on the reasons behind the experienced emotional states that users felt in this study, most of the users’ emotional states were triggered by individual lead control, in this case, emotions of self-control (user) and emotions of other-person control (system, i.e. persuasive technology). It seems possible that these results are due to the simultaneous experience of positive and negative emotional states that occurred from different source of control in which [4] claimed that emotional states trigger by self-control will have no influence on trust, while emotional states trigger by the used systems will determined users whether or not to trust. Hence, measuring emotional states as a whole seems to not affecting trust.

In Model 1 of the regression model on persuasion in Table I, user’s emotions is significantly ($p < 0.05$) related to persuasion with 25.1% variance in persuasion. However, to examine the mediation effect of trust on the relationship between user’s emotion and persuasion using persuasive technology, correlation coefficient in Model 2 suggests a strong linear relationship between user’s emotions and persuasion that mediates by user’s trust which can be explained by 74.3% of the variance in persuasion using persuasive technology, $F(2, 22) = 31.73, p < 0.001$. Since the user’s emotion variables exerts its total influence through the mediating variable, $p > 0.1$ ($\beta = .051, p = 0.69$), there exists full mediation by the trust variable giving support to H2. The equation suggesting that trust mediates the relationship between user’s emotions and persuasion $-0.758 + 0.015\text{Emotion} + 1.129\text{Trust}$. This finding suggests that users’ trust will increase when they felt positive emotions; thus the likelihood for the users to be persuaded is higher when they

TABLE I. SUMMARY OF REGRESSION FOR VARIABLES PREDICTING TRUST AND PERSUASION

Dependent Variables	Trust			Persuasion					
	Model 1			Model 1			Model 2		
	B	SE B	β	B	SE B	β	B	SE B	β
Independent Variable									
User emotions X ¹	.073	.028	.353**	.148	.054	.501**	.015	.038	.051
Mediator									
Trust in persuasive technology Y ₁							1.129	.174	.833***
R		.442			.501			.862	
R ²		.195			.251			.743	
Adj. R ²		.161			.218			.719	
R ² Change		.195			.251			.492	
F		5.692**			7.668**			31.732***	

Note: *** $p < 0.001$, ** $p < 0.05$, * $p < 0.1$

TABLE II. PEARSON PRODUCT-MOMENT CORRELATION BETWEEN VARIABLES

Variables	X	Y	Z
X. Emotion	-		
Y. Trust	.349*	-	
Z. Persuasion	.501*	.861**	-

Note: ** $p < 0.001$ (2-tailed), * $p < 0.05$ (2-tailed)

have high trust in the persuasive technology which corroborates the ideas of [9], who suggests design for persuasion, emotion and trust. Trust is important to persuasion in using persuasive technology as it may affect users' intention to use the technology as well as leverage the continuation used of the technology [34]. The effect of emotions that were exerted into trust should encourage user to develop confident towards the persuasive technology and at the same time increase the persuasion effect on them.

As shown in Table II, there was a moderate, positive correlation between user emotions and trust in persuasive technology ($r_{xy}=.35, p<0.05$), with positive emotions associated with high level of trust and negative emotions associated with low level of trust. A strong, positive correlation was found between user emotions and persuasion using persuasive technology ($r_{xz}=.50, p<0.05$) as well as between trust in persuasive technology and persuasion using persuasive technology ($r_{yz}=.86, p<0.001$) with the positive association indicates that increases in one variable correspond to increases in the other.

V. CONCLUSION

A persuasive technology should be design to be able to elicit positive emotions in users using different persuasion principles or strategies in making them to trust the technology and thus, successfully persuade users to the target attitude or behavior. As predicted, the relationship between users' emotions in using persuasive technology and trust towards persuasion is found to be positive. It can be presumed that positive emotions increase user's trust, while negative emotions will impact user's trust by decreasing it. The effect of emotions on persuasion is found to be mediated by trust positively and thus indicates the association between the three variables. Nevertheless, this study managed to shed some lights on the effect of users' emotions on users' trust and persuasion in using persuasive technology. The findings could encourage for further investigation on emotions elements used in persuasive technology that could leverage trust and persuasion.

ACKNOWLEDGEMENTS

This work is part of the research funded by Ministry of Education, Malaysia under the Fundamental research Grant Scheme (FRGS) (FRGS/2/2014/ICT01/UKM/02/3).

REFERENCES

- [1] B. J. Fogg, *Persuasive Technology. Using Computers to Change What We Think and Do*, Morgan Kaufman, San Francisco (2003).
- [2] W. N. Wan Ahmad and N. Mohamad Ali, Engendering Trust Through Emotion in Designing Persuasive Application, in *Advances in Visual Informatics*, Edited H. Badioze Zaman, P. Robinson, P. Olivier, T. K. Shih and S. Velastin, Springer International Publishing, (2013), Lecture Notes in Computer Sciences, Vol. 8237, pp. 707-717.
- [3] M. Harjumaa and H. Oinas-Kukkonen, *Persuasion Theories and IT Design*, Edited Y.de Kort, W. IJsselsteijn, C. Midden, B. Eggen, B.J Fogg, Springer-Verlag Berlin Heidelberg, (2007), Lecture Notes in Computer Sciences, Vol. 4744, pp 311-314.
- [4] J. R. Dunn and M. E. Schweitzer, Feeling and Believing: The Influence of Emotion on Trust, *Journal of Personality and Social Psychology*, 88, 5, (2005).
- [5] W. Lee and M. Selart, The Impact of Emotion on Trust Decision, in *Handbook on Psychology of Decision-Making*, Edited N. Moore, K.O. and Gonzalez, Nova Science Publishers, NY (2011), pp. 1-16.
- [6] D. Myers and D. Tingley, *The influence of Emotion on Trust*, Academic Dissertation, Princeton University, NJ (2011).
- [7] W. N. Wan Ahmad, Development of Emotion-based trust Model for Desinging Persuasive Application, 9th International Conference on Persuasive Technology, PERSUASIVE 2014, (2014) May 21-23; Padova, Italy.
- [8] W. N. Wan Ahmad and N. Mohamad Ali, Investigation into Trust and Emotion, *New Zealand Journal of Human-Computer Interaction*, 1 (2016).
- [9] E. Schaffer, *PET Research: Looking Deeper to Understand Motivations*. Human Factors International (2010).
- [10] M. Lewis, J. M. Haviland-Jones and L. F. Barrett, *Handbook of Affective Sciences*, Oxford University Press, Oxford (2009).
- [11] P. N. Johnson-Laird and K. Oatley, *The Language of Emotions: An Analysis of a Semantic Field*. *Cognition and Emotion*, 3,(1989).
- [12] A. Ortony, G. L. Clore and A. Collins, *The Cognitive Structure of Emotions*. Cambridge University Press, New York (1988).
- [13] K. R. Scherer, What are emotions? and How Can They be Measured?, *Social Science Information*, 44, 4 (2005).
- [14] G. Jones and J. M. George, The Experience and Evolution of Trust: Implications for Cooperation and Teamwork, *Academy of Management Review*, 23, 3 (1998).
- [15] W. N. Wan Ahmad and N. Mohamad Ali, Trust Perceptions in Using Persuasive Technologies, 3rd International Conference on Computer and Information Sciences (ICCOINS 2016), (2016) August 15-17; Kuala Lumpur, Malaysia.
- [16] M. Williams, In Whom We Trust: Group membership as an Affective Context for Trust Development, *The Academy of Management Review*, 26, 3 (2001).
- [17] E. C. Tomlinson and R. C. Mayer, The Role of Causal Attribution Dimensions in Trust Repair, *The Academy of Management Review*, 34, 1 (2009).
- [18] P. H. Andersen and R. Kumar, Emotions, Trust and Relationship Development in Business Relationships: A Conceptual Model for Buyer-Seller Dyads, *Industrial Marketing Management*, 35, 4 (2006).
- [19] C. Frith, R. Perry and E. Lumer, The Neural Correlates of Conscious Experience: An Experimental Framework, *Trends in Cognitive Sciences*, 3, 3 (1999).
- [20] J. Jokinen, Emotional User Experience: Traits, Events and States, *International Journal of Human-Computer Studies*, 76 (2015).
- [21] D. J. McAllister, Affect and Cognition-based Trust as Foundations for Interpersonal Cooperation in Organizations, *Academy of Management Journal*, 38, 1(1995).
- [22] D. M. Rousseau, S. B. Sitkin, R. S. Burt and C. Camerer, Not So Different After All: A Cross-discipline View of Trust, *Academy of Management Review*, 23, 3 (1998).
- [23] P. Verbeek, Persuasive Technology and Moral Responsibility towards an Ethical Framework for Persuasive Technologies, First International Conference on Persuasive Technology for Human Well-Being, Persuasive06, (2006) May 23-26; Eindhoven, The Netherlands.
- [24] D. H. McKnight and L. N. Chervany, What Trust Means in E-Commerce Customer Relationships: An Interdisciplinary Conceptual Typology, *International Journal of Electronic Commerce*, 6, 2 (2002).
- [25] M. Madsen and S. Gregor, Measuring Human-Computer Trust, *Proceedings of the Eleventh Australasian Conference on Information Systems*, (2000) December 6-8; Brisbane, Australia.
- [26] J. Riegelsberger, M. A. Sasse and J. D. McCarthy, The Mechanics of Trust: A Framework for Research and Design, *International Journal of Human-Computer Studies*, 62, 3 (2005).
- [27] T. Lehto, H. Oinas-Kukkonen and F. Drozd, Factors Affecting Perceived Persuasiveness of a Behavior Change Support System, *Proceedings of 33rd International Conference on Information Systems*, (2012) December 16-19; Orlando, USA.
- [28] W. D. Crano and R. Prislin, Attitudes and Persuasion, *Annu. Rev. Psychol.*, 57 (2006).
- [29] K. R. Scherer, V. Shuman, J. R. J. Fontaine and C. Soriano, The GRID meets the Wheel: Assessing Emotional Feeling via Self-report, in *Components of Emotional Meaning: A Sourcebook* Edited J. R. J. Fontaine, K. R. Scherer, and C. Soriano, Oxford University Press, Oxford (2013), pp. 281-298.
- [30] D. Watson, L. A. Clark, and A. Tellegen, Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales, *Journal of Personality and Social Psychology*, 47 (1988).
- [31] D. H. McKnight, M. Carter, J. B. Thatcher and P. F. Clay, Trust in a Specific Technology: An Investigation in Its Components and Measures, *ACM Transactions on Management Information Systems*, 2, 2 (2011).
- [32] E. Schniter and R. M. Sheremeta, Predictable and Predictive Emotions:

Explaining Cheap Signals and Trust Re-Extension, *Frontiers in Behavioral Neuroscience*, 8, November (2014).

- [33] J. S. Lerner and D. Keltner, *Beyond Valance: Toward a Model of Emotion-specific Influences on Judgement and Choice*, *Cognition and Emotion*, 14, 4 (2000).
- [34] H. Sun and P. Zhang, *The Role of Moderating Factors in User Technology Acceptance*, *International Journal of Human-Computer Studies*, 64, 2 (2006).



Wan Nooraishya Wan Ahmad

A lecturer at Universiti Malaysia Sabah who is currently a Ph.D student at Institute of Visual Informatics (IVI), Universiti Kebangsaan Malaysia studying about emotion and trust for persuasive applications. She received her Master degree in Creative Software System from Heriot-Watt University in the year of 2009. Her current research interests include affective computing, trust in human-

computer interaction and persuasive technology.



Nazlena Mohamad Ali

She received her Ph.D. degree from Dublin City University, Ireland in 2009. She is currently a senior research fellow at Institute of Visual Informatics (IVI), Universiti Kebangsaan Malaysia. Her major field of study is Human Computer Interaction in particular: interaction design, user evaluation and usability.

EEG Signal Analysis of Writing and Typing between Adults with Dyslexia and Normal Controls

Harshani Perera¹, Mohd Fairuz Shiratuddin¹, Kok Wai Wong¹, Kelly Fullarton² *

¹ School of Engineering and Information Technology, Murdoch University, Murdoch (Australia)

² DSF Literacy and Clinical Services, (The Dyslexia-SPELD Foundation), South Perth (Australia)

Received 12 November 2017 | Accepted 14 February 2018 | Published 20 April 2018



ABSTRACT

EEG is one of the most useful techniques used to represent behaviours of the brain and helps explore valuable insights through the measurement of brain electrical activity. Hence, plays a vital role in detecting neurological conditions. In this paper, we identify some unique EEG patterns pertaining to dyslexia, which is a learning disability with a neurological origin. Although EEG signals hold important insights of brain behaviours, uncovering these insights are not always straightforward due to its complexity. We tackle this using machine learning and uncover unique EEG signals generated in adults with dyslexia during writing and typing as well as optimal EEG electrodes and brain regions for classification. This study revealed that the greater level of difficulties seen in individuals with dyslexia during writing and typing compared to normal controls are reflected in the brainwave signal patterns.

KEYWORDS

Dyslexia, Electroencephalography, Machine Learning, Classification, Support Vector Machines.

DOI: 10.9781/ijimai.2018.04.005

I. INTRODUCTION

DYSLEXIA is a hidden specific learning disability that affects a significant amount of the world population, it is neurological in origin and causes difficulties in reading and spelling despite average or above average intelligence and acceptable exposure to literacy instructions [1-3]. Dyslexia screening assessments are based on a review of biographical information, educational history, behavioural aspects and academic indicators such as reading, spelling, writing, working memory and processing abilities [4]. Common symptoms of dyslexia include failure to attain sufficient reading skills and poor writing skills compared to peers despite conventional teaching guidelines. Typing is a modern-day task that often replaces writing, but still, affects people with dyslexia in a similar manner when it comes to spelling.

In addition to these behavioural symptoms seen externally, past studies have uncovered neurological differences in individuals with dyslexia. These include unique brain structures as well as distinctions in brain behaviours compared to normal controls [5]. Electroencephalography, commonly known as EEG, is a technique that helps to capture neurological behaviours. In our previous work we have covered past work carried out to identify unique brainwave activation patterns using EEG and identified gaps to be filled in the literature about these unique EEG signal patterns pertaining to dyslexia, in particular the EEG patterns while performing tasks that are more challenging for individuals with dyslexia [1, 6]. Hence, in this paper we have selected two of such tasks that are more challenging

for individuals with dyslexia, namely writing and typing and aim to identify unique EEG signal patterns generated while performing these tasks. The purpose of the analysis is to identify whether the difficulties seen in individuals with dyslexia during writing and typing are reflected in the brainwave signal patterns. This paper covers the EEG signal acquisition, signal processing, classification using machine learning and future work.

II. EEG SIGNAL ACQUISITION

The EEG headset used for this research was the Cognionics 32-channel dry EEG headset, and the EEG was recorded at a sampling rate of 300Hz. The EEG channel map is depicted in Fig. 1 where the channels used on this specific EEG headset are indicated in grey. This research was carried out with a total of 32 participants, where 17 participants were individuals with dyslexia (7 males and 10 females) and 15 participants were normal controls (8 males and 7 females). The number of participants was determined using the Altman's Nomogram sample size calculation as shown in Fig. 2. Therefore, for a power of 0.80 (p-value significance of 0.05) and a standardised difference value between 0.8 and 1.0 (Cohen's d effect size), the total amount of participants would range between 30-50. Hence, the number of participants per group would range between 15-25. The inclusion and exclusion criteria of participants was 18 years and above, right-handed, fluent in English, have a normal or corrected-to-normal vision and normal hearing. The participants in the group with dyslexia had to be diagnosed by a psychologist as having dyslexia and the control group had to be free from motor and neurological conditions such as dyslexia, ADHD and autism. The participants with dyslexia were recruited with the help of DSF Literacy and Clinical Services in Western Australia (The Dyslexia-SPELD Foundation).

* Corresponding author.

E-mail addresses: h.perera@murdoch.edu.au (H. Perera), f.shiratuddin@murdoch.edu.au (M. F. Shiratuddin), k.wong@murdoch.edu.au (K. W. Wong), kellyfullarton@dsf.net.au (K. Fullarton).

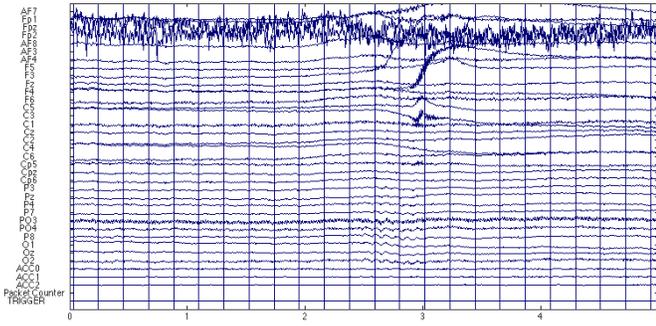


Fig. 5. Raw EEG.

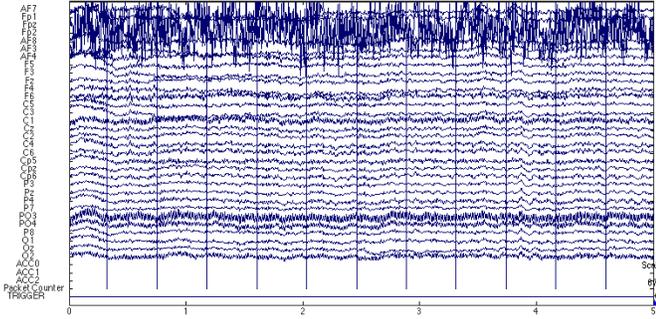


Fig. 6. ASR filtered EEG.

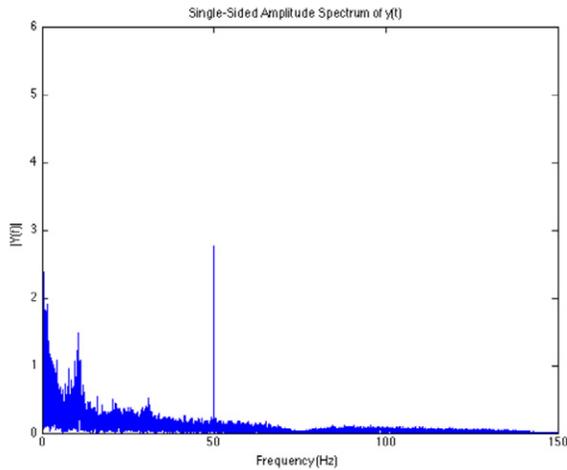


Fig. 7. Electric power noise at 50Hz.

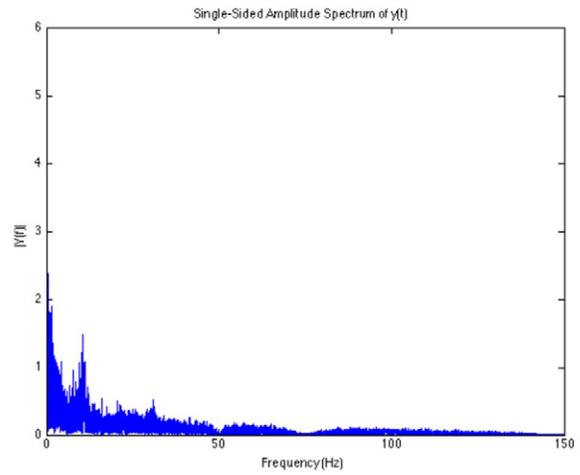


Fig. 8. Filtered EEG.

B. Sub-band Decomposition

In this research, the EEG signals are analysed by decomposing the EEG signals into pre-defined sub-bands (Fig. 9). The sub-bands are namely delta, theta, alpha, beta and gamma. The sub-band decomposition was performed using band-pass FIR digital filters. Next, the frequency domain transformation was performed using MATLAB's FFT function. This function returns the Discrete Fourier Transform (DST) computed using a FFT algorithm.

C. Feature Extraction

A total of 8 features mean, median, mode, standard deviation, maximum, minimum, skewness and kurtosis were calculated for each participant, for each task, at each of the 5 frequency sub-bands (delta, theta, alpha, beta and gamma) in each of the 32 channels. All of these features collectively represent important characteristics of the EEG signal datasets. This adds up to a total of 1280 predictors per participant, which will be the input for the classifiers.

IV. EEG CLASSIFICATION

Previous studies [1, 10, 11] show that Support Vector Machines (SVM) is one of the most suitable classifiers to be used for EEG classifications. Hence, in this research we perform the classification of EEG using Cubic Support Vector Machines. Further, in addition to creating classifiers with all the EEG channels as a whole, classifiers were also created for different segments of the brain as illustrated in

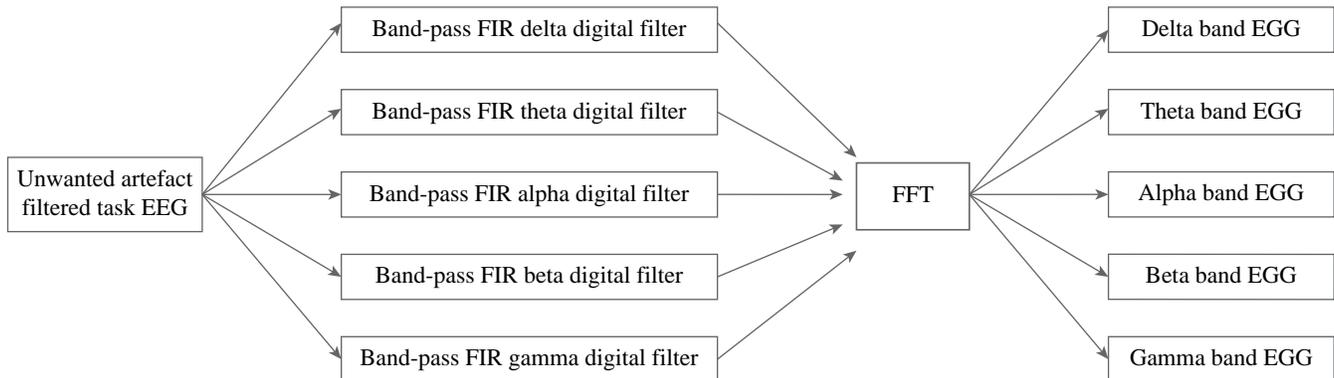


Fig. 9. Overview of sub-band decomposition and frequency domain transformation.

Table I. This helps to identify sections of the brain that have more prominent EEG activation patterns.

TABLE I. FEATURE GROUPING

Area		Channels
Brain Left Hemisphere		Fp1, AF7, AF3, F5, F3, C5, C3, C1, Cp5, P3, P7, PO3, O1
Brain Right hemisphere		Fp2, AF8, AF4, F4, F6, C2, C4, C6, Cp6, P4, P8, PO4, O2
Brain Centre		Fpz, Fz, Cz, Cpz, Pz, Oz
Frontal Lobe	Frontal Pole	Fp1, Fpz, Fp2
	Anterior-Frontal	AF7, AF3, AF4, AF8
	Frontal	F5, F3, FZ, F4, F6
Central Lobe	Central	C5, C3, C1, Cz, C2, C4, C6
	Centro-Parietal	Cp5, Cpz, Cp6
Parietal Lobe	Parietal	P3, Pz, P4, P7, P8
	Parieto-Occipital	PO3, PO4
Occipital Lobe		O1, Oz, O3

The classifier outputs were measured based on the Validation Accuracy (VA), Sensitivity/True Positive Rate (TPR) and Specificity/True Negative Rate (TNR) that were calculated using the resulting confusion matrix as shown in Fig. 10 and (1), (2) and (3).

True Class	dyslexic	True Positive (TP)	False Negative (FN)
	non-dyslexic	False Positive (FP)	True Negative (TN)
		dyslexic	non-dyslexic
		Predicted Class	

Fig. 10. Confusion matrix.

$$TPR = \frac{TP}{(TP | FN)} \times 100 \quad (1)$$

$$TNR = \frac{TN}{(TN | FP)} \times 100 \quad (2)$$

$$VA = \frac{TP | TN}{(TP | FP | FN | TN)} \times 100 \quad (3)$$

V. RESULTS AND DISCUSSION

Poor writing skills are one of the commonly seen difficulties in individuals with dyslexia. The classifier results from the writing task, which is summarised in Table II, verify that adults with dyslexia produce unique brainwave signal patterns compared to normal controls. The peak VA of 71.88%, a sensitivity of 76.47% and specificity of 66.67% was produced from the anterior frontal classifier, which included the EEG electrodes AF7, AF3, AF4 and AF8. However, this outcome has not previously been reported in previous similar studies, and a possible explanation for this might be that because those studies had not used the EEG electrodes AF7, AF3, AF4 and AF8. The channels used in these similar studies were C3, C4, P3 and P4 [12-14]. Therefore, these results contribute towards to the pool of knowledge as a new finding. Fig. 11 depicts the positions of AF7, AF3, AF4 and AF8.

TABLE II. WRITING TASK CLASSIFIER RESULTS

Brain Area	VA %	Sensitivity %	Specificity %
All	59.38	64.71	53.33
Left Hemisphere	65.63	70.59	60.00
Right Hemisphere	50.00	64.71	33.33
Frontal Lobe	56.25	64.71	46.67
Central Lobe	59.38	64.71	53.33
Parietal Lobe	59.38	64.71	53.33
Occipital Lobe	62.50	64.71	60.00
Parieto-Occipital	46.88	58.82	33.33
Parieto-Occipital Left	46.88	52.94	40.00
Parieto-Occipital Right	59.38	58.82	60.00
Anterior Frontal	71.88	76.47	66.67

TABLE III. TYPING TASK CLASSIFIER RESULTS

Brain Area	VA %	Sensitivity %	Specificity %
All	78.13	88.24	66.67
Left Hemisphere	71.88	94.12	46.67
Right Hemisphere	62.50	76.47	46.67
Frontal Lobe	68.75	88.24	46.67
Central Lobe	68.75	82.35	53.33
Parietal Lobe	65.63	76.47	53.33
Occipital Lobe	56.25	82.35	26.67
Parieto-Occipital	62.50	70.59	53.33
Parieto-Occipital Left	68.75	76.47	60.00
Parieto-Occipital Right	68.75	76.47	60.00
Anterior Frontal	65.63	88.24	40.00
Central	68.75	76.47	60.00
Centro Parietal	59.38	76.47	40.00
Frontal Pole	68.75	94.12	40.00
Frontal	78.13	88.24	66.67
Frontal Left	68.75	82.35	53.33
Frontal Right	68.75	82.35	53.33

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we conducted research to identify whether adults with dyslexia produced unique brainwave signal patterns during writing and typing. The results show that adults with dyslexia show unique brainwave activation patterns during each task compared to normal controls. Although similar writing tasks had been investigated in past studies, the current research was conducted with additional EEG sensors and discovered a new optimal brain region anterior frontal, which has not been reported in past studies. On the other hand, the research results also uncovered novel findings for typing as this task that had not been analysed in past similar studies. This research contributes vital insights to the pool of knowledge about the unique brainwave patterns of adults with dyslexia, which could serve as a base for future studies, and could even one day, help complement the conventional dyslexia diagnosis process by giving a better view of the disability through the introduction of neurological aspects.

These preliminary findings can be further examined by making variations in parameters such as input features, channels, frequency sub-bands, kernels and more advanced classifiers such as Fuzzy SVM. This could perhaps lead towards the enhancement of result accuracies similar to how the current research obtained better results by making variations in the EEG sensors used for each classifier. The scope of this research was limited to right-handed adults. Further studies can be carried out in order to compare EEG signals of individuals below 18 years and left-handed. Comparisons of the EEG signals could also be made between the genders male and female. Further, this research can be expanded in order to identify unique brainwave signal patterns of other specific learning disabilities such as dysgraphia and dyscalculia. Lastly, the function of each brain region needs to be compared with the result outcomes in order to identify the neurological reason behind each discovery.

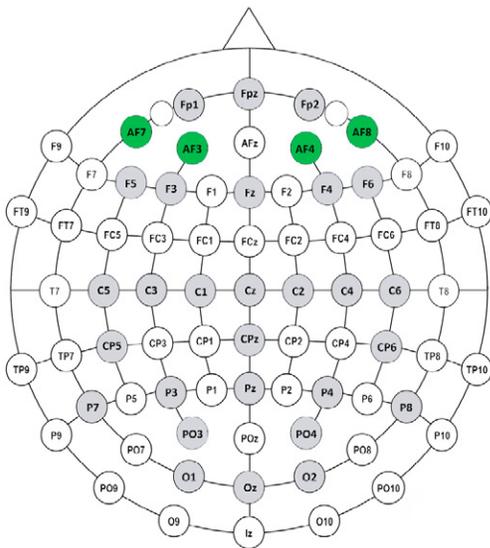


Fig. 11. Optimal channels for writing.

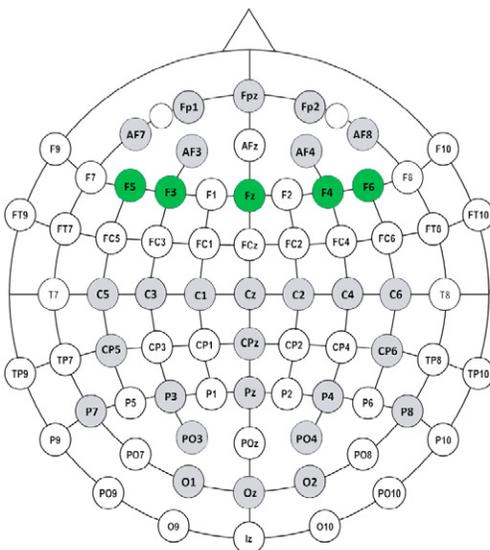


Fig. 12. Optimal channels for typing.

Typing can be considered as the modern-day replacement to writing and is yet another task found more challenging by individuals with dyslexia. Table III illustrates the behaviour of seventeen classifiers built to analyse the typing task. We examined the left hemisphere, right hemisphere, frontal lobe, central lobe, parietal lobe and the occipital lobe. Except for the parietal lobe, others showed a substantial difference between the sensitivity and specificity rates, which is not preferable. The classifiers from parietal and parieto-occipital performed fairly well. The frontal classifier showed the top VA of 78.13% with a fairly balanced specificity and sensitivity. Interestingly, this was close to the most significant region identified for writing, which was the anterior-frontal. The most significant EEG channels responsible for producing unique brainwave signals in individuals with dyslexia compared to normal controls were F5, F3, Fz, F4 and F6. Fig. 12 depicts the position of these four channels. All these findings show that EEG signals generated while typing produce unique brainwave signal patterns in adults with dyslexia compared to normal controls. Further, comparison of EEG signal patterns between persons with and without dyslexia during typing is a gap to be filled in the literature; therefore, we did not find any research results that could be directly compared against our results.

REFERENCES

- [1] H. Perera, M. F. Shiratuddin, and K. W. Wong, «A Review of Electroencephalogram-Based Analysis and Classification Frameworks for Dyslexia,» in International Conference on Neural Information Processing, 2016, pp. 626-635.
- [2] J. M. Fletcher, G. R. Lyon, L. S. Fuchs, and M. A. Barnes, Learning disabilities: From identification to intervention: Guilford Press, 2006.
- [3] M. F. C. A. Rani, R. Rohizan, and N. A. A. Rahman, «Web-based learning tool for primary school student with dyscalculia,» in Proceedings of the 6th International Conference on Information Technology and Multimedia, 2014, pp. 157-162.
- [4] The Dyslexia-SPELD Foundation of WA. (n.d., 14 May). Consultations and Assessments. Available: <http://dsf.net.au/consultations-assessments/>
- [5] S. Mohamad, W. Mansor, and K. Y. Lee, «Review of neurological techniques of diagnosing dyslexia in children,» in System Engineering and Technology (ICSET), 2013 IEEE 3rd International Conference on, 2013, pp. 389-393.
- [6] H. Perera, M. F. Shiratuddin, and K. W. Wong, «Review of the Role of Modern Computational Technologies in the Detection of Dyslexia,» in Information Science and Applications (ICISA) 2016, K. J. Kim and N. Joukov, Eds., ed Singapore: Springer Singapore, 2016, pp. 1465-1475.
- [7] Cognionics Inc. (n.d., 17 July). Channel Diagram. Available: <http://cognionics.com/index.php/products/hd-eeeg-systems/quick-20-dry-headset>
- [8] M. Bland, «Sample size for clinical trials,» ed: University of York, York, UK, 2011.
- [9] T. Mullen, C. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, et al., «Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG,» in Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE, United States, 2013, pp. 2184-2187.
- [10] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, «Comparison of linear, nonlinear, and feature selection methods for EEG signal

classification.» *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 11, pp. 141-144, 2003.

- [11] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, «A review of classification algorithms for EEG-based brain-computer interfaces.» *Journal of neural engineering*, vol. 4, 2007.
- [12] N. Fuad, W. Mansor, and K. Y. Lee, «Wavelet packet analysis of EEG signals from children during writing.» in *Computers & Informatics (ISCI), 2013 IEEE Symposium on*, 2013, pp. 228-230.
- [13] C. W. N. F. Che Wan Fadzal, W. Mansor, and L. Y. Khuan, «An analysis of EEG signal generated from grasping and writing.» in *Computer Applications and Industrial Electronics (ICCAIE), 2011 IEEE International Conference on*, 2011, pp. 535-537.
- [14] A. Zabidi, W. Mansor, Y. K. Lee, and C. W. N. F. Che Wan Fadzal, «Short-time Fourier Transform analysis of EEG signal generated during imagined writing.» in *System Engineering and Technology (ICSET), 2012 International Conference on*, 2012, pp. 1-4.



Kelly Fullarton

Kelly Fullarton is a Senior Clinical Psychologist at DSF Literacy and Clinical Services (The Dyslexia-SPELD Foundation of Western Australia). Kelly has extensive experience in the field of educational psychology and learning disabilities and Kelly has been working at DSF for 12 years assessing children, teenagers and adults struggling with the acquisition of literacy and numeracy skills. Kelly is an AHPRA approved supervisor and has been supporting Psychologists across Australia to develop their ability to diagnose specific learning disorders utilising best practice guidelines and developing a consistent approach to assessment, diagnosis and support.



Harshani Perera

Harshani Perera received her BSc (Hons) in Software Engineering from the University of Wales, United Kingdom and proceeded with her career in software engineering. With her passion for research, she then chose to pursue a PhD in developing a detection mechanism of learning disabilities using intelligent data analysis and classification of brainwave signals and received her PhD in Information Technology from Murdoch University, Australia. Dr Perera's entrepreneurial passion also motivated her to co-found Inqbaytor Pty Ltd and currently continues to serve as the CEO.



Mohd Fairuz Shiratuddin

Mohd Fairuz Shiratuddin is a Senior Lecturer in the School of Engineering and Information Technology. In his early careers, he was trained as an engineer mainly dealing with computers and its applications in the United Kingdom and Malaysia. Then he decided to pursue a career in academia. He graduated with a Bachelor of Engineering degree in Electrical & Electronics from Northumbria University at Newcastle-upon-Tyne in the UK, a Master of Science degree in Information Technology (by research specialising in Virtual Reality) from Universiti Utara, Malaysia, a Master of Science degree in Architecture from Virginia Tech, USA, and a Doctor of Philosophy degree in Environmental Design & Planning also from Virginia Tech. Prior to working at Murdoch University, he was a Lecturer at Universiti Utara, Malaysia, and an Assistant Professor at the University of Southern Mississippi, USA. He always considers himself as a multi-disciplinary person with a lot of appreciation for arts and design, and technical development. His research interests include computer-based systems for children's teaching and learning, and stroke survivors' rehabilitation, Virtual/Mixed/Augmented Reality, Natural User Interfaces, Games Design, Development and Technologies, and Artificial Intelligence; for practical, real-world uses. He is currently leading Project Neuromender:: A Low-Cost Home-Based Stroke Rehabilitation System. He is also working on a project to utilise augmented reality, image processing and facial recognition technologies for security purposes. He has numerous publications in national and international conference proceedings, journals, books, book chapters and reports.



Kok Wai Wong

Kok Wai Wong is currently working as an Associate Professor with the School of Engineering and Information Technology at Murdoch University in Western Australia. He is a Senior Member of IEEE, a member of ACS, and Certified Professional of ACS. He is the current Vice President (membership) for The Asia Pacific Neural Network Society (APNNS). He has also held executive positions such as the section chair for IEEE Western Australia Section, and has also served as a member for the Emergent Technologies Technical Committee (ETTC) and Games Technical Committee (GTC) of the IEEE Computational Intelligence Society (CIS) in the past. His current research interests include Intelligent Data Mining and Data Science, Artificial Intelligence and Machine Learning, and Game and Virtual Reality Technology.

Applying Bayesian Regularization for Acceleration of Levenberg-Marquardt based Neural Network Training

Azizah Suliman¹, Batyrkhan S. Omarov^{1,2} *

¹ College of Computer Science & Information Technology, Universiti Tenaga Nasional, Kuala Lumpur (Malaysia)

² International Information Technologies University, Almaty (Kazakhstan)

Received 12 November 2017 | Accepted 29 March 2018 | Published 13 April 2018



ABSTRACT

Neural network is widely used for image classification problems, and is proven to be effective with high successful rate. However one of its main challenges is the significant amount of time it takes to train the network. The goal of this research is to improve the neural network training algorithms and apply and test them in classification and recognition problems. In this paper, we describe a method of applying Bayesian regularization to improve Levenberg-Marquardt (LM) algorithm and make it better usable in training neural networks. In the experimental part, we qualify the modified LM algorithm using Bayesian regularization and use it to determine an appropriate number of hidden layers in the network to avoid overtraining. The result of the experiment was very encouraging with a 98.8% correct classification when run on test samples.

KEYWORDS

Image Classification, Levenberg-Marquardt Method, Neural Network, Pattern Clustering.

DOI: 10.9781/ijimai.2018.04.004

I. INTRODUCTION

CURRENTLY the theories of artificial neural networks (ANN) are interdisciplinary in nature and it is one of the fastest growing disciplines which are used in various scientific and applied fields. ANN has been successfully applied in a wide range of applications - from household appliances to complex computer systems. Artificial neural network methods are also widely used in classification problems. A classification problem is a task to include the sample to one of several disjoint sets. When solving classification problems, ANN should include the existing object characteristics (observable data) to one or more specific classes.

However, there are a lot of problems that remain open in classification. For example, network topology selection problem, determining the number of hidden layers and neurons interpretation of weighting coefficients and bias, their evaluation of optimality, etc. The main idea for a neural network research was to develop mathematical and software tools for modeling processes of human thinking in solving various applied problems. Solving the recognition problems, using conventional methods with strict algorithms and limitations, is not possible. An automated system which solves such problems should not be programmed, it should be learnt. Thus, in our research, in order to recognize and classify we chose Levenberg-Marquardt algorithm.

The classical Levenberg-Marquardt algorithm copes poorly with the situation where the training set contains elements that stand out from the

general population. However such types of problems do occur in the practice tasks. In this study, we apply a modification in Levenberg-Marquardt algorithm on the basis of Bayesian regularization to make the algorithm more acceptable for practical tasks in recognition and classification, develop direct distributed neural network that trained modified LM algorithm and test it for recognition and classification problems.

II. RELATED WORKS

There are a lot of research to improve Levenberg-Marquardt algorithm to train neural networks, in last years, for example, (An Ru et. al., 2016) in [1] suggested improving the LM algorithm by direct calculation of quasi-Hessian and adopted gradient vector, that does not require storage of the Jacobian matrix. In own case, it causes to reduce the computation time of a neural network training. However this kind of improvement cannot decrease the testing error. Henri P. Gavin, uses LM method for nonlinear least squares curve-fitting problems [2]. In his research, he considers the Levenberg-Marquardt curve-fitting method as a combination of methods for minimization such as the Gauss-Newton method and the gradient descent method.

Thus, Levenberg-Marquardt method improvement comes from improvements of the named two methods. Gauss-Newton method assumes that the least square function is locally quadratic, and it causes to reduce the squared error. The LM method is more similar to the method of gradient descent, when its parameters are far from their optimal value. In contrast, the LM method is more similar to Gauss-Newton method, in case of its parameters are close to their optimum. So, in [2], Henri P. Gavin explains these methods and represents the applying of software to solve problems of approximating of nonlinear least squares. Also, Liyan

* Corresponding author.

E-mail address: batyahan@gmail.com

Qi proposed LM-method with self adjusting parameters for solving a nonlinear system of equations [3]. By their approach, at each step, the LM parameter μ_k is automatically adjusted on the basis of the correlation between actual reduction and predicted reduction. Thus, Under the BD-regular condition, they prove that PSA-LMM is locally superlinearly convergent, for semi-smooth equations and locally quadratically convergent for strongly semi-smooth equations. In [4], Haddout and Rhazi try solving the problem of non-linear least squares, based upon LM and Gauss–Newton methods by minimizing the sum of squares of errors between the data and model prediction.

Some researchers use Levenberg-Marquardt algorithm for problems related to prediction. In [5], Murat Kayri did a comparative analysis to Levenberg-Marquardt and Bayesian regularization algorithms from the point of view of predictive abilities.

Despite its effectiveness, the LM algorithm requires large computational cost. To provide the practical suitability, the method demands to diminish the working time for neural networks of large size.

In our research we propose our approaches to improve the method and illustrate experimental results to prove the improvement of the algorithm.

III. ARCHITECTURE

To provide image recognition and classification using neural network training we create a program complex. Flowchart of the program complex is presented in Fig. 1. Firstly, we create a mathematical model of the Levenberg-Marquardt algorithm and its improvement. After, using the developed mathematical model, we construct a neural network that, following, is trained and its performance is evaluated. After getting the result we compare the improved LM algorithm with classical LM method and the improvements of the other researches.

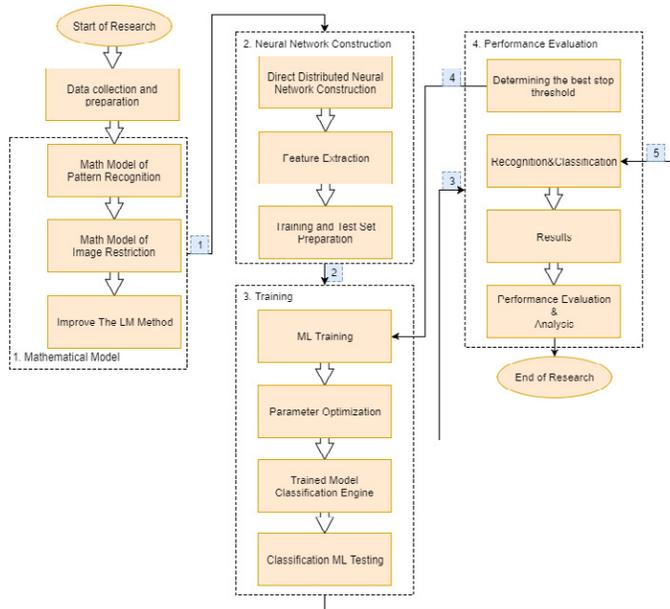


Fig. 1. Architecture of the Image Recognition and Classification Framework.

IV. BAYESIAN REGULARIZATION BASED IMPROVEMENT OF LM METHOD

The standard LM algorithm copes poorly with the situation where the training set contains elements that stand out from the general

population, usually such a situation arises when using data obtained by experiment. It means that, the algorithm is inconvenient for practical training. In order to provide the practical suitability, we propose Applying Bayesian Regularization for Acceleration of Levenberg-Marquardt based Neural Network Training. Moreover, applying Bayesian Regularization allows accelerating the LM based Neural Network Training. The essence of the approach is transition from searching the minimum point of mean square error to searching the minimum point of the function expressed by equation (1).

$$F(Y) = \alpha E_{\theta} + \beta E_D \quad (1)$$

Here E_D – network error, E_{θ} – the sum of the squares of the network weights, α and β – hyperparameters.

As a result, the algorithm seeks to minimize the network error and to prevent unlimited growth of its weights. The weight of the neural network can be considered as random variables and their distribution density can be expressed by the formula:

$$P(\theta | D, \alpha, \beta, M) = \frac{P(D | \theta, \beta, M)P(\theta | \alpha, M)}{P(D | \alpha, \beta, M)} \quad (2)$$

here D – training set, M – neural network model (in our case, it is feed forward neural network that learnt on the basis of Levenberg-Marquardt algorithm), θ – weight vector of the neural network. $P(\theta | \alpha, M)$ – priori probability, reflecting our knowledge of the initial weights of the network. $P(D | \theta, \beta, M)$ – likelihood function, which is the probability that the neural network with weights θ correctly responds to a set D . $P(D | \alpha, \beta, M)$ – normalization factor, that provides the equality of the total probability 1. If we assume the training set is noisy with Gaussian noise and the network weights distribution is Gaussian distribution, then formula (2) will be transformed as:

$$P(\theta | D, \alpha, \beta, M) = \frac{1}{Z_{\theta}(\alpha)} \frac{1}{Z_D(\beta)} e^{-(\alpha E_{\theta} + \beta E_D)} = \frac{e^{-F(\theta)}}{Z_F(\alpha, \beta)} \quad (3)$$

Here,

$$Z_D(\beta) = \left(\frac{\pi}{\beta}\right)^{\frac{p}{2}}, \quad Z_{\theta}(\alpha) = \left(\frac{\pi}{\alpha}\right)^{\frac{s}{2}} \quad (4)$$

Normalization coefficient can be expressed from the formula (5):

$$P(D | \alpha, \beta, M) = \frac{Z_F(\alpha, \beta)}{Z_{\theta}(\alpha)Z_D(\beta)} \quad (5)$$

Coefficient Z_F remains unknown. However it can be approximated by using the same assumptions as in the Levenberg-Marquardt method as equation (6):

$$Z_F \approx (2\pi)^{\frac{s}{2}} \sqrt{|H^{-1}|} e^{-F(\theta)} \quad (6)$$

here H^{-1} – inverse matrix to approximate Hessian matrix. Then

$$\alpha = \frac{\gamma}{2E_{\theta}}, \quad \beta = \frac{p - \gamma}{2E_D} \quad (7)$$

Here $\gamma = S - 2\alpha \text{trace}([H]^{-1})$, $\gamma \in [0, S]$ - parameter reflecting the number of weights of a neural network taking part in decreasing a function $\text{trace}(H^{-1})$ that is the sum of the diagonal elements of the inverse matrix to approximate Hessian matrix.

In this work to calculate hyperparameters α and β we use the formula proposed by Jan Poland in the work [6, 7]:

$$\begin{aligned} \gamma &= S - \alpha \text{trace}([H]^{-1}) \\ \alpha &= \frac{\gamma}{2E_w + \text{trace}([H]^{-1})} \\ \beta &= \frac{pN}{2E_d} \end{aligned} \quad (8)$$

As a result, neural network is less susceptible to fluctuation in the training set and accurately approximates the function that is given by the training set.

V. EXPERIMENT RESULTS

To test the proposed approach, three well known problems such as face recognition, lung cancer classification and object classification problem, have been applied. To test the problem the data set was divided into two parts, a training set and a test set. The training set data were 70% of instances of each species, and in the test were about 30% of instances data. Fig. 2 demonstrates samples of the each training dataset.

Face recognition experiments were carried out upon the basis of a facial images database by Dr. Libor Spacek [16]. The database contains face images of 395 different people, with 20 images of each person. The test sample consists of 20 images of each person (only $20 \times 395 = 7900$ images).

For image classification, in the presented research we use JSRT database [17] that includes 147 digital images of lung cancer image samples of two types (small-cell and non-small-cell types), with the 512x512 size.

For objective recognition as human, car, and other objects, we use Omnidirectional and panoramic digital images dataset. The dataset includes 30 omnidirectional images for human detection and 50 omnidirectional images for car detection [18].



Fig. 2. Samples of the training dataset.

In particular, the problem of choosing the number of hidden layer neurons was considered in [7-13, 19-21]. In [11] authors state that the optimal number of neurons in the hidden layer (N^h) should be calculated from the formula (9).

$$N^h = \sqrt{N^{(i)} N^{(o)}} \quad (9)$$

Here $N^{(i)}$ is number of input layer neurons, $N^{(o)}$ is the number of output layer neurons.

So, we should check this in practice. In this work, we carried out a practical study of the influence of the number of neurons in the hidden layer of the neural network in the learning rate and recognition quality. As a selection criteria of number of neurons, number of training epochs of neural network and recognition quality were chosen.

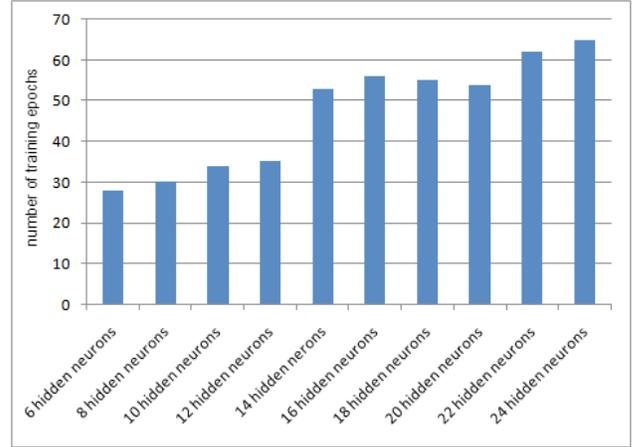


Fig. 3. Variable number of hidden neurons and learning rate.

Fig. 3 shows the ratio of neural network training epochs, when the number of neurons in the hidden layer varied from 6 to 24, in increments of 2 neurons. As it can be seen from the ratio, number of teaching epochs for the neural networks with different number of neurons in the hidden layer, the increase of this number reduces the speed of learning. As a result, not only the number of training epochs increase due to the growth of the Jacobian matrix of the neural network, also the total training time increases. However, this does not mean that the neural network with 6 neurons in the hidden layer will give the best results.

The graphs in Fig. 4 confirm that the minimization of the number of neurons in the hidden layer of the neural network does not improve the recognition quality. 9 and 15 neurons in the hidden layer give the best result. However, the number of training epochs in neural network with 18 hidden layers is substantially greater than others. Therefore, we assume that the best results in the “learning rate recognition quality” ratio are given by the neural network with 9 neurons in the hidden layer.

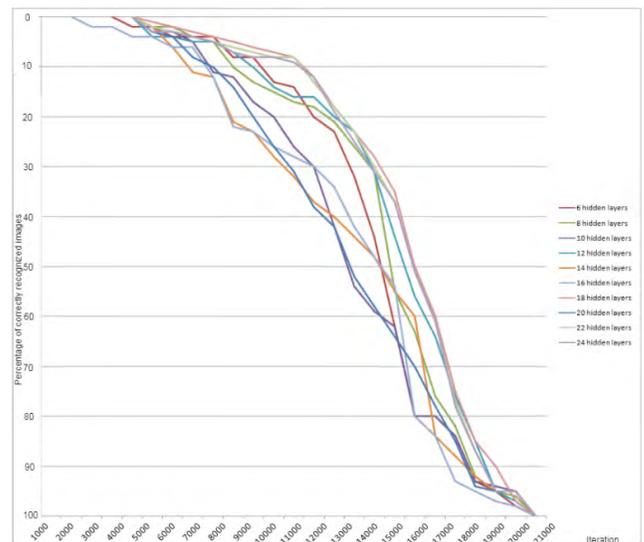


Fig. 4. Testing of neural networks with different number of neurons in the hidden layer.

For objective recognition as human, car, and other objects, we use Omnidirectional and panoramic digital images dataset. The dataset includes 30 omnidirectional images for human detection and 50 omnidirectional images for car detection [18].

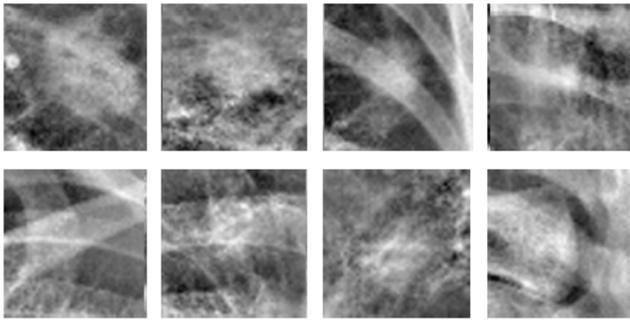


Fig. 5. Classification example of lung cancer image samples.

The results of the classification experiments consider correctly recognized training and test sets that were classified by using modified LM algorithm and standard LM algorithms, which are proposed by other researchers. As the experiment results show, modified Levenberg-Marquardt algorithm gives the highest training and test accuracy with the minimum sensitivity. Fig. 5 illustrates the classification of small-cell and non-small-cell lung cancer images of JSRT database. Table I illustrates the results of comparison with the other researchers.

TABLE I.
PERFORMANCE EVALUATION OF THE IMPROVEMENT IN THE LM ALGORITHM

Classifier	Percentage of Correctly Classified Training Samples%	Percentage of Correctly Classified Test Samples%	Train Validate to Test Partition %
The Proposed LM	99.3	98.8	60-40
Satish and Ritu, 2014 [11]	92.8	87.5	-
Aimi et. al., 2013 [12]	94.75	92.53	70-30
Devesh, 2014 [13]	84.36	-	-
Muhammad et. al, 2013 [14]	90.53	88.4	-
Abdel-Zaher et.al, 2016 [15]	99.03		54.76-45.24

VI. CONCLUSION

A mathematical model of pattern recognition using neural networks was proposed. Modification of Levenberg-Marquardt algorithm using Bayesian regularization was described and tested. Using the modified method feed forward neural network was constructed and tested for classification problems. The proposed method fulfills better in digital image classification, and maintains a good quality trade-off between the classification rate and sensitivity. The proposed method is effective from the point of view of computationally cost. Accordingly, the proposed method can be a valuable tool for medical data analysis and classification.

The predicted method better suits the classical problem, and also supports a trade-off between sensitivity and specificity. The proposed method is also effective in cooperation. Therefore, the proposed method can be a useful tool for classification in the field of data mining.

REFERENCES

- [1] AN Ru, LI Wen Jing, Han Hong Gui, QIAO Jun Fei. An Improved Levenberg-Marquardt Algorithm with Adaptive Learning Rate for RBF Neural Network. Proceedings of the 35th Chinese Control Conference July 27-29, 2016.
- [2] Henri P. Gavin. The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems. March 22, 2017.
- [3] Liyan Qi, Xiantao Xiao and Liwei Zhang. A Parameter-Self-Adjusting Levenberg-Marquardt Method For Solving Nonsmooth Equations. Journal of Computational Mathematics Vol.34, No.3, 2016, 317-338.
- [4] Soufiane Haddout and Mbarek Rhazi. Levenberg-Marquardt's and Gauss-Newton algorithms for parameter optimisation of the motion of a point mass rolling on a turntable. European Journal Of Computational Mechanics Vol. 24 , Iss. 6, 2015.
- [5] Murat Kayri. Predictive Abilities of Bayesian Regularization and Levenberg-Marquardt Algorithms in Artificial Neural Networks: A Comparative Empirical Study on Social Data. Mathematical and Computational Applications, May 2016.
- [6] Altayeva, A.B., Omarov, B.S., Aitmagambetov, A.Z., Kendzhaeva, B.B., Burkitbayeva, M.A.. Modeling and exploring base station characteristics of LTE mobile networks. Life Science Journal 11(6),31, pp. 227-233. 2014.
- [7] Poland J. On the Robustness of Update Strategies for the Bayesian Hyperparameter alpha. Jan Poland, November, 2001.
- [8] Omarov, B., Suliman, A., Kushibar, K. Face recognition using artificial neural networks in parallel architecture. Journal of Theoretical and Applied Information Technology 91 (2), pp. 238-248. (2016). Islamabad.
- [9] Omarov, B., Suliman, A., Tsoy, A. Parallel backpropagation neural network training for face recognition. Far East Journal of Electronics and Communications. Volume 16, Issue 4, December 2016, Pages 801-808. (2016).
- [10] A. Altayeva, B. Omarov, H.C. Jeong, Y.I. Cho. Multi-step face recognition for improving face detection and recognition rate. Far East Journal of Electronics and Communications 16(3), pp. 471-491.
- [11] Satish Saini, Ritu Vijay. 2014. Optimization of Artificial Neural Network Breast Cancer Detection System based on Image Registration Techniques. International Journal of Computer Applications (0975 – 8887) Volume 105 – No. 14, November 2014.
- [12] Aimi Abdul Nasir, Mohd Yusoff Mashor, and Rosline Hassan. 2013. Classification of Acute Leukaemia Cells using Multilayer Perceptron and Simplified Fuzzy ARTMAP Neural Networks. The International Arab Journal of Information Technology, Vol. 10, No. 4, July 2013.
- [13] Devesh Batra, 2014. Comparison Between Levenberg-Marquardt And Scaled Conjugate Gradient Training Algorithms For Image Compression Using MLP . International Journal of Image Processing (IJIP), Volume (8) : Issue (6) : 2014.
- [14] Muhammad Ibn Ibrahimy, Md. Rezwanul Ahsan, Othman Omran Khalifa. 2013. Design and Optimization of Levenberg-Marquardt based Neural Network Classifier for EMG Signals to Identify Hand Motions. MEASUREMENT SCIENCE REVIEW, Volume 13, No. 3, 2013.
- [15] Abdel-Zaher Ahmed, Eldeib Ayman. 2016. Breast cancer classification using deep belief networks. Expert Systems With Applications, 46 (2016) 139-144.
- [16] Dr Libor Spacek. 2016. Collection of Facial Images. Designed and maintained by Dr Libor Spacek. Electronic resource. <http://cswww.essex.ac.uk/mv/allfaces/index.html>, June 2016.
- [17] JSRT Database. 2016. Public Interest Incorporated Association Japanese Society of Radiological Technology. Electronic resource. <http://www.jsrt.or.jp/jsrt-db/eng.php>, 2016.
- [18] Pinz A. 2016. Human, car, other object database, Electronic resource, <http://cvrg.iyte.edu.tr/datasets.htm>, 2016.
- [19] Bouldid, Y., A. Souhar, and M. E. Elkettani, "Handwritten Character Recognition Based on the Specificity and the Singularity of the Arabic Language", International Journal of Interactive Multimedia and Artificial

Intelligence, vol. 4, issue Regular Issue, no. 4, pp. 45-53, 2017.

- [20] Harish, B. S., and S. V. A. Kumar, "Anomaly based Intrusion Detection using Modified Fuzzy Clustering", *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, issue Regular Issue, no. 6, pp. 54-59, 2017.
- [21] Reguieg, S., and N. Taghezout, "Supporting Multi-agent Coordination and Computational Collective Intelligence in Enterprise 2.0 Platform", *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, issue Regular Issue, no. 6, pp. 70-80, 2017.



Azizah Suliman

Azizah Suliman is an Associate Professor at College of Information Technology, Universiti Tenaga Nasional, Malaysia. She received her BCS from Southern Illinois University, USA in 1985, MCS from Universiti Malaya in 1990 and PhD in Computer Science (AI) from Universiti Putra Malaysia in 2011. Her interests includes

Soft Computing, Parallel and GPU Programming, Image Processing and Embedded Systems.



Batyrkhan S. Omarov

Batyrkhan S. Omarov is an Assistant Professor at International Information Technologies University, Almaty, Kazakhstan. He received his BCS and MSC from Al-Farabi Kazakh National University, Almaty, Kazakhstan in 2008 and 2010 respectively. From 2014 till 2017 did PhD research at Unversiti Tenaga Nasional, Kuala Lumpur, Malaysia. Research Interests include Computer Vision,

Machine Learning and Data Science.

Object Detection and Tracking using Modified Diamond Search Block Matching Motion Estimation Algorithm

Apurva S. Samdurkar¹, Shailesh D. Kamble², Nileshsingh V. Thakur³, Akshay S. Patharkar⁴

¹ PG Scholar, Computer Science & Engineering, Yeshwantrao Chavan College of Engineering (India)

² Computer Science & Engineering, Yeshwantrao Chavan College of Engineering (India)

³ Computer Science & Engineering, Prof Ram Meghe College of Engineering & Management (India)

⁴ Computer Technology, K.D.K. College of Engineering (India)

Received 5 June 2017 | Accepted 13 September 2017 | Published 30 October 2017



ABSTRACT

Object tracking is one of the main fields within computer vision. Amongst various methods/ approaches for object detection and tracking, the background subtraction approach makes the detection of object easier. To the detected object, apply the proposed block matching algorithm for generating the motion vectors. The existing diamond search (DS) and cross diamond search algorithms (CDS) are studied and experiments are carried out on various standard video data sets and user defined data sets. Based on the study and analysis of these two existing algorithms a modified diamond search pattern (MDS) algorithm is proposed using small diamond shape search pattern in initial step and large diamond shape (LDS) in further steps for motion estimation. The initial search pattern consists of five points in small diamond shape pattern and gradually grows into a large diamond shape pattern, based on the point with minimum cost function. The algorithm ends with the small shape pattern at last. The proposed MDS algorithm finds the smaller motion vectors and fewer searching points than the existing DS and CDS algorithms. Further, object detection is carried out by using background subtraction approach and finally, MDS motion estimation algorithm is used for tracking the object in color video sequences. The experiments are carried out by using different video data sets containing a single object. The results are evaluated and compared by using the evaluation parameters like average searching points per frame and average computational time per frame. The experimental results show that the MDS performs better than DS and CDS on average search point and average computation time.

KEYWORDS

Background Subtraction, Block Matching Algorithm, Motion Estimation, Cross Diamond Search Algorithm, Diamond Search Algorithm.

DOI: 10.9781/ijimai.2017.10.004

I. INTRODUCTION

FROM the viewpoint of tracking in videos, the use of block matching algorithm can be made for the purpose of motion estimation and object tracking. Detection of an object is typically the first step when starting the tracking process. The object detection mechanism is carried out either in the sequence of frames or when the object first appears in the frame of the video for tracking [1, 2]. Background subtraction is a process of subtracting the second frame (current frame) from the first background (reference) frame, thus the dissimilarity between two frames and the position of moving object can be obtained [3, 4]. Object tracking means the process of locating the object of interest from a video sequence. The object is tracked by continuously monitoring the motion of an object in the video. Each of the frames can be divided into two set of objects, foreground and background objects [5]. An adaptive background subtraction technique can also be used for object detection [6]. The complete process of object tracking can be categorized into

the following three steps, object detection, object classification and object tracking [7]. Many application of tacking object using different approaches and techniques are implemented. Video surveillance is the basic and widely used application of object tracking [8]. The block matching algorithm is used for the purpose of tracking an object. The algorithms for motion estimation by block matching are used where similar blocks in a sequence of frames of the video are located for the purposes of motion vector estimation. Motion estimation is the process of finding the motion vectors from all frames in a video sequence [9]. The block matching algorithm aims at finding a matching block from a frame in some other frame. Block matching involves partitioning the current frame into a number of macro blocks and compares each macro block with the corresponding macro block. A vector is created that maps the movement of a macro block from one location to another [10]. These motion vectors provide the displacement in the block which can be used for object tracking. Block matching technique also helps in removing the redundancy in the frames. This process can also be carried out in real-time environment along with some hardware [11, 12]. The matching process is performed by minimizing a certain matching criterion (Mean of Absolute Difference), and the best-matched block, where the motion vector is found, which gives the minimum block-matching distortion (BMD) [13]. There are different approaches of block matching algorithm (BMA) amongst which some are advanced algorithms of the already

* Corresponding author.

E-mail addresses: aapurva2009@gmail.com (A. S. Samdurkar), shailesh_2kin@rediffmail.com (S. D. Kamble), thakurnisvis@rediffmail.com (N. V. Thakur), akshay.patharkar7@gmail.com (A. S. Patharkar).

existing BMAs [14]: Three step search (TSS); novel four step search [15], New three step search (NTSS) [16]; Simple and efficient search; Four step search (FSS) [17]; Diamond Search (DS) [18] & Cross Diamond search (CDS) [19] etc. In this paper, the modified diamond search algorithm (MDS) is presented by introducing the capabilities of DS and CDS. It starts by initializing the small cross shape in the first step, which grows into the large diamond shape (LDSP). Estimation of Motion vector (MV) by reducing the size of the blocks can also make the tracking process fast [20]. The various applications of block matching algorithm other than tracking are used in video compression [21], block matching motion estimation using automata theory for fractal coding [22,23] and also modified Three-Step Search (TSS) algorithm using fractal coding [24]. The various sections of this paper are organized as follows: Algorithm and analysis of DS and CDS are given in Section II. Section III describes the Proposed MDS algorithm. Section IV reports the significant experimental results. The conclusion and future scope is given in Section V.

II. DS AND CDS ALGORITHM

From the entire existing block matching algorithms the behavior of diamond search and cross diamond search are studied. These two algorithms are selected for study because these algorithms require less searching points than the other algorithms as seen from the studied literature survey. The DS and CDS algorithms are explained in detail below with its steps.

A. Diamond Search Algorithm

The steps of DS algorithm with different cases are explained below.

Step 1: Start at the center by initializing 9 search points on the search window. Check all the 9 search points, if the minimum cost found to be at the center of the search window then go to Step 3, else go to Step 2.

Step 2: The point marked with minimum cost in the Step 1 is re-set as the new center by forming LDSP around it. If the newly obtained minimum cost point is located at the center position then go to Step 3, else repeat Step 2.

Step 3: Shift the searching pattern from LDSP to Small Diamond Search Patter (SDSP). The minimum cost point found in Step 3 is said to be the final solution for the motion vector, which points to the best matching block.

Diamond search pattern has different cases depending upon the point of concern. The presence of the least cost point at different location is processed differently as shown in Fig. 1.

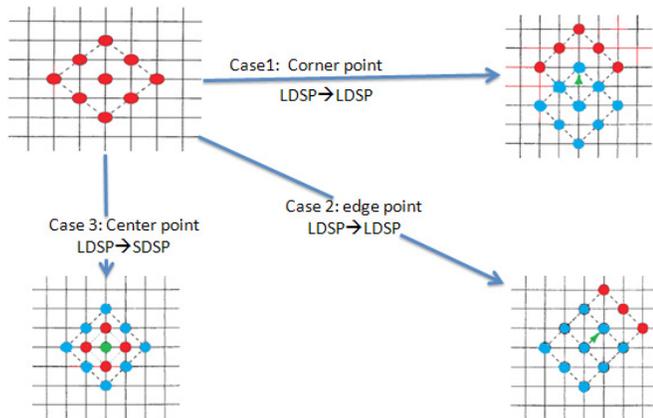


Fig. 1. Search Pattern of Diamond Search Algorithm.

B. Cross-diamond Search Algorithm

Cross diamond shaped search (CDS) pattern implies a cross search pattern (CSP) over the diamond search (DS) pattern as shown in Fig. 2.

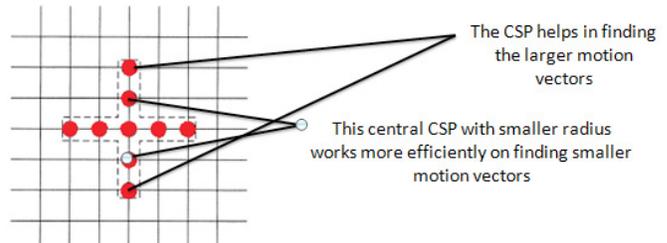


Fig. 2. Cross (+) Shaped Search Pattern.

The CDS Algorithm works as follows:

Step 1: The minimum cost of each of the 9 search points of the CSP is found. If the point with minimum cost is found at the center of the CSP then stop the search as shown in Fig. 3(a), else go to Step 2.

Step 2: 2 more closest search points to the current minimum cost point of the central CSP are checked. If the minimum cost point from the previous step 1 is located at the center of the CSP and if the new minimum cost point found in this step coincides with this point then, stop the search as in Fig. 3(b) else go to Step 3.

Step 3: The point with minimum cost found in previous Step 2 is the new center of the LDSP. If the new point with minimum is still at the center of the newly formed LDSP then go to step 4. Else repeat this step.

Step 4: A new SDSP is formed with the minimum cost point in the previous step as the center. The new minimum cost point, in this step, is the final solution to motion vector.

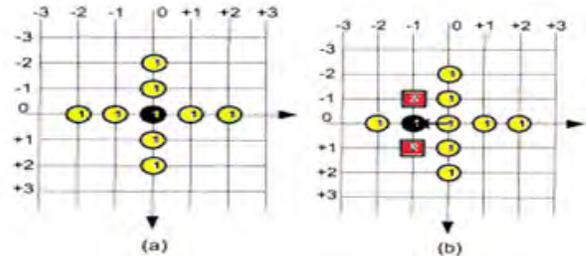


Fig. 3. (a) First step stop (b) Second step stop

For an in-depth analysis of the block matching motion estimation (BMME) algorithm, the experiments are carried out using the existing block matching algorithm: Diamond search (DS) and Cross diamond search (CDS). The experiments are performed on eight video sequences listed in Table I, among which the first four are the well-known sequences such as “Traffic”, “Pet”, “Walk”, “Ant”, and the remaining four “Bottle1”, “Bottle2”, “Walk1”, “Walk2” are the user defined video data sets. All these data sets consist of only a single object moving throughout the frames. Also for all the video data sets used, the camera is stationary. The video sequences such as “Traffic”, “Walk”, “Walk1” involves higher motion than other videos. The other remaining video sequences involve gentle or slower motion. The DS and CDS algorithms are performed on a 16x16 macro block, with a window size (w) of -7 to +7. The diamond search algorithm starts by initializing 9 points at the center of the search window. It then carries out the steps according to the algorithm and finds out the motion vectors. The motion vectors are found using the block matching criterion such as MAD. The Cross diamond search algorithm also starts by initializing 9 points in the first step, but it differs from the diamond search, in the shape of searching points initialized. The

results, after performing DS and CDS on the eight video sequences, the searching points per frame and computational time per frame are tabulated as required by the DS and CDS algorithms, given in Table II and Table III.

TABLE I. VIDEO SEQUENCE WITH SPECIFICATION USED FOR EXPERIMENTATION

Sr. No	Name of video sequence	No. of frames	Frame Rate	Frame Height	Frame Width
1	Traffic	120	15	120	160
2	Pet	520	25	288	352
3	Walk	113	25	120	160
4	Ant	430	8	288	352
5	Bottle1	123	15.08	240	320
6	Bottle2	76	15.08	240	320
7	Walk1	220	30	320	240
8	Walk2	180	30	240	240

The diamond search pattern starts by initializing nine points, arranged in a diamond shape pattern, on 5x5 grids, to cover the search points in both the directions (up and down). The diamond search pattern may require 13 search points at its best case and 30 search points in the worst case. On an average, it may range between 13 to 30 search points. Table II are the results of the searching points and computational time computed, using the diamond search algorithm, on the video sequences. In the same way, the cross diamond search algorithm also requires nine points in the first step but the search points are arranged such that it forms a cross like structure (+). In the best case, CDS may require 9 points and 29 search points in the worst case. On an average, it may range between 9 to 29 search points. Table III are the results of the searching points and computational time computed, using the Cross Diamond search algorithm, on the video sequences. The experiments are performed on the following sequences of video, "Traffic", "Pet", "Walk", "Ant", "Bottle1", "Bottle2", "Walk1", "Walk2". The first 10 frames of each of these video sequences are considered for analysis. The obtained results are studied and analyzed for further designing of modified diamond search algorithm. Fig. 4 and Fig. 5 represent the graph of the obtained results in terms of searching points and computational time, for DS and CDS, respectively. The main aim is to design an algorithm that could match or to improve the performance of the DS and CDS, in terms of the search points required and speed up the tracking process.

TABLE II. COMPARISON OF ALL VIDEOS IN TERMS OF SEARCHING POINTS PER FRAME AND COMPUTATIONAL TIME PER FRAME USING DIAMOND SEARCH ALGORITHM

Diamond Search Algorithm																
Fr. No	Search Points per Frame								Computational Time per Frame							
	Traffic	Pet	Walk	Ant	Bottle1	Bottle2	Walk1	Walk2	Traffic	Pet	Walk	Ant	Bottle1	Bottle2	Walk1	Walk2
2	11.56	12.52	13.74	13.59	13.93	13.92	14.13	13.65	8.75	49.46	8.59	47.79	36.36	36.2	36.53	36.52
3	11.77	12.52	14.67	13.48	12.65	12.87	13.76	13.57	8.62	50.24	8.54	48.23	36.08	36.38	36.32	36.37
4	11.7	12.43	16.62	12.93	12.7	12.8	14.06	13.46	8.74	51.04	8.56	48.01	36.08	36.15	36.35	36.39
5	11.65	12.41	15.25	13.39	13.18	12.89	13.65	13.7	8.81	48.32	8.54	48.72	36.64	36.15	36.28	36.44
6	11.65	12.39	14.12	12.92	12.75	12.53	13.75	13.24	8.99	48.1	8.59	48.7	37.44	36.09	36.31	36.39
7	11.65	12.36	14.32	13.34	12.51	12.59	13.61	13.49	8.59	47.61	8.53	48.69	40.25	36.08	36.38	36.46
8	11.7	12.45	12.94	13.01	12.76	12.63	13.95	13.74	8.81	47.92	8.61	47.98	43.06	36.14	36.43	36.43
9	12.17	12.39	15.6	13.28	12.86	12.59	13.55	13.34	8.73	47.64	8.58	47.83	42.92	36.28	36.32	36.39
10	8.73	12.41	15.18	13.09	12.89	12.77	13.81	13.19	8.5	47.74	8.65	47.76	36.5	37.03	36.34	36.39

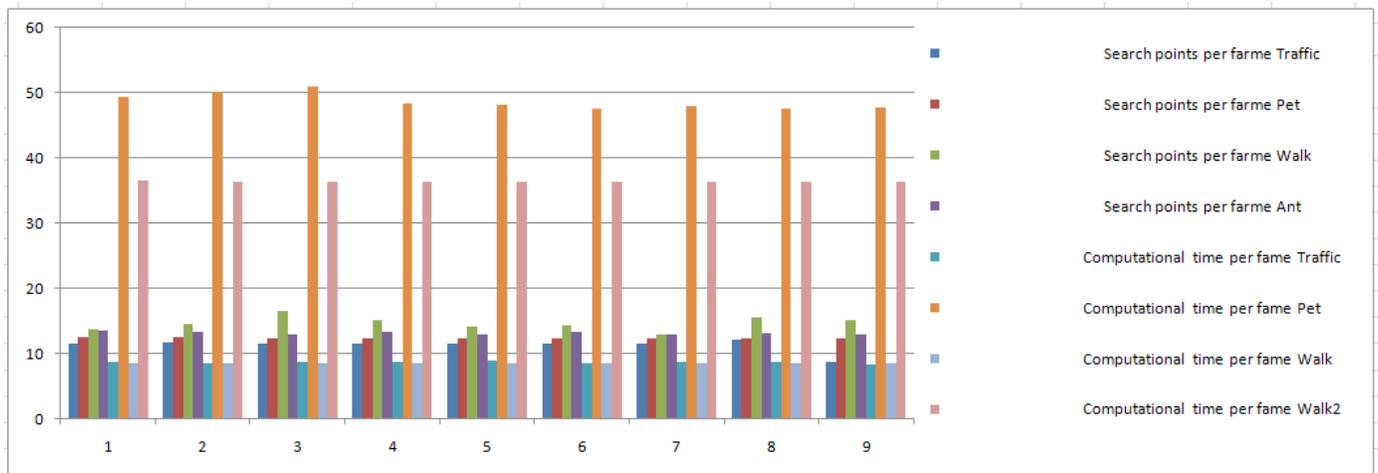


Fig. 4. Graphical representation of videos in terms of search points/frame and computational time/frame of DS algorithm.

TABLE III. COMPARISON OF ALL VIDEOS IN TERMS OF SEARCHING POINTS PER FRAME AND COMPUTATIONAL TIME PER FRAME USING CROSS DIAMOND SEARCH ALGORITHM

Cross-Diamond Search Algorithm																
Fr. No	Search Points per frame								Computational time per frame							
	Traffic	Pet	Walk	Ant	Bottle1	Bottle2	Walk1	Walk2	Traffic	Pet	Walk	Ant	Bottle1	Bottle2	Walk1	Walk2
2	8.35	9.54	12.67	11.65	12.6	12.65	12.49	11.79	7.25	40.25	7.23	39.78	34.78	32.01	31.2	31.16
3	8.4	9.48	14.05	11.38	10.46	10.64	12.56	11.74	7.66	39.65	7.33	39.87	29.95	29.22	31.26	30.96
4	8.45	9.36	17.87	10.44	10.47	10.48	12.72	11.39	7.24	39.44	8.32	39.1	29.26	29.38	30.75	30.24
5	8.31	9.37	15.41	11.31	11.52	10.44	11.97	11.9	8.1	38.96	8.15	39.41	30.93	29.12	30.26	31.21
6	8.38	9.74	14.2	10.47	10.42	9.68	11.83	10.9	7.99	43.26	7.86	39.64	29.2	28.92	30.41	29.96
7	8.31	9.08	13.39	11.18	9.73	9.71	11.54	11.16	11.46	43.45	7.91	39.45	28.4	28.33	29.82	29.82
8	8.45	9.25	12.91	10.58	10.6	9.8	12.2	11.91	8.52	40.16	7.61	38.96	29.48	28.37	29.87	30.57
9	8.94	9.27	15.27	11.25	10.81	9.73	11.44	10.81	7.2	41.45	8.26	39.37	29.42	28.35	29.83	29.58
10	8.68	9.16	15.5	10.63	10.72	10.1	12.12	10.86	7.79	45.41	8.28	38.72	29.54	28.57	30.22	29.88

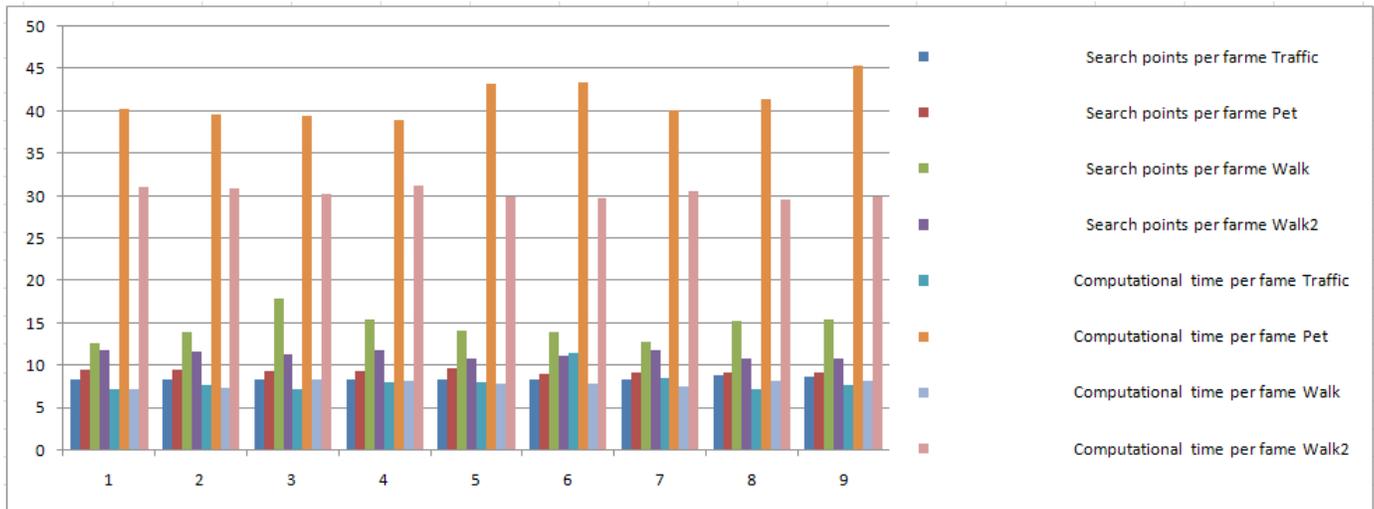


Fig. 5. Graphical representation of videos in terms of search points/frame and computational time/frame of CDS algorithm.

III. PROPOSED WORK

The proposed work is concerned with designing an algorithm that could match or improve the performance of the DS and CDS, in terms of the search points required and speed up the tracking process. The tracking process can be made fast if the search points required by the proposed algorithm can be reduced.

Flowchart of the object tracking system is shown in Fig. 6. That helps in visualization of the stepwise working of the proposed approach. It represents the process for single object tracking using block matching motion estimation techniques.

- As a preprocessing step video is converted into YUV color space, where only the “Y” component (luminance) is considered as it is sensitive to the human eye.
- Later the video is divided into many frames. Only two frames are considered at a time of processing. From these two frames one is current frame and the other is reference frame.
- On all the frames of the video sequence the background subtraction method for object detection is applied. In background subtraction,

the current frame gets subtracted from the reference frame. The result after subtraction is the object in the foreground.

- After detecting the object using background subtraction, the Block Matching Algorithm (BMA) is applied.
- BMA divides the block into number of macroblocks.
- BMA calculates the motion on a block by block basis. For every current frame block, a block from previous frame is found and matching is based on certain criteria. (MAD here to find best match).
- A vector is created that keeps track of the movement of a macro block within different locations. Motion vector provides displacement in the block in terms of MV_x and MV_y.
- Using all these motion vectors MV_x and MV_y the object is tracked from the video.

A. Diamond Search Pattern

The diamond search algorithm uses a Large Diamond Shape Pattern (LDSP) and a Small Diamond Shape pattern (SDSP) as shown in Fig. 7. The Fig. 7 (b) is used as the initial step to the Modified Diamond Search (MDS) algorithm, followed by the Fig. 7 (a).

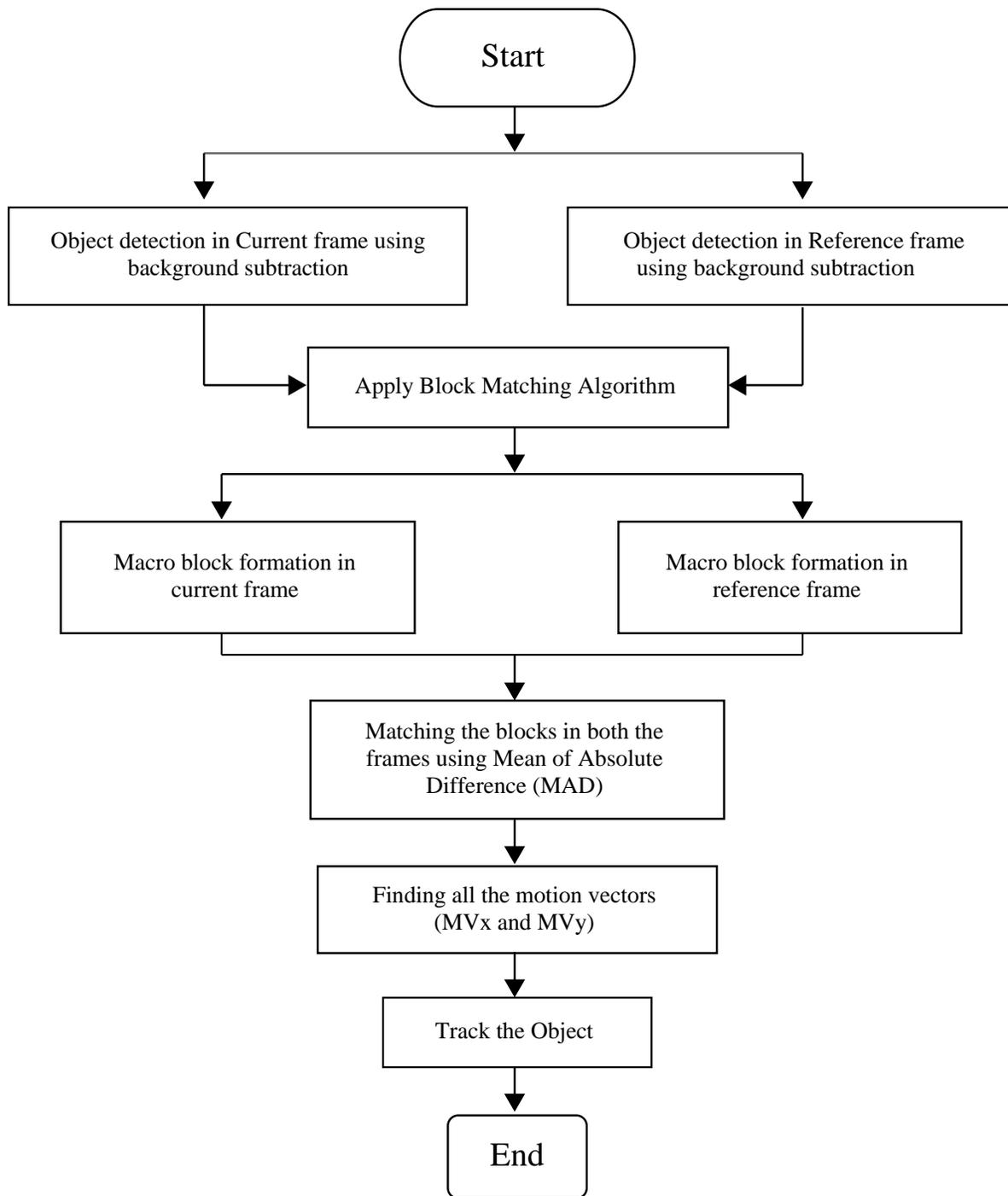


Fig. 6. Flow of object tracking system.

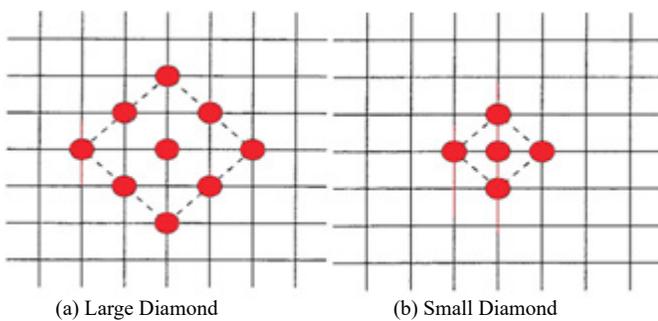


Fig. 7. Search Pattern used in MDS Algorithm.

1) The Modified Diamond Search Algorithm

Following are the steps of MDS algorithm:

Step 1: Start with five points of SDSP as the initial step from the center of the search window and the 5 checking points of SDSP, are tested. If the cost with minimum point after calculation is located at the center position then stop the search. Otherwise, go to Step 2.

Step 2: The minimum cost point found in the Step 1 is re-positioned as the center point and forms a new LDSP. If the new point with minimum cost obtained is located at the center position then go to Step 3; else repeat this step in recursion until the minimum point occurs at the center of the LDSP and then go to Step 3.

Step 3: With the new minimum point as the center, shift the search pattern, from LDSP to SDSP. The minimum cost point, found in this step, is the final solution of the motion vector, which is the best matching block.

Fig. 8 shows the flowchart of the MDS algorithm and the various cases of MDS algorithm are shown in Fig. 9.

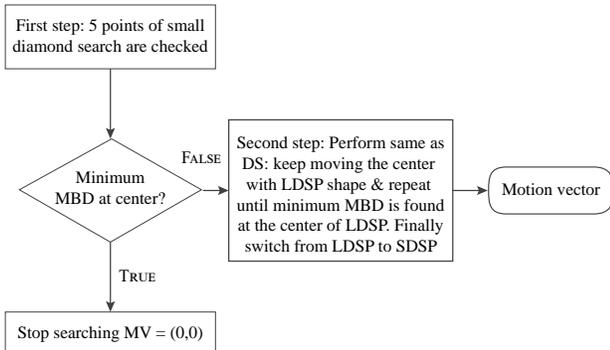


Fig. 8. Flowchart of the MDS algorithm.

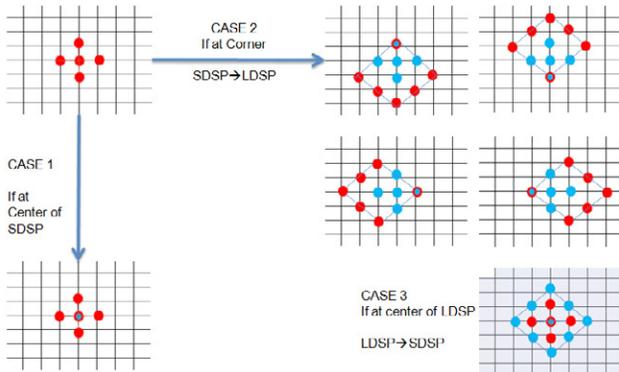


Fig. 9. Different cases of searching pattern in MDS algorithm.

B. Analysis of MDS Algorithm

Fig. 9 shows the various cases of MDS algorithm, in which the modified diamond search algorithm starts, by initializing only five

points in the first step in a 3x3 window and nine points in the second step on the 5x5 search window respectively, when compared to the nine searching points used in the first step of DS and CDS. With this characteristic, it can facilitate the optimization of computation of MDS over DS and CDS. The design of MDS also helps in finding smaller motion vectors more efficiently. MDS algorithm behaves like DS only if the min MBD is not found at the center in the first step and keeps on modifying between successive LDSP, by three points or five points for searching. Thus the total number of search points, in the best case, are 5 and varies from 5 to (5+8) =13, in the worst case, whereas, on an average, it requires search points in between 5 and 13.

The MDS algorithm is quite different from any other fast BMAs. i.e. (i) The search patterns used in MDS have the minimum number of points; (ii) the directional search patterns are used, and (iii) the switching strategy of the diamond search patterns is adopted in the last stage. MDS starts with a small searching pattern and then grows into a large shape and again shrinks to the small shape, replacing the methods, in cross search pattern and simple diamond pattern.

IV. EXPERIMENTAL RESULTS

A. Experimental Results on MDS

The proposed MDS algorithm is simulated and tested using the various standard video sequences and use created video sequences consisting of different motion types. For all the video sequences we have used the Mean of Absolute Difference (MAD) as the block matching measure (criterion), block size of 16x16, and search window of size 7 (i.e. $w = -7$ to $+7$). All these video sequences used for the experiment, consist of only a single object moving throughout the frames. Also for all the video data sets used, the camera is stationary. The video sequences, such as “Traffic”, “Walk”, “Walk1”, involves higher motion than other videos. The other remaining video sequences involve a slow moving object. The results of the Modified diamond search algorithm are listed in Table IV.

The results in Table IV are graphically represented in Fig. 10. A lot of improvement is seen in the result of MDS over the DS and CDS. The modified algorithm for diamond search is compared against the existing diamond search and cross diamond search algorithm, in the aspect of searching points.

TABLE IV. COMPARISON OF ALL VIDEOS IN TERMS OF SEARCHING POINTS PER FRAME AND COMPUTATIONAL TIME PER FRAME USING MDS

Modified Diamond Search Algorithm																
Search points per frame									Computational time per frame							
Fr. No	Traffic	Pet	Walk	Ant	Bottle1	Bottle2	Walk1	Walk2	Traffic	Pet	Walk	Ant	Bottle1	Bottle2	Walk1	Walk2
2	8.35	9.12	10.04	10.07	10.83	11	10.56	10.5	8.57	48.39	8.68	49.36	36.76	36.51	36.41	36.44
3	8.41	9.08	10.31	9.98	9.84	9.79	10.53	10.37	8.57	48.04	8.60	89.16	35.99	36.44	36.35	36.75
4	8.41	9.05	11.6	9.57	9.96	9.7	10.46	10.04	8.53	47.94	8.68	49.02	36.01	36.92	36.3	36.43
5	8.31	9.07	11.64	9.84	10.62	9.75	10.21	10.25	8.56	47.93	8.68	48.23	36.03	36.43	36.41	36.38
6	8.34	8.86	10.98	9.56	9.91	9.21	10.18	9.73	8.58	48.24	8.64	48.75	36.63	36.67	36.4	36.38
7	8.31	8.84	11.24	9.87	9.4	9.24	9.91	9.69	8.54	47.87	8.73	48.96	36.05	36.32	36.39	36.47
8	8.34	9.02	10.67	9.63	9.84	9.24	9.88	9.99	8.68	48.10	8.68	49.01	36.04	36.25	36.41	36.42
9	8.48	8.94	11.64	9.89	9.87	9.28	9.88	9.7	8.57	48.04	8.69	48.69	36.13	36.43	36.31	36.36
10	8.44	8.92	11.67	9.61	9.91	9.49	10.15	9.73	8.56	47.85	8.58	48.69	36	36.2	36.37	36.54

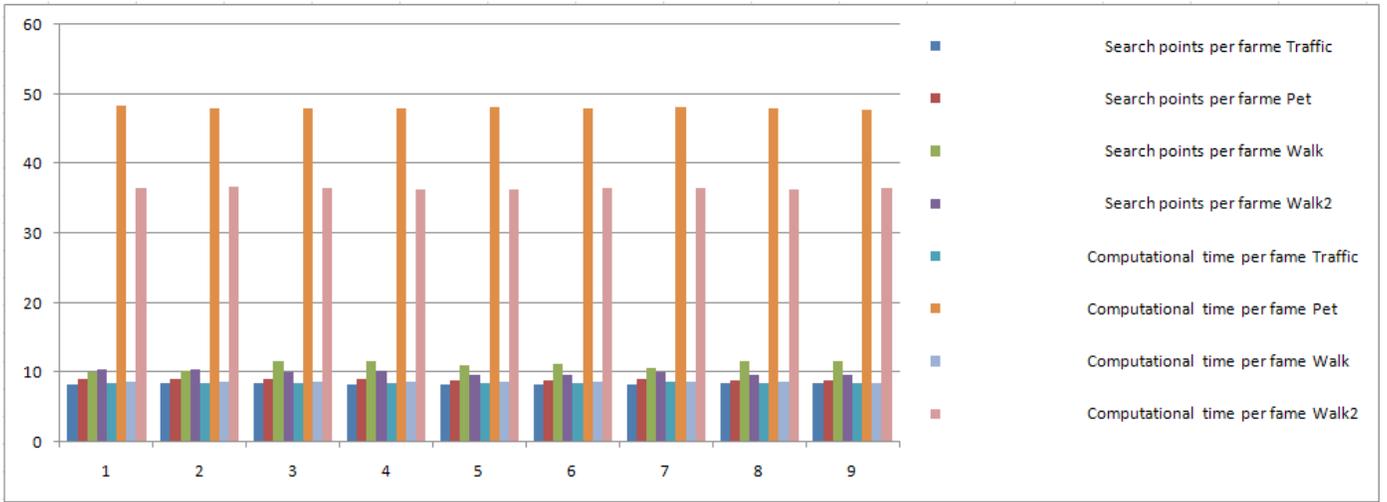


Fig. 10. Graphical representation of videos in terms of search points/frame and computational time/frame of MDS algorithm.

Table V shows the comparison of MDS with DS and CDS algorithm.

TABLE V. PERFORMANCE COMPARISONS OF DS, CDS AND MDS

Using sequence "Traffic"			
BMA	DS	CDS	PDS
Avg. SP	11.4	8.47	8.38
Using sequence "Pet"			
BMA	DS	CDS	PDS
Avg. SP	12.43	9.36	8.99
Using sequence "Walk"			
BMA	DS	CDS	PDS
Avg. SP	14.71	14.63	11.09
Using sequence "Ant"			
BMA	DS	CDS	PDS
Avg. SP	13.22	10.98	9.78
Using sequence "Bottle1"			
BMA	DS	CDS	PDS
Avg. SP	12.91	10.81	10.02
Using sequence "Bottle2"			
BMA	DS	CDS	PDS
Avg. SP	12.84	10.35	9.63
Using sequence "Walk1"			
BMA	DS	CDS	PDS
Avg. SP	13.8	12.07	10.19
Using sequence "Walk2"			
BMA	DS	CDS	PDS
Avg. SP	13.48	11.38	10

The Table V compares the searching points among different block matching algorithms. It shows that the Modified diamond search algorithm (MDS) always consumes the smaller number of search points, as compared to the Diamond Search (DS) and Cross Diamond Search (CDS). Compared to DS, CDS approximately saves about 1.98 – 2.24 search points. When CDS is compared to MDS, it also saves 0.36 – 3.55 search/check points, which is much more than the other two BMAs. The average search points, per frame with observations, $MDS < CDS < DS$, is detected for the video sequences, with $w = +7$ to -7 . With such an improvement in the searching point, the experimental results, using the video sequence, shows that the MDS algorithm achieves less searching points, than that of DS

and CDS. For the video sequence "Traffic", the MDS algorithm saves up to 0.36 and even 3.02 search/check points per frame, as compared to CDS and DS respectively. For sequence "Pet" the proposed algorithm saves 0.37 to 3.44 search points, "Walk" saves 3.55 to 3.62 searching points, "Ant" search points ranges from 1.2 to 3.44, "Bottle1" sequence saves 0.79 to 2.89 searching points, video sequence "Bottle2" saves 0.72 to 3.21 search points, and the video sequences "Walk1" and "Walk2" saves 1.88 to 3.61 and 1.38 to 3.48 searching points per frame. This implies that the searching pattern of proposed Diamond Search (MDS) algorithm, with a search area of 3×3 , in the first step, starts with 5 searching points, instead of with 9 search points as in DS and CDS. This feature outperforms the existing Diamond search and Cross Diamond search that use a larger search area and more search/check points in the first step. The smaller shape of MDS helps in finding smaller Motion Vector (MV) efficiently. MDS starts with a small searching pattern and then grows into a large shape and again shrinks to the small shape, replacing the methods in cross search pattern and simple diamond pattern. Thus, MDS performs better than DS and CDS. Fig. 11 and Fig. 12 plots the average searching points per frame using sequence "Traffic" and "Bottle1" respectively, on each frame. These two figures clearly show the superior performance of MDS over DS and CDS algorithm, in terms of the search points used. Fig. 13 shows the performance comparison of DS, CDS, and MDS in terms of average search points.

B. Experiment Results using MDS for Object Tracking

This research has applied the capabilities of Modified diamond search (MDS) block matching algorithm for the purpose of tracking an object. The block matching motion estimation algorithm is used, where similar blocks in a sequence of frames of the video are located, for the purposes of motion vector estimation. The purpose of a block matching algorithm is to find a matching block, from a frame, in some other frame. Block matching involves partitioning the current frame into a number of macro blocks and compares each macro block, with the corresponding block. A vector is created that maps the movement of a macro block, from one location to another. These motion vectors provide the displacement in the block, which can be used for object tracking. In the preprocessing, the RGB video sequence is converted into YUV space and all the experimentation is done, using the luminance of video sequence. Thus this research performs background subtraction for detection of the object and then applies the MDS algorithm to it, where the motion vectors are calculated and the object is tracked. Fig.14 shows the step wise results of the tracking. Final tracked object is shown by the blue bounding box.

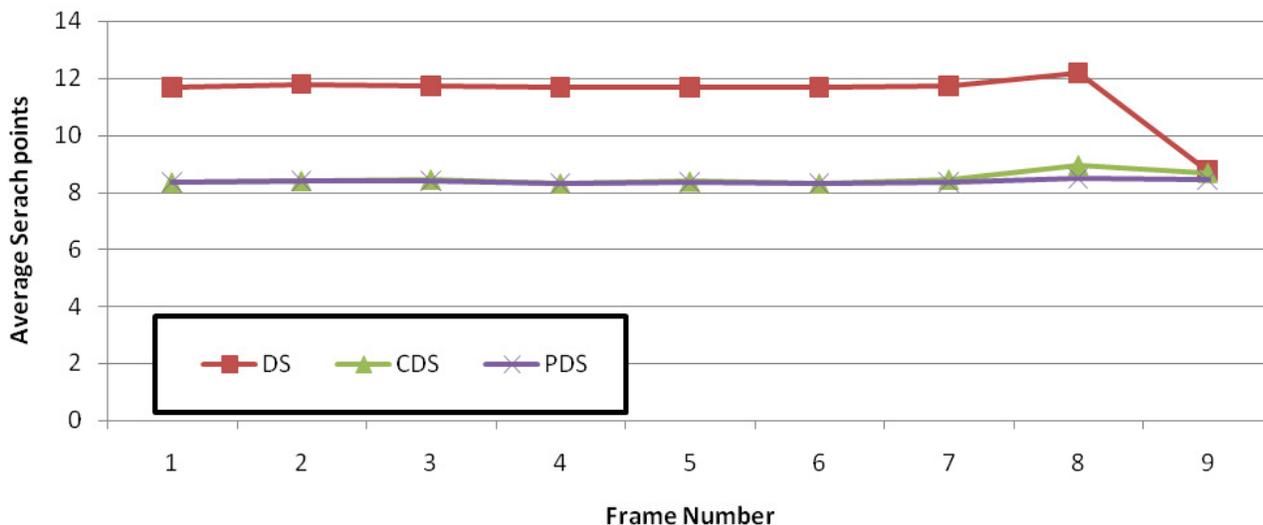


Fig. 11. Frame-wise performance comparison between DS, CDS and MDS on “Traffic” sequence by average search point.

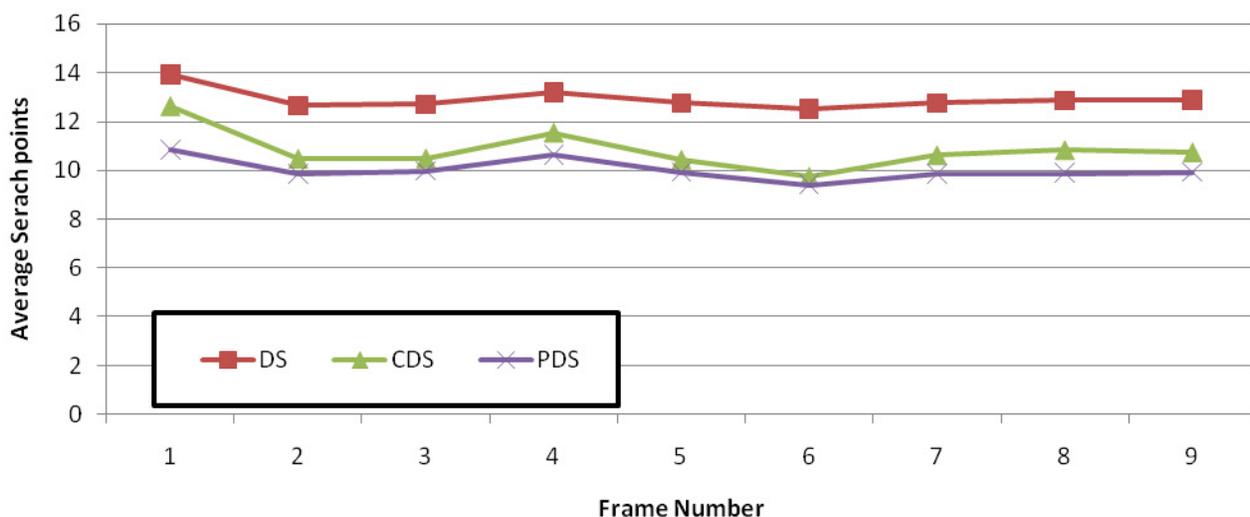


Fig. 12. Frame-wise performance comparison between DS, CDS and MDS on “Bottle1” sequence by average search points.

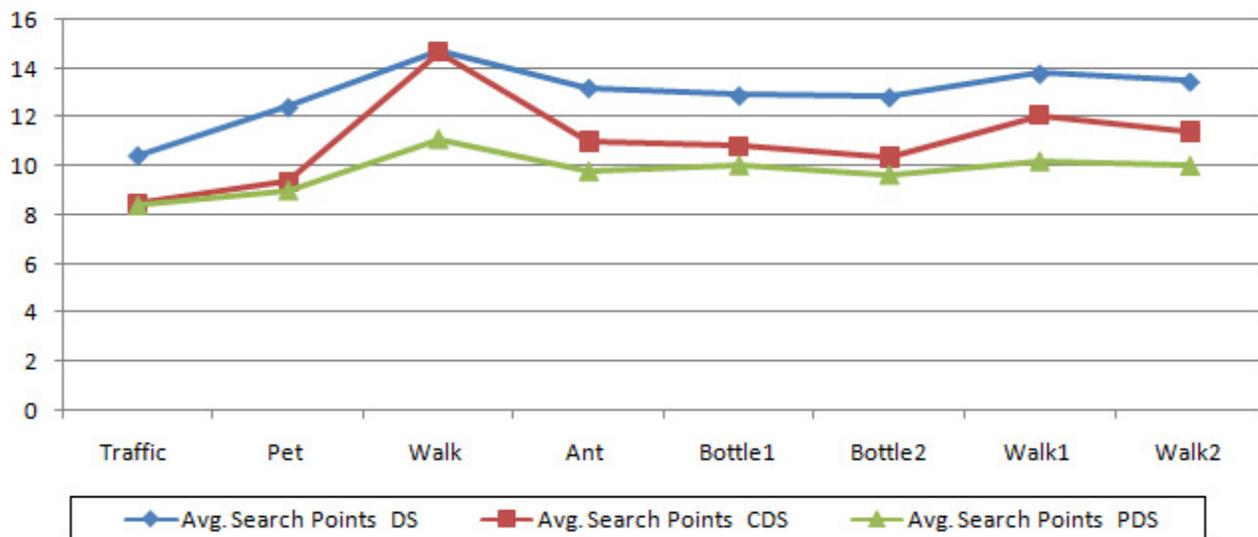


Fig. 13. Performance comparison of DS, CDS and MDS by average search points.

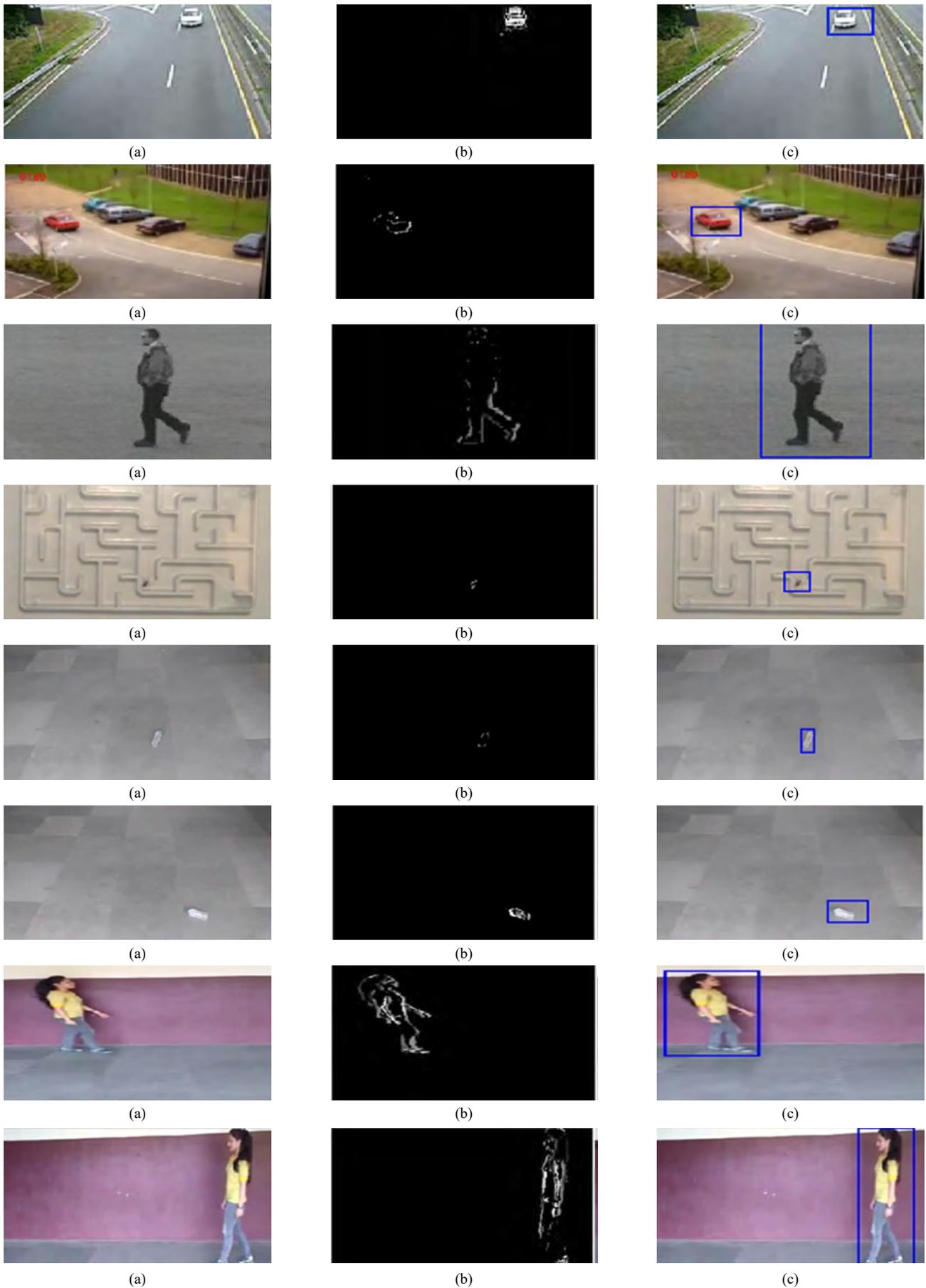


Fig. 14. (a) Original video Frame (b) Background Subtraction (c) Object Tracking Result of the video sequences.





Fig. 15. (a) Original video frame, (b) YUV converted frame, (c) Output after background subtraction, (d) Result of object tracking.

C. Results of Object Tracking

The experiments start with the background subtraction and have implemented the Modified Diamond search algorithm on all the video sequences, to obtain the results. The first 20 frames of the "Walk1" sequence are processed and the step wise output is given, as shown in Fig. 15. The first image is the original video frame, followed by its YUV converted frame. Third image shows the output after the background subtraction and the last image shows the result after object tracking.

D. Simulator

The proposed approach is implemented using (MATLAB 8.1.0.604 (R2013a)). The experiments are carried out on Intel (R) Core (TM) 3 Duo T6570, 2.10 GHz processor. The RAM of 4GB is used. The operating system is 32-bit installed on Windows 7 platform.

V. CONCLUSION AND FUTURE SCOPE

In this paper, a modified diamond search algorithm (MDS) is proposed by combining the capabilities of Diamond search (DS) and Cross diamond search (CDS) block matching algorithm for finding motion vectors between two macro blocks. These motion vectors are then used to track the object in video sequence. The proposed algorithm uses a 3x3 area as the first most step and diamond shape as the further steps. The proposed work is concerned with designing an algorithm that could match or improve the performance of the DS and CDS, in terms of the search points required and speed up the tracking process. The experimentation performed and obtained results show that the proposed algorithm requires fewer searching points as compared to DS and CDS algorithms. The tracking process can be made fast if the search points required by the proposed algorithm can be reduced. Hence this research concludes that the Modified diamond search Algorithm is better, when compared with DS and CDS, based on the experimental results.

The literature on object tracking is very rich. Developing a new way of tracking object from the video scene by combining the object detection approach with the block matching algorithm gives a scope of improvement to existing object tracking techniques. In future this research can extend the block matching algorithm for object tracking applications using real-time object tracking system.

REFERENCES

- [1] H.S. Parekh, U.K. Jaliya, D.G. Thakore, "A Survey on Object Detection and Tracking Methods", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 2, 2014.
- [2] B. Karasulu and S. Korukoglu, "Moving Object Detection and Tracking in Videos", Springer Briefs in Computer Science, 2013.
- [3] S. H. Shaikh, N.Chaki, K. Saeed, "Moving Object Detection Using Background Subtraction", Springer Briefs in Computer Science, 2014.
- [4] Helly, M. Desai, V. Gandhi, "A Survey: Background Subtraction Techniques", International Journal of Scientific & Engineering Research, Vol. 5, Issue 12, 2014.
- [5] M.Piccardi, "Background subtraction techniques: a review", IEEE International Conference on Systems, Man and Cybernetics, 2004.
- [6] R.Zhang, J.Ding, "Object Tracking and Detecting Based on Adaptive Background Subtraction", International Workshop on Information and Electronics Engineering, 2012.
- [7] Duc Phu Chau, Francois Bremond, Monique Thonnat, "Object Tracking in Videos: Approaches and Issues", the International Workshop "Rencontres UNS-UD", 2013.
- [8] A.K.Chauhan and D.Kumar, "Study of Moving Object Detection and Tracking for Video Surveillance", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, Issue 4, 2013.
- [9] Chia-Ming Wu and Jen-Yi Huang, "A New Block Matching Algorithm for Motion Estimation", Applied Mechanics and Materials, Trans Tech Publications, 2017.
- [10] N.Singh, A.Mishra, "Block Matching Algorithm for Motion Estimation using Previous Motion Vector Pattern", International Journal of Computer Applications, Vol. 150, No.8, September, 2016.
- [11] H. A.Surrah, Mohd. J. Haque, "A Comparative Approach for Block Matching Algorithms used for Motion Estimation", International Journal of Computer Science Issues, Vol. 11, Issue 3, No. 2, 2014.
- [12] Hsiang-Kuo Tang, Tai-Hsuan Wu, Ying-Tein Lin, "Real-time object Image Tracking based On Block Matching Algorithm", Project Report available at <https://homepages.cae.wisc.edu/~ece734/project/s06/lintangwuReport.pdf>
- [13] M.H.Sherie, I.Ashimaa, Mahmoud Imbaby and A. Elam, "Experimental Comparison among Fast Block Matching Algorithms (FBMAs) For Motion Estimation and Object Tracking", in Proceeding of 28th National Radio Science Conference (NRSC 2011), National Telecommunication Institute, Egypt, 2011.
- [14] B.N. Tejas, S.H. Bharathi, K.N. Vidya Sagar, "Block Based Algorithms for Estimating Motion", International Research Journal of Computer Science, Vol. 3, Issue 05, 2016.
- [15] B.Sugandi, H.Kim., J.K.Tan and S.Ishikawa, "A Block Matching Technique for Object Tracking", 2008 International Conference on Computer and Communication Engineering, Kuala Lumpur, 2008.
- [16] N. Verma, T. Sahu, P.Sahu, "Efficient Motion Estimation by Fast Three Step Search Algorithms", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 1, Issue 5, 2012.
- [17] Bassam, Ait-Boudaoud, Djamel, "Improved diamond half-pel hexagon search algorithm for block-matching motion estimation", *EI 2017 proceedings*. Society for Imaging Science and Technology, IS&T International Symposium on Electronic Imaging, 2016.
- [18] S.Banchhor, D.Shukla, "An Improved Diamond Search Pattern for Motion Estimation", IEEE Transaction on Image Processing, Vol. 9, Issue 2, 2016.
- [19] C.H. Cheung, L.M.Po, "A Novel Cross-Diamond Search Algorithm for Fast Block Motion Estimation", i-manager's Journal on Pattern Recognition, 2002.
- [20] S. Acharjee, N. Dey, D. Biswas, P. Das, and S. Chaudhuri, "A novel Block Matching Algorithmic Approach with smaller block size for motion vector estimation in video compression", 12th IEEE International Conference on Intelligent Systems Design and Applications (ISDA), 2012.
- [21] S.Kamble, N.Thakur, L.Malik, P.Bajaj, "Color video compression based on fractal coding using quad tree weighted finite automata", Information system design and Intelligent application, Proceedings of Second International Conference INDIA 2015, Advances in Intelligent System and Computing, Springer India, Vol. 340, pp. 649-658, 2015.
- [22] S.Kamble, N.Thakur, L.Malik, P.Bajaj, "Quad Tree Partitioning and Extended Weighted Finite Automata Based Fractal Color Video Coding", Int.J.Image Mining, Vol. 2, No. 1, pp. 31-56, 2016.
- [23] S.Kamble, N.Thakur, P.Bajaj, "A Review on Block Matching Motion Estimation and Automata Theory based Approaches for Fractal Coding", International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 4, No. 2, pp. 91-104, 2016.
- [24] S.Kamble, N.Thakur, L.Malik, and P.Bajaj, "Fractal Video Coding using Modified Three-Step Search Algorithm for Block Matching Motion Estimation", Computational Vision & Robotics in Proceedings of International Conference on Computer Vision & Robotics, Advance in Intelligent Systems & Computing, Vol. 332, pp. 151-162, 2015.



Apurva S. Samdurkar

Apurva S. Samdurkar received Bachelor of Engineering degree in Computer Science and Engineering from Pankaj Laddhad Institute of Technology and Management Studies Buldana, India under the Sant Gadge Baba Amravati University, Amravati, India. She received Master of Technology degree from Yeshwantrao Chavan College of Engineering, Nagpur, India. Presently she is working as a Software Engineer at MRR Soft LLC. She is member of CSI, India and IEANG.



Shailesh D. Kamble

Shailesh D. Kamble received Bachelor of Engineering degree in Computer Technology from Yeshwantrao Chavan College of Engineering, Nagpur, India under the Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur, India. He received Master of Engineering degree from Prof Ram Meghe Institute of Technology and Research, formerly known as College of Engineering, Bandera under Sant Gadge Baba Amravati University, Amravati, India. Presently, he is Ph.D. candidate in department of Computer Science and Engineering from G.H. Raison College of Engineering (An Autonomous Institution Affiliated to Rashtrasant Tukadoji Maharaj Nagpur University), Nagpur, India. He is working as the Assistant Professor in Department of Computer Science and Engineering at Yeshwantrao Chavan College of Engineering, Nagpur, India. He is having over 15 years of teaching and research experience. His current research interests include image processing, video processing and language processing. He is the author or co-author of more than 20 scientific publications in International Journal, International Conferences, and National Conferences. He is the life member of ISTE, India, IE, India.



Nilesingh V. Thakur

Nilesingh V. Thakur received Bachelor of Engineering degree in Computer Science Engineering from Government College of Engineering, Amravati, India and Master of Engineering degree from College of Engineering, Bandera under Amravati University and Sant Gadge Baba Amravati University in 1992 and 2005 respectively. He received Ph.D. degree in Computer Science Engineering under Department of Computer Science Engineering from Visvesvaraya National Institute of Technology, Nagpur, India on 1st February, 2010. His research interest includes image and video processing, sensor network, computer vision, pattern recognition, artificial neural network and evolutionary approaches. He is having over 24 years of teaching and research experience. Presently, he is Professor and Head in Department of Computer Science and Engineering and Dean, PG Studies at Prof Ram Meghe College of Engineering and Management, Bandera- Amravati, Maharashtra, India. He is the author or co-author of more than 70 scientific publications in refereed International Journal, International Conferences, National Journal, and National Conferences. He is a member of editorial board of over eight International journals; also, he is the life member of ISTE, India, IAENG and IAEME. He also worked as the reviewer for refereed international journals and conferences.

Spatial and Textural Aspects for Arabic Handwritten Characters Recognition

Y. Boulid^{1*}, A. Souhar², Mly M. Ouagague¹

¹ Research and Development Unit, Maxware Technology, Kénitra (Morocco)

² Department of Computer Science, Faculty of Science, University Ibn Tofail, Kénitra (Morocco)

Received 4 September 2017 | Accepted 30 November 2017 | Published 19 December 2017



ABSTRACT

The purpose of the present paper is the recognition of handwritten Arabic characters in their isolated form. The specificity of Arabic characters is taken into consideration, each of the proposed feature extraction method integrates one of the two aspects: spatial and textural. In the first step, a modified Bitmap Sampling method is proposed, which converts the character's images into a binary Matrix and then constructs a Mask for each class. A matching rate is used between the input binary matrix and the masks to determinate the corresponding class. In the second step we investigate the use of an Artificial Neural Network as classifier with the binary matrices as features and then the histograms of Local Binary Patterns to capture the texture aspect of the characters. Finally, the results of these two methods are combined to take into consideration both aspects at the same time. Tested on the Arabic set of the Isolated Farsi Handwritten Character Database, the proposed method has 2.82% error rate.

KEYWORDS

Handwritten Recognition, Arabic Documents, Local Binary Patterns, Modified Bitmap Sampling.

DOI: 10.9781/ijimai.2017.12.002

I. INTRODUCTION

ARCHIVES constitute the human documentary production; they reflect the activity of individuals and organizations in time.

The conservation of these archives is one of the major goals of the nations since they show their evolution. The use of the information contained in these archives are still a challenge since most documents are still in their original formats (paper) which makes searching and retrieving information a very hard task, even for scanned documents (saved as images on electronic media) because of the need to manually search through images one by one to find a specific information.

With the scientific progress in terms of computing resources of machines and the development of sophisticated algorithms, converting scanned documents to electronic versions has become possible, allowing indexing and searching content using queries feasible. Systems allowing the conversion of printed documents into an editable text are called Optical Character Recognition (OCR).

The conversion of historical handwritten documents into editable text is more challenging due to irregularity and diversity of handwritten styles, and also to the poor quality of the paper. The systems that correspond to this category are known as Intelligent Character Recognition (ICR).

As many object recognition problems, the process of hand written recognition often consists of five stages executed in a linear way: preprocessing, segmentation, data representation (feature extraction), classification/recognition and post-processing.

Each of these stages consists of a number of techniques and methods adapted to specific problems, such as: noise suppression, skew

and slope correction, text lines segmentation, words segmentation, characters recognition, lexical, syntax and semantic verification.

In addition to the complexity of a document, the language of which it is written with has an influence in the performance of the recognition result. For instance recognizing Latin script is much easier than Arabic or other cursive scripts.

Today, Arabic language is spoken by over 300 million people. The written direction is from right to left and the script is semi-cursive which means that some letters may connect to each other.

Depending on their position in the word the letters have shapes that may change, as they are preceded and/or followed by other letters or isolated.

The information about diacritical points is important in differentiating between letters sharing similar shape. The number and/or positions of these points differ, as for the three different letters 'ba', 'taa', 'thaa' (ب, ت, ث) that have the same basic shape but differ in the position and the number of points.

Arabic characters have different spatial disposition, which could be horizontal, vertical, diagonal or other. The texture of a character may change along its shape. The fact that Arabic script is written from right to left influences on the texture of the shape in the beginning which differs from the one in the ending.

We think that in the design of feature extraction methods for Arabic characters, one must take into consideration these two aspects spatial and textural.

The present paper is an extension of our work in [1]. We will first compare the recognition accuracy of two feature extraction methods respecting the spatial and textural aspects, namely modified Bitmap Sampling method and Local Binary Patterns (LBP). In the second step the combination of these two methods is investigated in order to take into consideration both aspects.

* Corresponding author.

E-mail address: y.boulid@gmail.com

The rest of the paper is organized as follows: Section 2 provides a brief overview of some related works on the recognition of Arabic letters. Section 3 describes the used feature extracted methods, provides results of each one and also investigates their combination, and compares the performance of the proposed method with existing works on IFHCDB dataset. Finally, section 4 concludes the paper.

II. RELATED WORKS

Generally, in handwritten recognition there are two category of feature extraction: either statistical [2] such as histograms of transition and projection profile, histograms of gray level distribution, Fourier descriptors and chain code; or structural [3, 4] such as the presence of loops, number and position of diacritical points and orientation of curves. Table I presents an overview of some related works dealing with the recognition of isolated Arabic and Farsi handwritten characters.

III. FEATURES EXTRACTION METHODS WORKS

The challenge in the recognition of the Arabic letters is that the used descriptor must take into consideration the existence of different characters written in a similar way, for example, some character shares the same body shape but differs only in the diacritical points. In the following, we will give details and results of the used feature extraction methods.

A. Dataset

In [19] a dataset called Isolated Farsi Handwritten Character Database (IFHCDB) was introduced and it was used in the ICDAR 2009 competition [15]. The distribution of characters is not uniform.

In this work, we are interested only in the Arabic portion, which represents approximately 97% of all character set (Arabic and Farsi), and which is divided into 35989 images for learning and 15041 for testing.

B. Modified Bitmap Sampling (MBS)

In order to take into consideration the spatial aspect of the characters, the first method of feature extraction is called Bitmap Sampling, which allows to sub-sample the character's image in a smaller dimension in order to be used as input into a classifier.

We consider the character as a set of filled and empty cells distributed along a grid. The grid corresponds to the smallest bounding box enclosing the character and its diacritics.

In order to make the descriptor invariant to the size of the characters, the number of lines and columns in the grid is fixed for all characters (Fig. 1).

TABLE I
AN OVERVIEW OF SOME METHODS OF ARABIC/FARSI HANDWRITTEN CHARACTER RECOGNITION

Ref	Classifiers	Features	Dataset	Recognition Rate
[1]	Artificial Neural Network	Local Binary Pattern, Spatial distribution of pixels	IFHCDB dataset	96.31%
[5]	Inductive learning program to generate Horn clauses	Structural features (lines, curves and loops)	30 samples for training and 10 samples for testing from each character	86.65%
[6]	Back propagation neural network	Statistical and Morphological and features from the main body and secondary components	CENPRMI	88%
[7]	Non-dominated Sorting Genetic Algorithm	Normalized central and Zernike moments features	Isolated handwritten characters of 48 persons	90%
[8]	Neural network	Wavelet coefficients	Isolated Arabic letters from more than 500 writers	88%
[9]	Vector quantization	Statistical feature such as distributive and concavity	3000 of Farsi letters	85.59%
[10]	Support Vector Machine	Fourier descriptor of the projection profile, Moments and centroid distance	1000 Arabic isolated characters	96.00%
[11]	Artificial Neural Network, Support Vector Machine, and Logistic Regression	HOG, GIST, LBP, SIFT and SURF	Isolated Arabic handwritten characters which contain about 8988 letters collected from 107 writers	94.28%
[12]	Support Vector Machine, K Nearest Neighbor, Neural Networks	Zoning and crossing counts	IFHCDB dataset	94.82%
[13]	Support Vector Machine	Under-sampled bitmaps and chain code directional frequencies	IFHCDB dataset	96.68%
[14]	Support Vector Machine, K Nearest Neighbor, Neural Networks	Gradient, shadow, directional, line-fitting, intersection and undersampled features	IFHCDB dataset	96.91%
[15]	Hierarchy of multidimensional recurrent neural networks	Raw data	CENPARMI and IFHCDB datasets	91.85%
[16]	Artificial neural network	Cosine Transform and Discrete Wavelet Transform	A dataset of 5600 characters	79.87%
[17]	Artificial neural network	Number of dots and angles between points in the contour	A Datasets of 39200 of Arabic isolated characters	96.91%
[18]	K Nearest Neighbor, Random forest, Linear discriminant analysis	Feature selection based on Moth-flame optimization including : gradient, right, left, vertical and horizontal projections, number of holes, aspect ratio, number and position of secondaries	CENPARMI dataset	99.25%

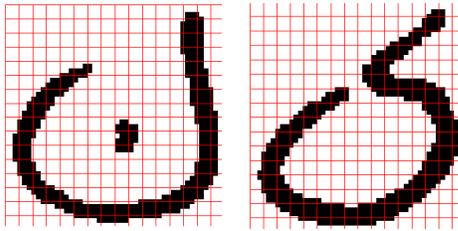


Fig. 1. Example of decomposition of the bounding box of a character (from the IFHCDB dataset) in a grid of 16x16 cells.

From this grid called G, a binary matrix called M with the same size, is generated as in (1):

$$M(i, j) = \begin{cases} 1 & \text{if } G(i, j) \text{ not empty} \\ 0 & \text{else} \end{cases} \quad (1)$$

From the training set and for all the samples of a class, the Matrices of all characters are superposed. The result is called “The Mask of the class”. This Mask contains for each cell, the number of times this cell is filled (Fig. 2).

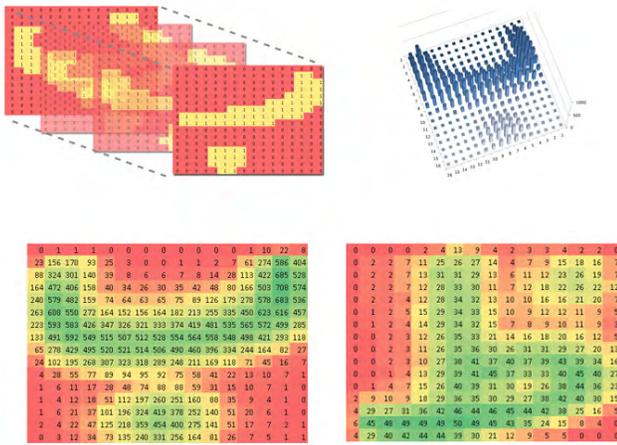


Fig. 2. Example of a 16x16 Mask reconstruction. (a) The superposition of all Matrices of the character (Baa, ب). (b) a 3D histogram of the Mask. (c) and (d) are respectively the resulted Masks of the classes (Baa ب, and thad, ا).

Since the number of samples in the dataset is not the same for all characters, the resulted Masks will contain values of different scales which will influence the result of the comparison. For example, the maximum value in the Mask in Fig. 2.c is 708, while the maximum value in the Mask in Fig. 2.d is 50.

To cope with this issue, the idea is to normalize all the Masks using one scale. Therefore, the algorithm below is used for this purpose.

The normalized versions of the Masks in Fig. 2.c and 2.d, are shown in the Fig. 3. As can be noticed the maximum value in both Masks after applying the normalization algorithm is equal to 256.



Fig. 3. Normalized Masks for the two characters (Baa ب, and thad, ا).

Algorithm 1 Normalization of a Mask

```

for row = 1 to WS do
  for col = 1 to WS do
    M[k, 1] = row
    M[k, 2] = col
    M[k, 1] = Mask[row,col]
  end for
end for
OM = SORT(M)
NbrCells = WS * WS
MaxVal = NbrCells
for k = 1 to NbrCells do
  NMask(OM(K, 1), OM(K, 2)) = MaxVal
  MaxVal = MaxVal - 1
end for
    
```

For a given character to be recognized, we extract its binary Matrix and compute the Matching Rate (MR) from all the Masks as in (2):

$$MR_k(B) = \frac{1}{1 + \sum_{i=1}^W \sum_{j=1}^W M_k(i, j) * B(i, j)} / k=1...28 \quad (2)$$

Where M_k is the Mask of the class k , and B is the binary Matrix of the character to be recognized, W is the number of rows and columns. $MR_k(B)$ is the amount of participation which measures how the input is close to the Mask of the class k . Therefore, computed over all the Masks, the corresponding class of an input character is the one with the smallest distance.

Fig. 4, illustrates the TOP3 measures (the percentage of samples that the true class is among the TOP3 positions in the list of candidates) recognition rates according to different configuration of the grid.

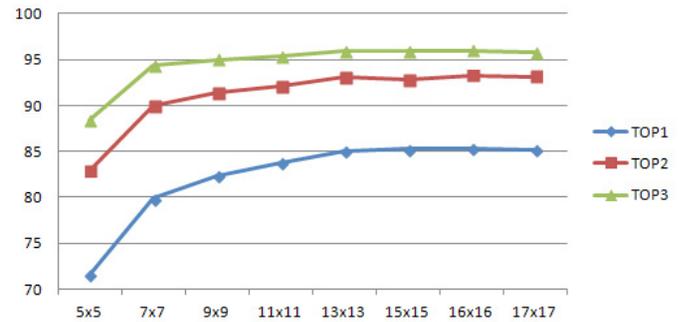


Fig. 4. TOP3 measures for different configurations of the size of the grid.

The higher recognition rate corresponds to the 16x16 configuration, which gives respectively for TOP1, TOP2 and TOP3 the rates 85.29%, 93.29% and 95.98%. This configuration allows certainly to resolve the inter-class problem since a given character will be certainly included in the Mask of some class, but to the detriment of the intra-class problem, since the variability of the shapes of characters allow some masks to be included in others, such as the case of (sin, س and yaa, ي) in Fig. 5.

As alternative, using the Binary Matrices of the characters (Fig. 2.a) as features with a feed-forward neural network with Scaled Conjugate Gradient as training algorithm allows us to reach respectively for TOP1, TOP2 and TOP3 the rates 94.56%, 97.62% and 98.51%.

Based on the result of the confusion matrix we have noticed that this method produces errors between some characters such as (ج, ا), (ل, ن), (ت, ن).



Fig. 5. Example of inclusion of the two normalized Masks of the characters (a) (sin, س) and (b) (ya, ي).

When writing a character, a writer does not put the same pressure at the beginning as the end. Certain characters such as (ش, ح, ط, ض, س) require more vigilance from the writer as other character like (ن, ر, د, ل) because they contain more details which result in changing the texture of their shape. This information is not captured by the Modified Bitmap Sampling method, hence the need to use another method that integrates this aspect.

The next section gives a brief overview of Local Binary Patterns method which is used to capture the textural aspect.

C. Local Binary Patterns (LBP)

Developed by Ojala et al [20], LBP for Local Binary Pattern, is a descriptor designed to analyze textures in terms of local spatial patterns and gray level contrast. In its basic form, a window is used to produce labels for each pixel by thresholding to neighboring pixels with the center pixel in the window and considers the result as a binary number as illustrated in Fig. 6.

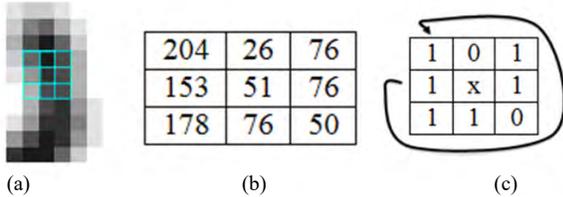


Fig. 6. Example of LBP computation: (a) a 3x3 window centered on a pixel, (b) the gray levels in the window, (c) thresholding the neighborhood compared to pixel x.

As in [1], the use of LBP histogram extracted from the smallest rectangle containing the character and its diacritical points with a feed-forward neural network and a Scaled conjugate Gradient as training algorithm gives a recognition rate of 88.47% which is reasonable, since in the computation of LBP Histogram the spatial information is lost.

So in order to integrate the spatial aspects, as in [1], we simply split the bounding box of the character into four regions from the center point between the centroid of the character and the centroid of diacritics and extract the histogram from each region and then concatenates them to construct the feature vector (Fig. 7). This way allows to discriminate between certain characters that have the same termination.

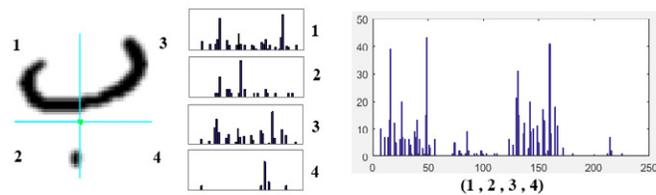


Fig. 7. Concatenation of the histograms extracted from the four regions from the centroid of the character and its diacritics (taken from [1]).

Using these histograms feed-forward neural network with Scaled Conjugate Gradient as training algorithm allows to increase the

recognition rate up to 96.31%, 98.91% and 99.44% respectively for TOP1, TOP2 and TOP3 which proves that the integration of both spatial and textural aspect is beneficial. Table II shows the recognition rates for some characters for the two feature extraction methods.

TABLE II
RECOGNITION RATES OF SOME CHARACTERS ACCORDING TO MBS AND LBP

Recognition rate (%)	ر	ا	ل	ن	ت	س
MBS	91.67	98.60	92.00	94.08	92.02	96.50
LBP	95.23	99.57	95.33	97.21	90.85	96.50

For the case of ت and س the spatial aspect discriminates better than LBP. This is related to the way the characters are written and to the speed of writing. Some details of the character such as texture may be ignored and therefore the spatial shape can be effective in these cases.

From these results, we can say that the spatial aspect deals with the global shape of the character, but ignores the details which could be captured by the textural aspect. While for LBP the calculation of the histogram loses the spatial information which we compensate by considering four regions of the character.

In what follows we investigate another way of combination that combines the results of the classifiers based on these features.

D. Combination of Classifiers

The goal of combining classifiers is to conceive a more efficient one which operates on the same data as the base classifiers and separates the same type of classes [21]. As stated in [21] if we denote the score assigned to the class i , by the classifier j : S_{ij} , Therefore, a combination rule is a function f , and the final combination of the scores of the class i is in (3):

$$S_i = f(\{S_{ij}^j, j=1 \dots M\}) \quad (3)$$

Then the sample is classified as belonging to the class with the maximum score. In this work we have two classifiers of the same type (neural networks) that are trained with different features. So we are interested in non-ensemble combinations [21].

Since we have the scores of the classifications for each classifier, we can use basic combination schemes at the score level, by applying simple rules such as: the maximum, average, multiplication. For an input image, respectively the maximum, the average and the multiplication rule simply calculates the maximum, the average, the product of the scores of each classifier for each class, and assigns the label of the class having the maximum score.

Table III shows the recognition rates according to the combination of MBS and LBP.

TABLE III
RECOGNITION RATES ACCORDING TO DIFFERENT COMBINATION RULES

Combination (%)	Maximum	Average	Multiplication
MBS + LBP (entire image)	94.32%	94.47%	95.75%
MBS + LBP (split image)	96.84%	96.95%	97.18%

The multiplication rule applied to the scores of the MBS classifier and the classifier based on LBP after splitting the image gives the best recognition rates which is about 97.18 %, 99.19% and 99.54% respectively for TOP1, TOP2 and TOP3.

E. Comparative Analysis

Is difficult to compare the performance of the proposed approach with most of related works since most existing datasets are not public or not representative (small number of characters). Therefore a comparative analysis (Table IV) is conducted in order to compare the proposed approach with the existing systems working on the same dataset (IFHCDB) when considering 28 (Arabic set), 32 and 33 (Arabic and Farsi sets) class problems.

TABLE IV
COMPARISON OF THE PROPOSED SYSTEM WITH OTHER METHODS ON IFHCDB DATASET

Method	Training size	Testing size	Number of classes	Reco rate (%)
[12]	36000	13320	33	94.82
[13]	36682	15338	32	96.68
[22]	36682	15338	32	96.91
[1]	36437	15233	32	95.87
[1]	36437	15233	33	96.04
[1]	35989	15041	28	96.31
Proposed method	36437	15233	32	96.51
Proposed method	36437	15233	33	96.50
Proposed method	35989	15041	28	97.18

F. Discussion

As can be seen from the result of the combination rules, taking into consideration both spatial and textural aspect improves more the recognition rate. This confirms that the two aspects are complementary since both classifiers participate equally without favoring one over the other as in the case of the maximum and the average rules.

As can be seen from the comparison the proposed method is competitive. The analysis of the confusion matrix (Fig. 8) shows that most errors are due mainly to :

- The confusion between different characters drawn in a similar way such as the case of (ر , ل), (ع , ح) and (ص , س).
- The presence of illegible characters due to the noise.
- The presence of misclassified characters due to human errors of classification during the assembly of the database.

Fig. 9 shows some of unsuccessful recognition of the proposed method which proves that even humans may find difficulty to recognize some of them.



Fig. 9. Example of some misclassified characters.

	ا	ب	ت	ث	ج	ح	خ	د	ذ	ر	ز	س	ش	ص	ض	ط	ظ	ع	!	"	#	\$	%	&	'	()
ا	2992						4		2													1			1	
ب		39						2	1					1								1			1	
ت			46	1															4	3	5			6	1	
ث			4																							
ج					149	16	1											2								
ح						364											5									
خ					2		12										1	3								
د	2							791		5																
ذ									7		2								1	2						
ر	37	4					2	3	1245	2	1													6	2	
ز			1				1	7	3	322		1														
س									1	7		5									1				7	
ش									1		3	16	1	1												
ص									1	1	23		114	1								1				
ض												6		12												
ط															59	1										
ظ															2	13										
ع	1				1	12												241	3					1	2	
!				1			4											1	36							
"																			268		5				2	
#			2									3							5	126	1		1		3	
\$										1		1									169	1			9	
%	3	1					2	1													1	858		34		
&	3				1			1						1									1483	1	1	
'			7	1				3										1	4	2	1	4		112	5	
()		2					5	1	3															397		
)		1	1	1	1					1	1	5		1					1			1			753	
)		1	1	1	1					1	1	5		1					1		4	1	3	2	5	
)																									27	
)																									13	

Fig. 8. Confusion matrix of the proposed method.

IV. CONCLUSION

The texture of a character changes from beginning to the end. Using LBP to take into consideration this information allows to tolerate the inter-class variability. And in order to recognize characters that differ only in certain details, we divided the image in four regions so as to differentiate those in interclass.

This technique is limited to treat different characters having similar shape but differ only in diacritical points. This motivated us to propose a method integrating the spatial aspect.

These two methods operate on the same object from different but complementary viewpoints. By combining them at the scores level using multiplication rule, the recognition rate is improved to 97.18% in TOP1 and 99.19% in TOP2.

REFERENCES

- [1] Y. Boulid, A. Souhar, and M. Y. Elkettani, "Handwritten character recognition based on the specificity and the singularity of the arabic language," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no.4, pp. 45-53 Regular Issue, 2017.
- [2] J. Cai and Z.-Q. Liu, "Integration of structural and statistical information for unconstrained handwritten numeral recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 3, pp. 263–270, 1999.
- [3] L. Heutte, T. Paquet, J.-V. Moreau, Y. Lecourtier, and C. Olivier, "A structural/statistical feature based vector for handwritten character recognition," *Pattern recognition letters*, vol. 19, no. 7, pp. 629–641, 1998.
- [4] M. T. Parvez and S. A. Mahmoud, "Arabic handwriting recognition using structural and syntactic pattern attributes," *Pattern Recognition*, vol. 46, no. 1, pp. 141–154, 2013.
- [5] A. Amin, "Recognition of hand-printed characters based on structural description and inductive logic programming," *Pattern recognition letters*, vol. 24, no. 16, pp. 3187–3196, 2003.
- [6] A. Sahlol and C. Suen, "A novel method for the recognition of isolated handwritten arabic characters," *arXiv preprint arXiv:1402.6650*, 2014.
- [7] G. Abandah and N. Anssari, "Novel moment features extraction for recognizing handwritten arabic letters," *Journal of Computer Science*, vol. 5, no. 3, p. 226, 2009.
- [8] A. Asiri and M. S. Khorsheed, "Automatic processing of handwritten arabic forms using neural networks.," in *IEC (Prague)*, pp. 313–317, 2005.
- [9] J. Shanbehzadeh, H. Pezashki, and A. Sarrafzadeh, "Features extraction from farsi hand written letters," *Proceedings of Image and Vision Computing*, pp. 35–40, 2007.
- [10] F. Bouchareb, R. Hamdi, and M. Bedda, "Handwritten arabic character recognition based on svm classifier," in *Information and Communication Technologies: From Theory to Applications, 2008. ICTTA 2008. 3rd International Conference on*, pp. 1–4, IEEE, 2008.
- [11] M. Toriki, M. E. Hussein, A. Elsallamy, M. Fayyaz, and S. Yaser, "Window-based descriptors for arabic handwritten alphabet recognition: A comparative study on a novel dataset," *arXiv preprint arXiv:1411.3519*, 2014.
- [12] M. Rajabi, N. Nematbakhsh, and S. A. Monadjemi, "A new decision tree for recognition of persian handwritten characters," *International Journal of Computer Applications*, vol. 44, no. 6, pp. 52–58, 2012.
- [13] A. Alaci, P. Nagabhushan, and U. Pal, "A new two-stage scheme for the recognition of persian handwritten characters," in *Frontiers in Handwriting Recognition (ICFHR), 2010 International Conference on*, pp. 130–135, IEEE, 2010.
- [14] A. Alaci, U. Pal, and P. Nagabhushan, "A comparative study of persian/arabic handwritten character recognition," in *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on*, pp. 123–128, IEEE, 2012.
- [15] S. Mozaffari and H. Soltanizadeh, "Icdar 2009 handwritten farsi/arabic character recognition competition," in *Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on*, pp. 1413–1417, IEEE, 2009.
- [16] A. Lawgali, A. Bouridane, M. Angelova, and Z. Ghassemlooy, "Handwritten arabic character recognition: Which feature extraction method?," *International Journal of Advanced Science and Technology*, vol. 34, pp. 1–8, 2011.
- [17] Hammad, N. H., & Musa, M. E. (2016). The Impact of Dots Representation in Recognition of Isolated Arabic Characters. *International Journal of Information Engineering and Electronic Business*, vol. 8, no 6, p. 37.
- [18] Ewees, A. A., Sahlol, A. T., & Amasha, M. A (2017). A Bio-inspired Moth-Flame Optimization Algorithm for Arabic Handwritten Letter Recognition. *International Conference on Control, Artificial Intelligence, Robotics and Optimization (ICCAIRO 2017)*.
- [19] S. Mozaffari, K. Faez, F. Faradji, M. Ziaratban, and S. M. Golzan, "A comprehensive isolated farsi/arabic character database for handwritten ocr research," in *Tenth International Workshop on Frontiers in Handwriting Recognition*, Suvisoft, 2006.
- [20] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51–59, 1996.
- [21] S. Tulyakov, S. Jaeger, V. Govindaraju, and D. Doermann, "Review of classifier combination methods," in *Machine Learning in Document Analysis and Recognition*, pp. 361–386, Springer, 2008.
- [22] M. Askari, M. Asadi, A. Asilian Bidgoli, and H. Ebrahimpour, "Isolated persian/arabic handwriting characters: Derivative projection profile features, implemented on GPUs," *Journal of AI and Data Mining*, vol. 4, no. 1, pp. 9–17, 2016.



Youssef Boulid

artificial intelligence.



Abdelghani Souhar

A. Souhar received M.S. degree in applied Mathematics in 1992, PhD degree in computer science in 1997 from University Mohammed 5 in Rabat - Morocco. Now he works as a Professor at university Ibn Tofail in Kenitra – Morocco. His research interests include CAE, CAD and Artificial Intelligence.



Mly Moustafa Ouagague

Mly.M Ouagaguereceived his M.S. degree in Decision Support Systems in 2001 from Blaise Pascal University. Now he works as a researcher at Maxware Technology Kénitra - Morocco. His research interests include artificial intelligence, complex system, image processing, serious games.

A Novel Smart Grid State Estimation Method Based on Neural Networks

Mohamed Abdel-Nasser*, Karar Mahmoud, Heba Kashef

Department of Electrical Engineering, Aswan University, 81542 Aswan (Egypt)

Received 27 December 2017 | Accepted 22 January 2018 | Published 29 January 2018



ABSTRACT

The rapid development in smart grids needs efficient state estimation methods. This paper presents a novel method for smart grid state estimation (e.g., voltages, active and reactive power loss) using artificial neural networks (ANNs). The proposed method which is called SE-NN (state estimation using neural network) can evaluate the state at any point of smart grid systems considering fluctuated loads. To demonstrate the effectiveness of the proposed method, it has been applied on IEEE 33-bus distribution system with different data resolutions. The accuracy of the proposed method is validated by comparing the results with an exact power flow method. The proposed SE-NN method is a very fast tool to estimate voltages and re/active power loss with a high accuracy compared to the traditional methods.

KEYWORDS

Neural Network, Smart Grid, Renewable Energy, Power Loss, Voltage Profile.

DOI: 10.9781/ijimai.2018.01.004

I. INTRODUCTION

THE smart grid technologies are a step towards developing the utility industry, whereas it could overcome the drawbacks of the existing power grid [1]. The high demand for electricity cannot serve the high increasing population rate and the low response. The manual control and the high losses during transmission made the manufacturers think about alternative solutions. Smart grids can deal with the needs of the high population [2], achieve reliability, automated control which leads to the high response, and decrease the losses in transmission lines by providing efficiency of data transmission from head to tail [3]–[5].

The feeding of the smart grid [6] can be renewable energy resources (e.g., photovoltaic (PV) and wind) in parallel with the utility grid to provide a low overall cost. On the other side, these resources have negative impacts on the distribution system, such as the high power loss due to the back feeding of the power flow [7], fluctuation and unbalanced voltage [8], and voltage rise [9].

Many researchers evaluated the steady-state conditions of distribution system by using the state-of-the-art techniques, such as Gauss-Seidel and Newton-Raphson, which supposed that the system is already stable [10]. The power flow techniques have been applied to determine the steady-state operation conditions during transmitting the electricity, Kirchhoff's current law and other methods to determine the voltage magnitude, angle, and active/reactive power. Other iterative methods are detailed in [11], [12]. A technique for the multiphase unbalanced system, which was developed to deal with different types of loads (constant current, constant power, and constant impedance), is presented in [13]. To model the harmonics resulting by the variable load with the time variant, a simulator is modeled based on the openDSS program to estimate the value of harmonics [14].

Impact of PV integration was studied in [15]. A study for modeling and simulating the penetration effects by two feeders in different scenarios is illustrated in [16].

A quasi-static time-series (QSTS) is presented for analyzing and simulating the operation of PV distribution system and evaluating the effects of the PV integration during a specified duration (day or year).

In [17], several works have been proposed to reduce the execution time of QSTS. The authors presented two methods for shortening the time-series for power flow calculations in the presence of the distributed generation. The size of the input data, as well as the power flow calculation, was decreased. The first method depends on down-sampling data through reducing time resolution. This method has the disadvantage of losing the peak values during down-sampling. The second method chooses the similar intervals, and then performs the calculation on them. This method also has the disadvantage of missing the rest of data, which makes the simulations inaccurate.

In [18] and [19], the neural networks control the electrical systems with nonlinear dynamic characteristics. Ref [20] explained the use of neural networks with fully connected neuron learning combined with power flow for optimizing smart grids.

The limitation of the methods reported in the literature is that they take very long execution time to estimate smart grid state. In this paper, we propose a fast method (SE-NN) to calculate the state of the smart grid (voltage, active power, and reactive power) in a short time and accurate way using a neural network model.

II. BACKGROUND

Power flow analysis is essential for monitoring system distribution by calculating active/reactive power, currents, and voltages at each time step for a specified duration. Iterative power flow methods are used for determining the operation of the system at the steady-state condition. In [21], a Newton-Raphson method algorithm was used for solving non-linear equations. The problem with this method is that the

* Corresponding author.

E-mail address: egnaser@gmail.com

memory space is not sufficient for large-scale systems. A modified Gauss-Seidel algorithm was presented in [22]. This algorithm solves the problem of the limited memory size with small execution time.

The state estimation is used for estimating the power flow solution of different real-time information because of the large number of nodes, few phase measurements, and disability of the system to be hardly provided with a meter at each node and each branch. The commonly used state estimator is weighted least square criteria (WLS). A study on state estimation depending on Tikhonov regularization compared with the WLS is detailed in [23]. Due to the changeable nature of the energy environment, the distribution system needs new system modeling that can adapt to the new advanced technologies. In [24], the authors discussed the new methodologies for wide area monitoring and analyzing of the smart grids. Another approach for wide-area network estimation called multi-area state estimation (MASE) is illustrated in [25]. This approach relies on a two-step procedure, the overall network area is divided into subareas according to its geographical characteristics and measurement estimator. The first step is performing all the available measurements for each area. The subareas have the same nodes number and the estimator is working in parallel for low execution time. In the second step, the data of the first step is processed and refined for different operation conditions. The previous methods have drawbacks of the long execution time [26]. In general, iterative power flow methods require a long time to estimate the states of the smart grid. In this paper, we propose a fast method for real-time analysis with high accuracy. To demonstrate the effectiveness of our method, we have compared it with a well-known power flow method.

III. THE PROPOSED METHOD

A. Neural Networks

The artificial neural networks are a close simulation of the human brain. We use the neural network in state estimation because of its high efficiency in processing signals and predicting output in a fast and accurate way. The neural network model either contains one layer (input and output are directly connected), two layers (the first layer for receiving data, and the other layer for output results) or multilayer whereas three or more layers are connected in parallel (input layer, hidden layer(s), and output layer) [27]–[29].

There are many algorithms to train neural networks; the preferable one is the back-propagation algorithm in which the error is calculated by computing the difference between the target and the measured output, and the weights are adjusted and propagated backward from the output to the input until finding the weights that minimize the error. Then the network is being learned for the desired output. More details about learning of neural network can be found in [30]. Fig. 1 shows a neural network containing three layers. The first layer is the input layer which contains four neurons, the second layer is the hidden layer that contains three neurons, and the last one is the output layer.

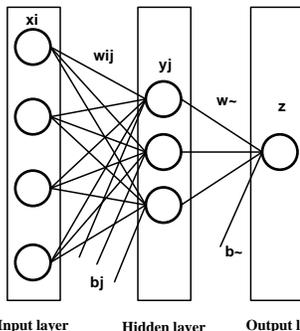


Fig. 1. The architecture of neural networks.

The connection lines w and b denote the weights and biases. The output of the first layer can be expressed as follow:

$$y_j = \sum_{i=1}^k x_i w_{ij} + b_j \quad (1)$$

where y_j is the summation of the inputs x_i multiplied by the weights w_{ij} , and added to the biases b_j [30]. Using equation (2), the weights are updated every iteration until the errors are minimized to a level equal or less than the threshold value (0.05).

$$w_{new} = w_{old} + \alpha (y_{actual} - y_{measured}) \quad (2)$$

where α is the learning rate which is adjusted to a small value. The output of the summation function (y_j) is the input to another function called activation function. The common types of activation function are shown in Fig. 2.

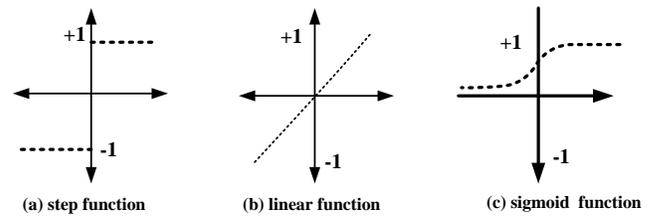


Fig. 2. The common types of activation functions.

The step function or threshold function in Fig.2 (a) is limited with (-1 or +1), If the output is less than or equal to 0, the output will be -1, if it is more than 0, the resulting output will be +1. In Fig.2 (b), the output of the linear function is increasing linearly from -1 to +1. In Fig.2 (c), the output of the sigmoid function is increasing from 0 to +1. The final output of the neural network (z) can be expressed using equation (3).

$$z = f \sum_{j=1}^l y_j w_{jz} + b_z \quad (3)$$

where f is the activation function, and the updated weights and biases are w^- and b^- .

B. SE-NN Method

In the proposed method, the network is constructed directly between the input data and the desired output. The input of the network is a training data set divided into samples within a limited duration called load factor. The load factor samples can be presented by the following equation:

$$LF = [LF_{t_1} \quad LF_{t_2} \quad \dots \quad LF_{t_m}] \quad (4)$$

where t is the time step, and m is the number of samples. The output of the network is the state estimation for the generated load profiles which are sampled by the load factor along with a specified duration. If we considered that the output is presented by a matrix A as shown in equation (5), each element in the matrix A represents the state of a known bus or line at a known sample.

$$A = \begin{bmatrix} X_{b_1,t_1} & X_{b_1,t_2} & \cdots & X_{b_1,t_m} \\ X_{b_2,t_1} & X_{b_2,t_2} & \cdots & X_{b_2,t_m} \\ \vdots & \vdots & \vdots & \vdots \\ X_{b_n,t_1} & X_{b_n,t_2} & \cdots & X_{b_n,t_m} \end{bmatrix} \quad (5)$$

The rows in the matrix A represent the bus number (in case of voltage calculation) or the line number (in case of losses calculations). Note that the number of lines is less than the number of buses by one. The columns represent the values of the input sample with regards to the data set. The symbol X can be replaced with V , P or Q . For example, if we want to find the voltage at bus 15 through all durations, we will extract the vector of 15th row of the matrix as below:

$$V_{b_{15}} = [V_{b_{15},t_1} \quad V_{b_{15},t_2} \quad \cdots \quad V_{b_{15},t_m}] \quad (6)$$

Another example, if we want to find the reactive losses at line 10 (the line from bus 10 to bus 11), we will extract the vector of 10th row of the matrix as the following equation:

$$Q_{l_{10}} = [Q_{l_{10},t_1} \quad Q_{l_{10},t_2} \quad \cdots \quad Q_{l_{10},t_m}] \quad (7)$$

In our approach, we used an efficient method for solving the power flow called quadratic based backward/forward sweep (QBBFS) [31]. Unlike the traditional quadratic based methods that cannot be employed for the unbalanced systems, the QBBFS method accommodates multiphase systems and different load types. Although QBBFS is an efficient method, it requires a long time to estimate the state of the grid. The QBBFS method is used for the large system, where the backward sweep is used for calculating the branch currents from the far ends, and the forward sweep is used for calculating the voltages from the slack bus. In QBBFS, the iteration is repeated until reaching to high convergence rate. Fig. 3 shows the steps of estimation the real-time voltage using a neural network model.

C. The Solution Steps

1. Read the data inputs from the load profiles.
2. Sample the input data and solve it using QBBFS method.
3. Redirect the QBBFS outputs (voltages or re/active power) to the SE-NN model.
4. Create the SE-NN model using the feed forward neural network with initial parameters and Levenberg-Marquardt algorithm that has high efficiency.
5. Train the SE-NN model with the initialize network created and training parameters. In the training process, the actual QBBFS output is compared with the measured SE-NN output until the network performance goal is met (maximum 300 epochs, minimum 5% gradient and 1e-3 goal error). If the goal is not met, the weights and biases are updated with the learning machine rate until the error is equal or less than the network goal.
6. Once the model is trained, it would be able to simulate the output results for any change in the input data set (Load Factor).
7. Finally, the results are printed.

The proposed algorithm overcomes the high complexity of the computational methods. The function of the QBBFS method is to solve the power flow and redirect the output for training the network, once the network model is learned, the model is saved and used to simulate and reproduce the output rapidly and accurately for any new input data without the need for the QBBFS method.

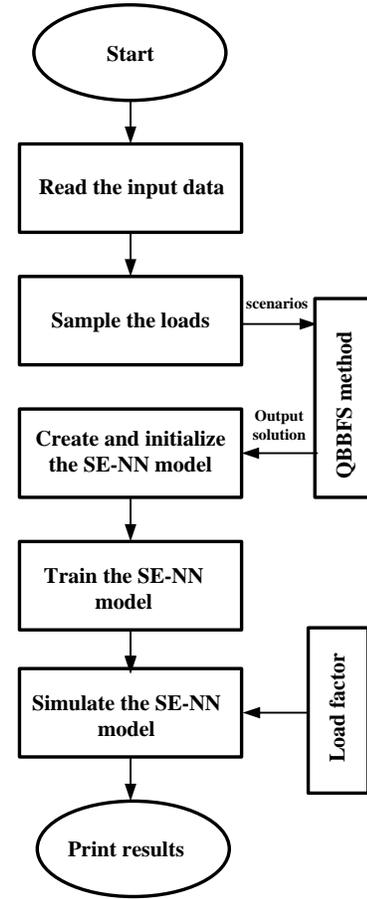


Fig. 3. The flowchart of the SE-NN method.

IV. ASSESSING THE ACCURACY OF THE PROPOSED METHOD

To validate the accuracy of the proposed method, we compute the difference between the actual output using the QBBFS method and the network output using the SE-NN method. To perform this we compute different types of errors such as the mean square error (MSE) which is expressed as the average of the square difference between the actual output and the network output.

$$MSE_i = \frac{1}{M} \sum_{t=t_1}^{t_m} (X_{i,t}^{act} - X_{i,t}^{SE-NN})^2 \quad (8)$$

where M is the size of the duration, i denotes to the bus number (for voltage calculations) or the line number (for losses calculations), and X represents the state estimation (e.g., voltage, active or reactive power), $X_{i,t}^{act}$ is the QBBFS output, and $X_{i,t}^{SE-NN}$ is the SE-NN output. In equation (9) we calculate the square root of the mean of the squared difference between actual and estimated output which is called the root mean square error (RMSE). The maximum value of the absolute difference between actual and measured output which is called the maximum absolute error (MAE) is expressed in (10). The mean absolute percentage error (MAPE) and the sum of the squared error (SSE) are expressed in (11) and (12), respectively.

$$RMSE_i = \sqrt{\frac{1}{M} \sum_{t=t_1}^{t_m} (|X_{i,t}^{act}| - |X_{i,t}^{SE-NN}|)^2} \quad (9)$$

$$MAE_i = \max_{t=t_1, \dots, t_m} (|X_{i,t}^{act} - X_{i,t}^{SE-NN}|) \quad (10)$$

$$MAPE_i = \frac{1}{M} \sum_{t=t_1}^{t_m} \left| \frac{X_{i,t}^{act} - X_{i,t}^{SE-NN}}{X_{i,t}^{act}} \right| \quad (11)$$

$$SSE = \sum_{i=1}^n \sum_{t=t_1}^{t_m} (X_{i,t}^{act} - X_{i,t}^{SE-NN})^2 \quad (12)$$

The maximum values of the previous errors are computed to evaluate the accuracy of the proposed method.

V. RESULTS AND DISCUSSIONS

This section presents the results of state estimation using the proposed SE-NN method in smart grids. This method has been implemented at 2.20GHZ CPU, Intel Core i5 and 4GB RAM using MATLAB. The test is applied to a 33-bus distribution system as shown in Fig. 4. To assess the performance of the proposed method, two training datasets with different sizes are provided (200 and 1440 data resolutions). To validate the accuracy of the proposed method, the following steps are carried out:

- We analyze the state estimation of the smart grid using two training datasets (200 and 1440 samples).
- We compare the proposed SE-NN method and the exact QBBFS method.
- To validate the accuracy of the proposed method, we calculate MSE, RMSE, MAE, MAPE, and MSSE.
- For outlining the contribution of the proposed method, we calculate the execution time for producing the SE-NN output and then compare it with the execution time of the exact one.

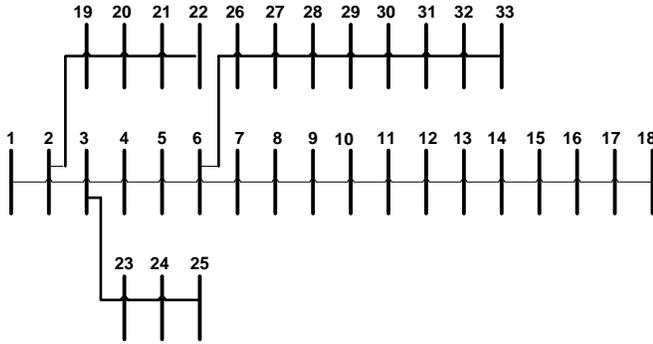


Fig. 4. The single-line diagram of the 33-bus distribution system.

A. Voltage Estimation

Fig. 5 shows a comparison between the actual voltages computed by the exact power flow method and the estimated voltage computed by the SE-NN method with data resolution of 200 for three buses (10, 25 and 32 buses). We use the per unit expression as a fraction of the base unit for simplifying the calculations. The other dataset is applied, which is a day data per minutes containing 1440 samples (24*60). Fig. 6 shows the estimated and exact voltages at 10, 25, and 32 buses. It is important to note that there is a little difference between the actual and estimated voltages for both data resolutions. This implies that the proposed method gives very accurate estimations of voltages.

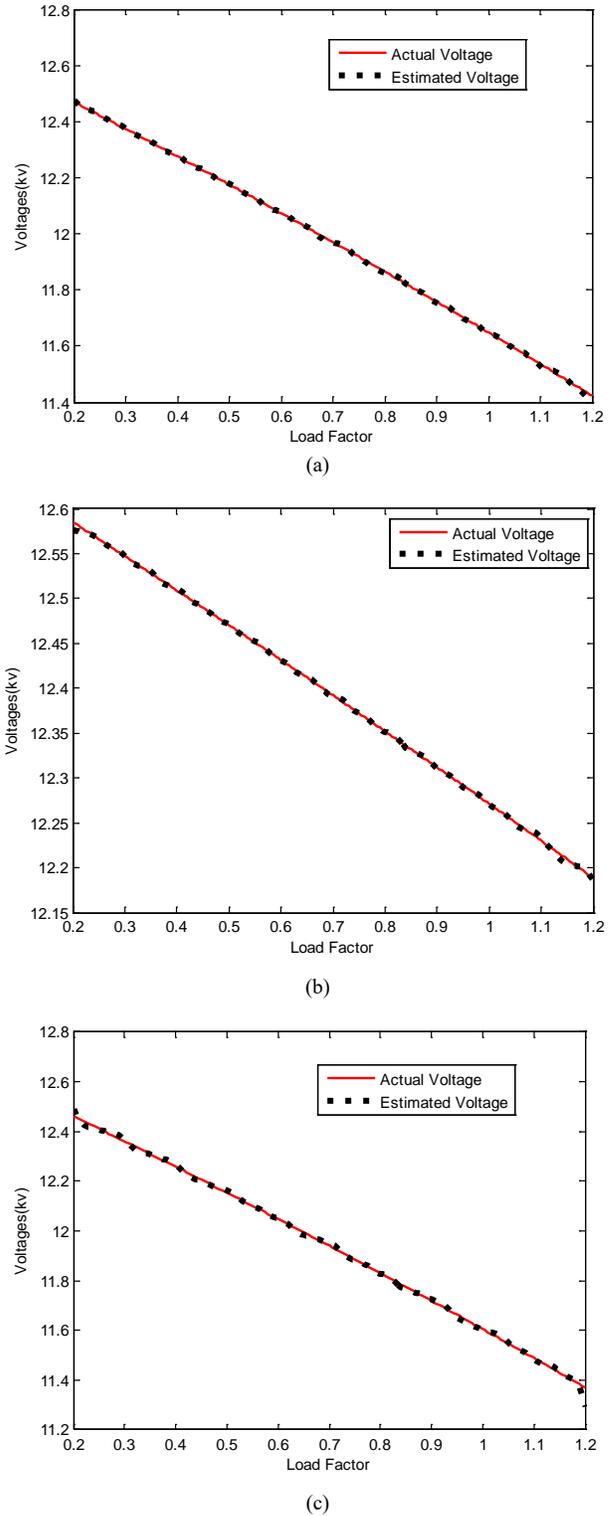
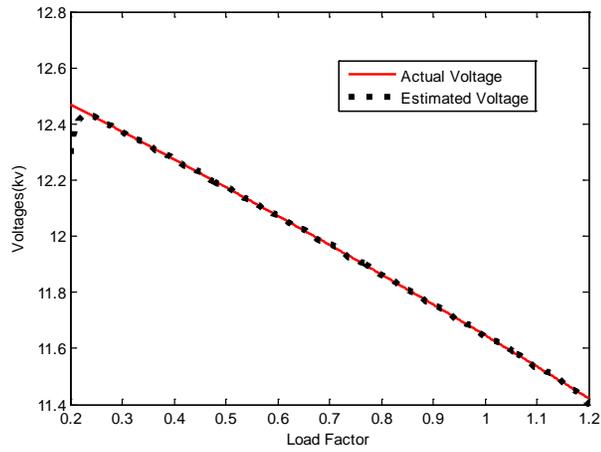
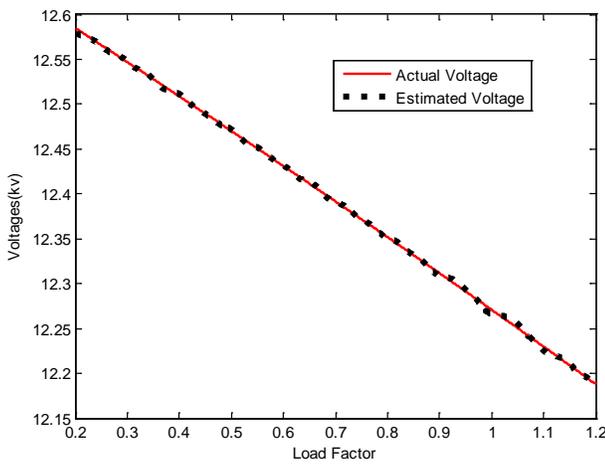


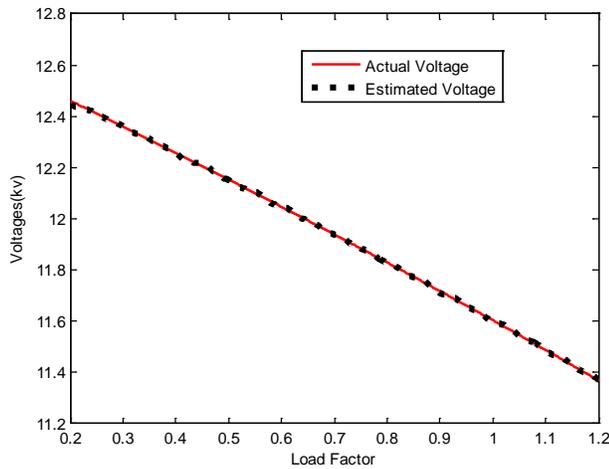
Fig. 5. The actual and estimated voltages magnitudes with 200 data resolution of different buses (a) bus 10, (b) bus 25, and (c) bus 32.



(a)



(b)



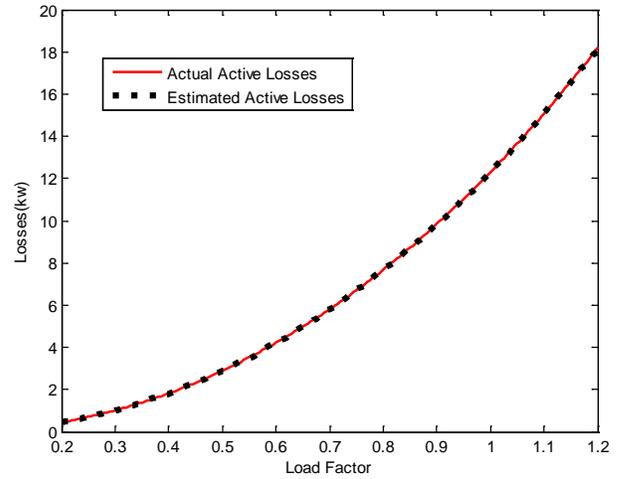
(c)

Fig. 6. The actual and estimated voltages magnitudes with 1440 data resolution of different buses (a) bus 10, (b) bus 25, and (c) bus 32.

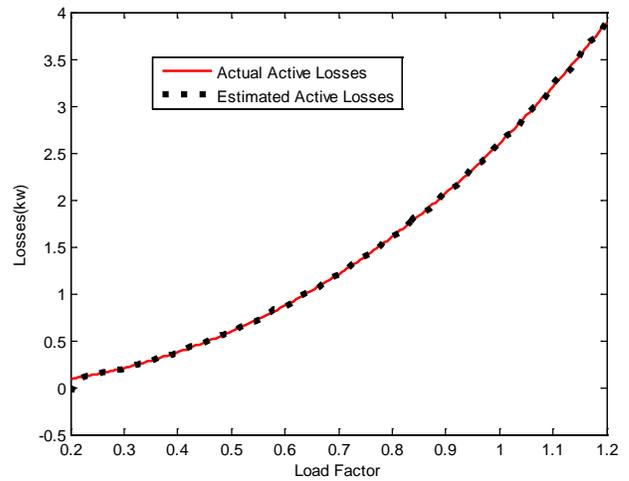
B. Active Power Estimation

The same data resolutions are applied for computing active power losses. We choose three lines (1, 25 and 30) for testing the datasets. The active power losses for the three lines with 200 and 1440 data sets are shown in Fig. 7 and Fig. 8, respectively. The total active losses for all the system are also shown in Fig. 9 for 200 and 1440 data resolutions.

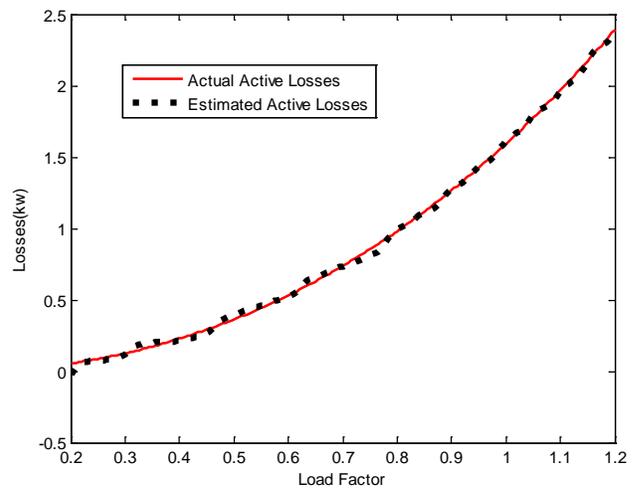
We notice a strong match between the losses computed by the proposed method and the exact one.



(a)

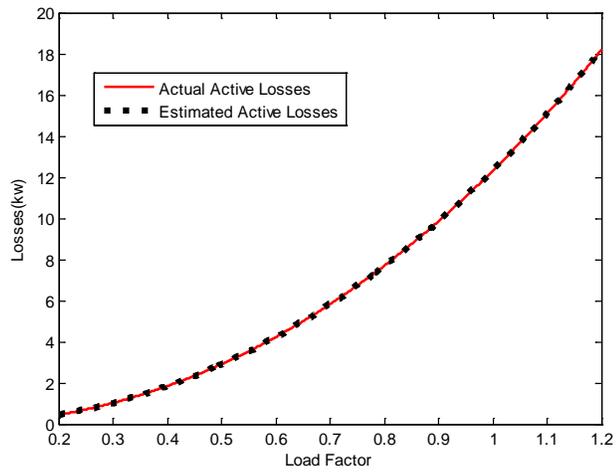


(b)

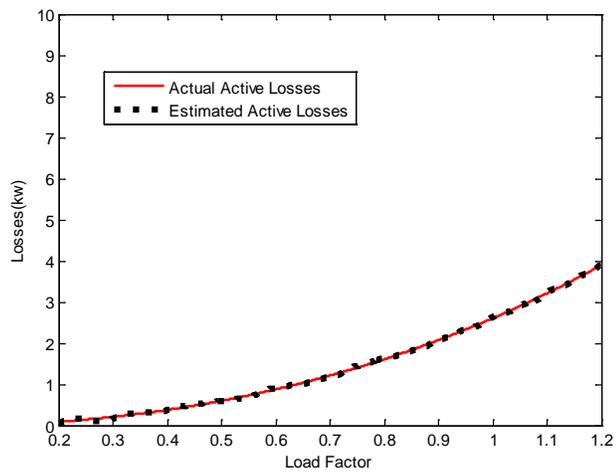


(c)

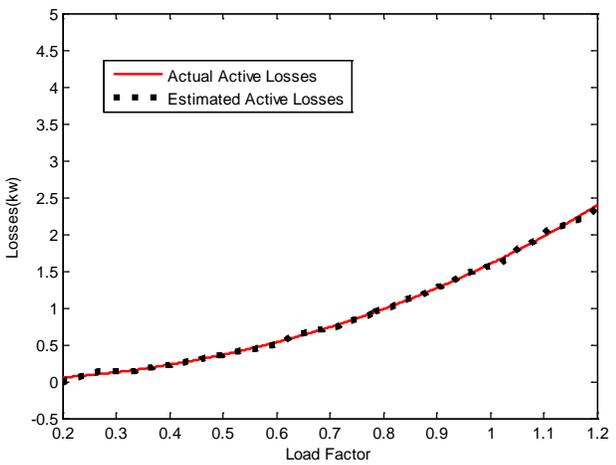
Fig. 7. The actual and estimated active power loss magnitudes with 200 data resolution of different lines (a) line 1, (b) line 25, and (c) line 30.



(a)

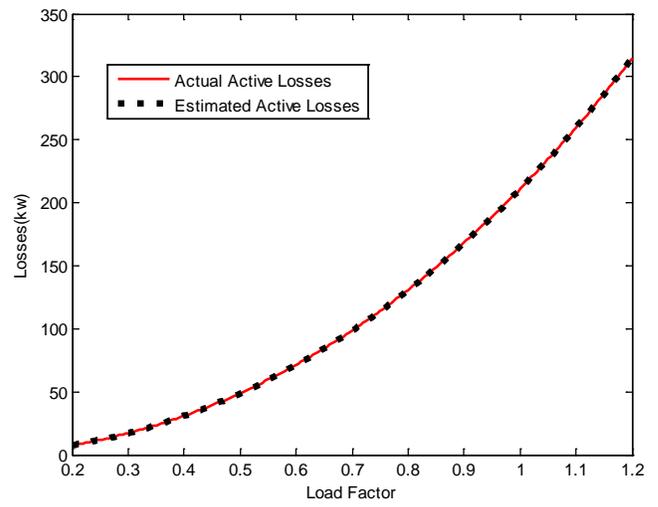


(b)

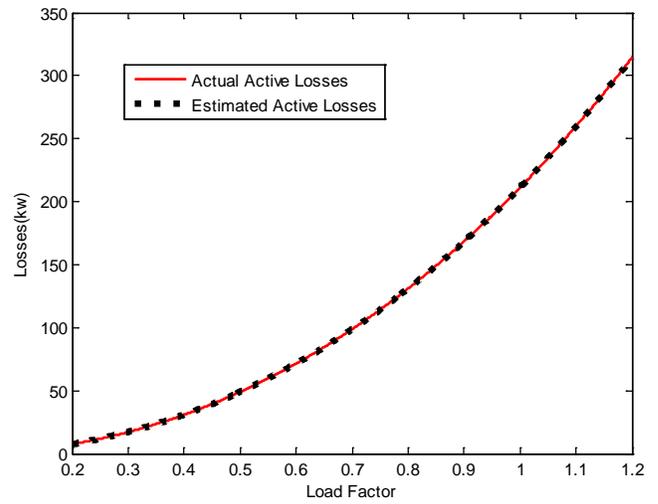


(c)

Fig. 8. The actual and estimated active power loss magnitudes with 1440 data resolution of different lines (a) line 1, (b) line 25, and (c) line 30.



(a)

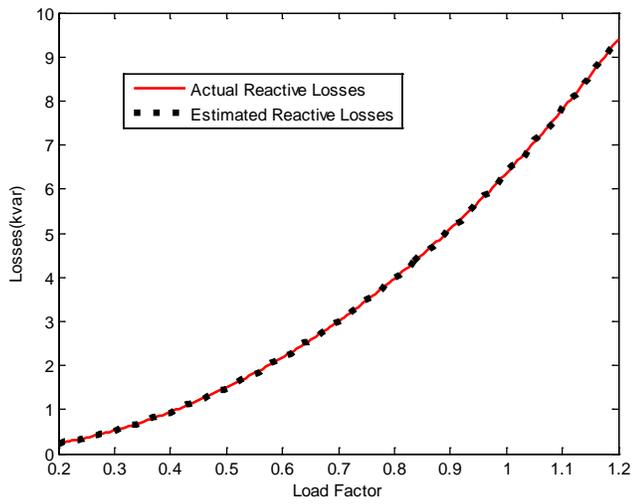


(b)

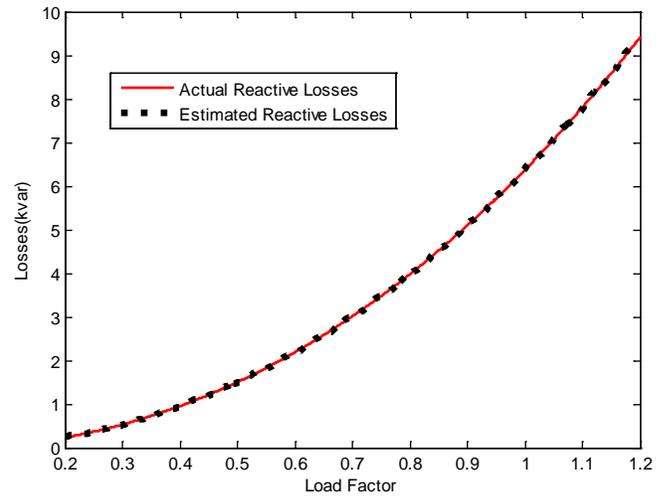
Fig. 9. The total actual and estimated active losses magnitudes for the two data resolutions (a) 200 data resolution, and (b) 1440 data resolution.

C. Reactive Power Estimation

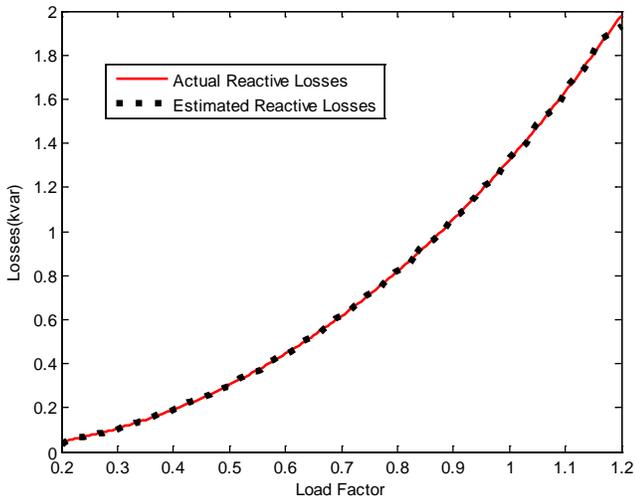
The reactive losses are also computed for testing the accuracy of the proposed method with respect to the exact method. A comparison between the actual reactive losses computed by the QBBFS method and the estimated reactive losses computed by the SE-NN method with 200 and 1440 samples at three lines (1, 25, and 30) are shown in Fig. 10 and Fig. 11, respectively. The total reactive losses in the distribution system are shown in Fig. 12 for both data resolutions. Similar to the active power loss estimation results, the reactive power estimation results of the proposed method are very close to the exact QBBFS method.



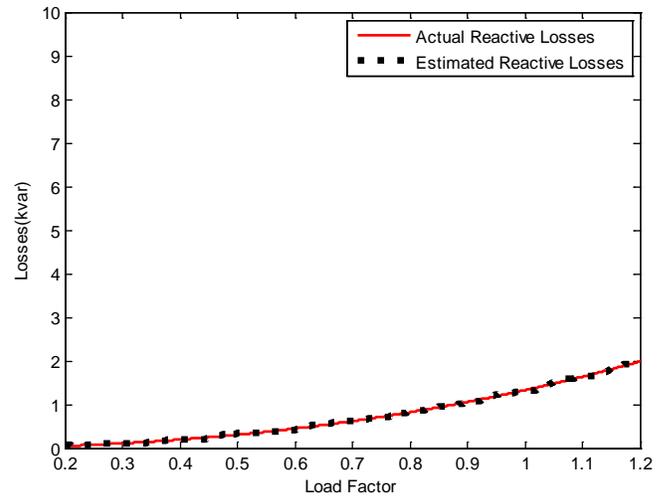
(a)



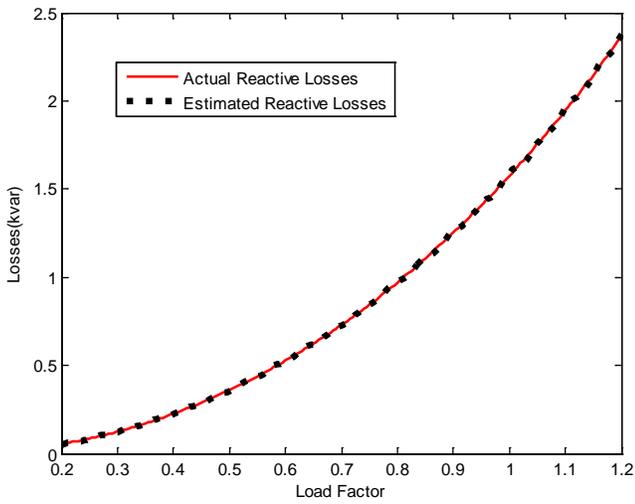
(a)



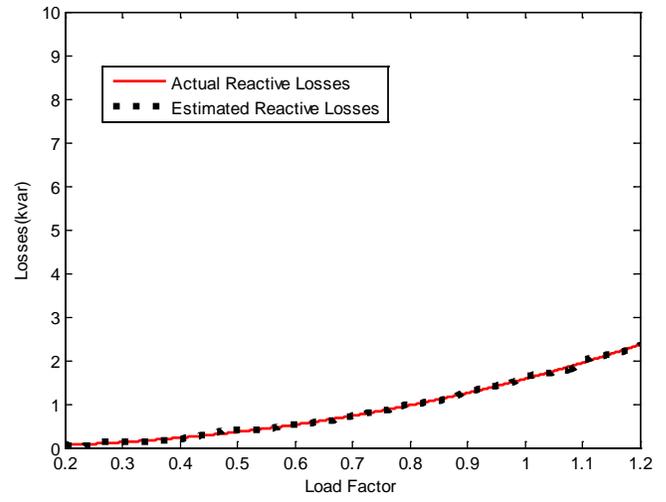
(b)



(b)



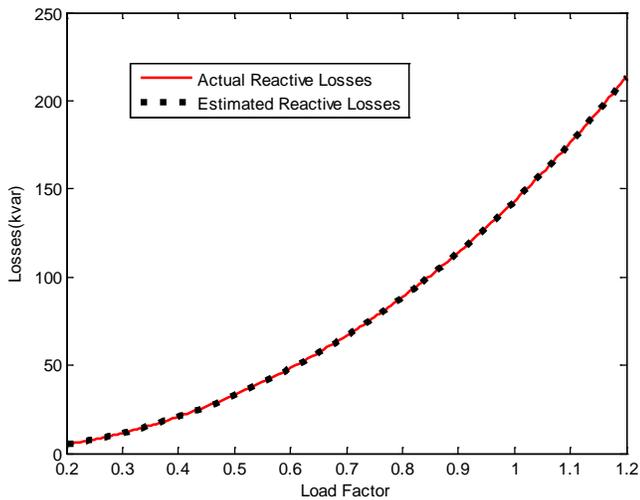
(c)



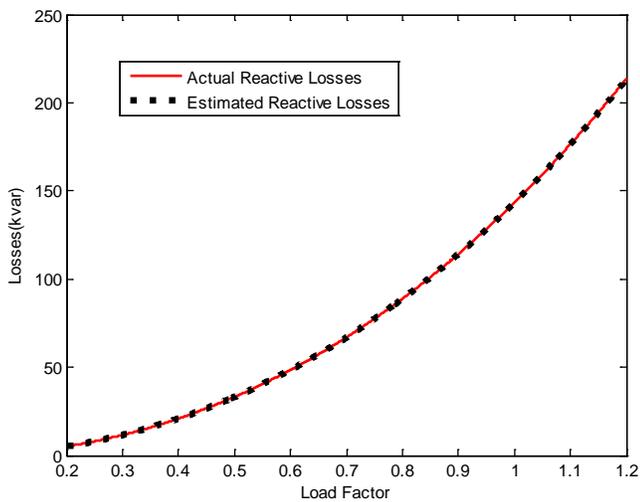
(c)

Fig. 10. The actual and estimated reactive power loss magnitudes with 200 data resolution of different lines (a) line 1, (b) line 25, and (c) line 30.

Fig. 11. The actual and estimated reactive power loss magnitudes with 1440 data resolution of different lines (a) line 1, (b) line 25, and (c) line 30.



(a)



(b)

Fig. 12. The total actual and estimated reactive losses magnitudes for the two data resolutions (a) 200 data resolution, and (b) 1440 data resolution.

D. Performance Comparison

To assess the validity of the proposed method, the different errors are computed for comparing the QBBFS and SE-NN outputs of voltages and re/active power loss. Table I demonstrates the maximum value of the MSE, RMSE, MAE, MAPE, and SSE at all busses for 200 samples. As we can see, the values of errors are very small. In the case of increasing data size to 1440, errors are still small, as illustrated in Table II. This demonstrates the high accuracy of the proposed method.

TABLE I. MAXIMUM VALUES OF ERRORS AT 200 DATA RESOLUTIONS

Errors	Voltages	Active losses	Reactive losses
MMSE	4.0765e-04	9.5776e-04	9.3902e-04
MRMSE	0.0202	0.0309	0.0306
MAE	0.2164	0.2440	0.2354
MAPE	5.8776e-04	0.3169	0.3074
MSSE	0.0819	0.1925	0.1887

TABLE II. MAXIMUM VALUES OF ERRORS AT 1440 DATA RESOLUTIONS

Errors	Voltages	Active losses	Reactive losses
MMSE	1.2011e-04	9.9229e-04	9.9254e-04
MRMSE	0.0110	0.0312	0.0315
MAE	0.0324	0.2638	0.1735
MAPE	6.4528e-04	0.5831	0.2910
MSSE	0.1730	1.4289	1.4293

For further illustrating the effectiveness and contribution of the SE-NN method, the required time for producing the output from QBBFS and SE-NN is compared. With the 200-data resolution, the QBBFS method takes longer execution time (around 16 sec) than the SE-NN method which is not exceeding 0.5 sec, as shown in Table III. For the data size of 1440, a slight change occurs in the execution time of the SE-NN (less than 0.6 sec), while the execution time of QBBFS increases dramatically (more than 115 sec), as shown in Table IV. It worth noting that the computational performance of the SE-NN method is not affected by increasing the data size as in the case of the iterative method.

TABLE III. EXECUTION TIME WITH 200 DATA RESOLUTION

Method	Voltages	Active losses	Reactive losses
QBBFS	16.546066	16.181725	16.247039
SE-NN	0.461741	0.223520	0.437750

TABLE IV. EXECUTION TIME WITH 1440 DATA RESOLUTION

Method	Voltages	Active losses	Reactive losses
QBBFS	128.972066	115.736379	116.249624
SE-NN	0.589995	0.239677	0.471077

VI. CONCLUSION

In this paper, the SE-NN method has been proposed for estimating the state of smart grid systems. The proposed method utilizes the feed-forward neural network to determine the state estimation of smart grids. The existing power flow methods take very long execution time, while the proposed method estimates the state at any point in a quick and accurate way. The experiments have been carried out at two different training sets (200 and 1440 data resolution). The proposed method has been compared with the QBBFS method for validation. SE-NN has the ability to estimate the state of smart grid rapidly with high accuracy rate. The proposed method can be a helpful tool for system operators for monitoring the real-time operation of smart grids. The future work will be directed to simulating large-scale distribution systems with renewable energy sources, such as photovoltaic and wind generation systems. In addition, we will use deep learning methods to estimate the state of smart grids

REFERENCES

- [1] C. González García, D. Meana Llorián, C. Pelayo G-Bustelo, and J. M. Cueva-Lovelle, "A review about Smart Objects, Sensors, and Actuators," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no. 3, pp. 7-10, 2017.
- [2] G. Jeon, Yun Bae Kim, and Jinsoo Park, "Agent based smart grid modelling," in *2015 Winter Simulation Conference (WSC)*, 2015, pp.

3114–3115.

[3] T. Halder, “A smart grid,” in *2014 6th IEEE Power India International Conference (PIICON)*, 2014, pp. 1–6.

[4] V. C. Gungor *et al.*, “Smart Grid Technologies: Communication Technologies and Standards,” *IEEE Trans. Ind. Informatics*, vol. 7, no. 4, pp. 529–539, Nov. 2011.

[5] H. Farhangi, “The path of the smart grid,” *IEEE Power Energy Mag.*, vol. 8, no. 1, pp. 18–28, Jan. 2010.

[6] M. Abdel-Nasser and K. Mahmoud, “Accurate photovoltaic power forecasting models using deep LSTM-RNN,” *Neural Comput. Appl.*, pp. 1–14, Oct. 2017.

[7] H. Mortazavi, H. Mehrjerdi, M. Saad, S. Lefebvre, D. Asber, and L. Lenoir, “A Monitoring Technique for Reversed Power Flow Detection With High PV Penetration Level,” *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2221–2232, Sep. 2015.

[8] A. P. Kenneth and K. Folly, “Voltage Rise Issue with High Penetration of Grid Connected PV,” *IFAC Proc. Vol.*, vol. 47, no. 3, pp. 4959–4966, 2014.

[9] G. K. Ari and Y. Baghzouz, “Impact of high PV penetration on voltage regulation in electrical distribution systems,” in *2011 International Conference on Clean Electrical Power (ICCEP)*, 2011, pp. 744–748.

[10] W. Tinney and C. Hart, “Power Flow Solution by Newton’s Method,” *IEEE Trans. Power Appar. Syst.*, vol. PAS-86, no. 11, pp. 1449–1460, Nov. 1967.

[11] K. Mahmoud and M. Abdel-Nasser, “Efficient SPF approach based on regression and correction models for active distribution systems,” *IET Renew. Power Gener.*, vol. 11, no. 14, pp. 1778–1784, Dec. 2017.

[12] Y. Zhang, R. Madani, and J. Lavaei, “Power system state estimation with line measurements,” in *2016 IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 2403–2410.

[13] H. Hooshyar and L. Vanfretti, “Power flow solution for multiphase unbalanced distribution networks with high penetration of photovoltaics,” in *2013 8th International Conference on Electrical and Electronics Engineering (ELECO)*, 2013, pp. 167–171.

[14] R. Dugan, U. S. A. Epri, and G. Ramos, “Harmonics Analysis Using Sequential-Time Simulation for Addressing,” *23th International Conference on Electricity Distribution (CIRED)*, pp. 15–18, 2015.

[15] T. Adefarati and R. C. Bansal, “Integration of renewable distributed generators into the distribution system: a review,” *IET Renew. Power Gener.*, vol. 10, no. 7, pp. 873–884, 2016.

[16] Danling Cheng, B. Mather, R. Seguin, J. Hambrick, and R. P. Broadwater, “PV impact assessment for very high penetration levels,” in *2015 IEEE 42nd Photovoltaic Specialist Conference (PVSC)*, 2015, pp. 1–6.

[17] C. D. López, “Thesis: Shortening time-series power flow simulations for cost-benefit analysis of LV network operation with PV feed-in,” 2015.

[18] K. S. Narendra and K. Parthasarathy, “Identification and control of dynamical systems using neural networks,” *IEEE Trans. Neural Networks*, vol. 1, no. 1, pp. 4–27, Mar. 1990.

[19] M. Al Shamisi, A. Assi, and H. Hejase, “Using MATLAB to Develop Artificial Neural Network Models for Predicting Global Solar Radiation in Al Ain City–UAE,” *Intechopen.Com*, pp. 219–238, 2009.

[20] P. Siano, C. Cecati, H. Yu, and J. Kolbusz, “Real Time Operation of Smart Grids via FCN Networks and Optimal Power Flow,” *IEEE Trans. Ind. Informatics*, vol. 8, no. 4, pp. 944–952, Nov. 2012.

[21] H. Le Nguyen, “Newton-Raphson method in complex form [power system load flow analysis],” *IEEE Trans. Power Syst.*, vol. 12, no. 3, pp. 1355–1359, 1997.

[22] J.-H. Teng, “A modified Gauss–Seidel algorithm of three-phase power flow analysis in distribution networks,” *Int. J. Electr. Power Energy Syst.*, vol. 24, no. 2, pp. 97–102, Feb. 2002.

[23] B. N. Rao and R. Inguva, “Power System State Estimation Using Weighted Least Squares (WLS) and Regularized Weighted Least Squares (RWLS) Method,” *International Journal of Engineering Research and Applications*, vol. 6, no. 5, pp. 1–6, 2016.

[24] A. Primadianto and C.-N. Lu, “A Review on Distribution System State Estimation,” *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3875–3883, 2016.

[25] C. Muscas, M. Pau, P. A. Pegoraro, S. Sulis, F. Ponci, and A. Monti, “Multiarea Distribution System State Estimation,” *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 5, pp.1140–1148, 2015.

[26] J. Deboever, X. Zhang, M. J. Reno, R. J. Broderick, S. Grijalva, and F. Therrien “Challenges in reducing the computational time of

QSTS simulations for distribution system analysis,” *Sandia National Laboratories*, SAND2017-5743, 2017.

[27] H. Ramchoun, M. Amine, J. Idrissi, Y. Ghanou, and M. Ettaouil, “Multilayer Perceptron: Architecture Optimization and Training,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no. 1, pp. 26–30, 2016.

[28] J. R. Machado Fernández and J. C. Bacallao Vidal, “Improved Shape Parameter Estimation in K Clutter with Neural Networks and Deep Learning,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 3, no. 7, pp. 96–103, 2016.

[29] K. Haddouch, K. Elmoutaoukil, and M. Ettaouil, “Solving the Weighted Constraint Satisfaction Problems Via the Neural Network Approach,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no. 1, pp. 56–60, 2016.

[30] S. C. Huang and Y.-F. Huang, “Learning algorithms for perceptions using back-propagation with selective updates,” *IEEE Control Syst. Mag.*, vol. 10, no. 3, pp. 56–61, Apr. 1990.

[31] N. Yorino and K. Mahmoud, “Robust quadratic-based BFS power flow method for multi-phase distribution systems,” *IET Gener. Transm. Distrib.*, vol. 10, no. 9, pp. 2240–2250, 2016.

Mohamed Abdel-Nasser



Dr. Mohamed Abdel-Nasser received his B.Sc and M.Sc degrees in Electrical Engineering from Aswan University (Egypt) and his Ph.D in Computer Engineering from Universitat Rovira i Virgili (Spain) in 2009, 2013, and 2016, respectively. He received the 2017 Marc Esteva Vivanco prize to the best PhD dissertation on Artificial Intelligence (Spain, 2017). His doctoral thesis was on the analysis of breast cancer images. He is currently an Assistant Professor at the Department of Electrical Engineering, Faculty of Engineering, Aswan University, Aswan, Egypt. He has participated in several projects funded by the Government of Spain and the Government of Egypt. He has published more than 40 papers in international conferences and journals. His research interests include the application of machine learning and deep learning to several real-world problems, including breast cancer detection, diabetic retinopathy image analysis, skin cancer detection, smart grid analysis, time-series forecasting, wireless communication, wireless link quality estimation and sediment modeling and estimation in Nasser Lake, Egypt. Dr. Abdel-Nasser is actively serving as a Reviewer to several journal and conference publications including IEEE Transactions. He is one of the founders of the International Conference on Innovative Trends in Computer Engineering.

Karar Mahmoud



Dr. Karar Mahmoud received the B.S. and M.Sc. degrees in electrical engineering from Aswan University, Aswan, Egypt, in 2008 and 2012, respectively. In 2016, he received the Ph.D. degree from the Electric Power and Energy System Laboratory (EPESL), Graduate School of Engineering, Hiroshima University, Hiroshima, Japan. Currently, he is working as an assistant professor in Aswan university. His research interests include modeling, analysis, control, and optimization of distributed systems with distributed generations.

Heba Kashef



Heba Kashef received the B.Sc. degree in Electrical Engineering from the Faculty of Engineering, Aswan University, Aswan, Egypt, in 2012. Her interests include machine learning and smart grid analysis.

Conceptual Model Development of Big Data Analytics Implementation Assessment Effect on Decision-Making

Cecilia Adrian*, Rusli Abdullah, Rodziah Atan, Yusmadi Yah Jusoh

Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 UPM Serdang (Malaysia)

Received 12 November 2017 | Accepted 14 February 2018 | Published 5 March 2018



ABSTRACT

The significance of big data advancement has benefited various organizations to leverage the potential insights and capabilities of big data in organizational performance and decision-making. However, the analytics outcome and quality of big data analytics (BDA) implementation has yet to be addressed. Therefore the aims of this paper are to identify and analyze the affecting factors and elements of BDA implementation and to propose a conceptual model for effective decision-making through BDA implementation assessment. The model is developed based on three dimensions such as performing data strategy (organization), collaborative knowledge worker (people) and executing data analytics (technology). The findings of this ongoing study proceeds with designing a proposed conceptual model with the research hypothesis and may provide a better assessment model for effective decision-making in the long run.

KEYWORDS

Big Data, Data Analysis, Decision-Level, DSS, E-assessment.

DOI: 10.9781/ijimai.2018.03.001

I. INTRODUCTION

MANY organizations have shifted to data-driven decision-making as a result of the advancements of big data benefits. The use of analytics in real-time decision-making has significant impacts on business activities and performance enhancement [1]. Most of them have benefited from big data analytics (BDA) implementation in many ways, such as information technology (IT) infrastructure, operational, managerial, strategic and organization benefits [2] [3]. The IT infrastructure benefits including using sharable and reusable IT resources that reduce operation cost for IT management and increased IT infrastructure capability. Indirectly, it helps to improve the operational activities by reducing information process cycle time, productivity and quality improvement [2]. Likewise, the analytics results used by business leaders have improved decision-making and facilitate better planning in business management activities. BDA implementation will also benefit in strategic long-term planning to support business growth and business value creation that leads to organizational performance enhancement [3][4].

Big data analytics implementation includes processes of managing the big data analytics capabilities and resources (such as technologies, people and analytics processes) [5], and transforming big data into valuable and understandable information [3] by using the analytics applications to gain insights for effective decision-making and enhance the organizational performance [6] [7]. Therefore, business leaders play an important role in achieving the goals of BDA implementation. The understanding of analytics deliverables and effective decision-making totally depends on the influencing factors. Several factors such as organizational capability, technology capability, analytics capability, talent or human capability, and the information quality are effecting the

success of its implementation [8].

Valuing the information systems (IS) investments for BDA implementation and assessing its application will be worthwhile in the organization's effort to develop new big data business model in the future, sustaining competitive advantages and to address new issues and challenges [9]. In the course of assessment during BDA implementation stage, the aim is to determine the relevance and fulfilment of objectives, sustainability of the current business model [10] and business survival for the long term. At this point in time, the assessment of BDA is only at the stage of readiness [11], which focusing only on the big data capabilities and resources. Nonetheless, the assessment of BDA implementation impact is still an understudied research topic. Therefore, the aims of this paper are: 1) to identify and analyze the affecting factors and elements of BDA implementation; and 2) to propose a conceptual model of BDA implementation assessment for effective decision-making.

The organization of this paper will be as follows: the second section focuses on the theoretical background and is followed by the third section that discusses on research methodology. The fourth section presents the discussion and findings of the affecting factors and proposes a conceptual model for BDA implementation assessment while the last section concludes the study by pinpointing the research contribution and suggesting further research work.

II. THEORETICAL BACKGROUND

There have been a number of studies that dwelled on developing BDA capabilities models by adopting resource-based view theory (RBV) [12], and the updated version of DeLone and McLean's Information Systems Success Model (ISSM) [13]. Capabilities and resources are the key components of RBV in achieving the organizational performance and sustained competitive advantage [14].

* Corresponding author.

E-mail address: cecilia.upm@gmail.com

In adopting a resource-based perspective, big data researchers have identified various BDA related capabilities and resources that serve as potential grounds of organizational performance. For example, Akter et al., 2016 [7] and Wamba et al., 2017 [15] pointed out that BDA capabilities covers management, talent and technology dimensions that positively influenced the organizational performance. Likewise, empirical findings by Gupta and colleague, 2016 [16] provided evidence on tangible, human skills and intangible resources supported by the BDA capabilities that led to market and operational performance.

Furthermore, the use of business analytics tools has enabled decision making effectiveness in health care through absorptive capability [17]. Several studies have adopted RBV theory to deepen their understanding of BDA capabilities in achieving organizational performance such as, in manufacturing [4] [18], health care [17] [19], retail [15], and non-specific domains [7] [16] [20] [21]. On the other hand, the adoption of ISSM theory focuses only on data, information and system quality factors towards organizational impact and benefits. Thus, system quality and information quality are significant to enhance business value and firm performance in big data surroundings [21].

A. Big Data Analytics Capabilities

Various concepts have been utilized in the big data literature to describe BDA capabilities. A systematic literature found that the following three concepts were the most commonly used to describe BDA capabilities. The first concept described BDA capability as the most competence to provide business insights using data management, talent (personnel context) and infrastructure (technology context) capability to transform information business into a competitive force [7] [15]. The second concept mainly used the context in health care, where BDA capabilities is regarded as having the ability to acquire, store, process and analyze a large amount of health data in various forms, and deliver meaningful information to users that allow them to discover business values and insights in a timely fashion [2]. The third concept used BDA capability to assemble, integrate, and deploy its big data-based resources in big firms [16]. This study points out the BDA capabilities factors as shown in Table I that includes management capability, organizational capability, technology capability, talent capability (data analytics personnel expertise), information processing capability and other related capabilities.

TABLE I. BIG DATA ANALYTICS CAPABILITIES FACTORS AND VARIABLES

Factors	Variables	Sources
Management capability	<ul style="list-style-type: none"> • Big data analytics planning • Investment decision-making 	<ul style="list-style-type: none"> • Coordination and control <p>[7] [15]</p>
Organizational capability	<ul style="list-style-type: none"> • Collaboration, BDA strategy • Information strategy • Top management support 	<ul style="list-style-type: none"> • Strategic groundwork • Organizational readiness <p>[22] [4] [5] [23] [24]</p>
Technology capability	<ul style="list-style-type: none"> • Infrastructure flexibility • Process integration and standardization • Integration of IT systems 	<ul style="list-style-type: none"> • Effective use of data aggregation tools • Effective use of data analysis tools • Effective use of data interpretation tools <p>[7] [15] [22] [23] [4] [5] [17]</p>
Talent capability	<ul style="list-style-type: none"> • Technical knowledge • Technology management capability • Relational knowledge • Managerial skills 	<ul style="list-style-type: none"> • Business knowledge • Analytics capabilities • People skills • Training people • Engaging people <p>[7] [15] [16] [23] [4] [5]</p>
Information processing capability	<ul style="list-style-type: none"> • Analytical capability • Patterns of care • Unstructured data analytical capability • Decision support capability • Traceability capability 	<ul style="list-style-type: none"> • Speed of decisions • Predictive analytics • Interoperability • Data analytics (data collection, data analysis and modelling, data visualization, insight generation) <p>[18] [19] [23] [2]</p>
Other capability	<ul style="list-style-type: none"> • Process-oriented dynamic capabilities 	<ul style="list-style-type: none"> • Absorptive capability <p>[15] [17]</p>

TABLE II. INTANGIBLE AND TANGIBLE RESOURCES FACTORS AND VARIABLES

Factors	Variables	Sources
<i>Tangible Resources</i>		
Data	<ul style="list-style-type: none"> • Data standardization • Data openness • Data quality • Data privacy & security 	<ul style="list-style-type: none"> • Data provisioning (include data sourcing, access, integration and delivery) <p>[16] [25] [4]</p>
Technology	<ul style="list-style-type: none"> • Data aggregation & processing • Data storage 	<ul style="list-style-type: none"> • Analytics platform • Application <p>[16] [25] [24]</p>
Basic resources	<ul style="list-style-type: none"> • Basic resources • Investment 	<ul style="list-style-type: none"> • Support (include laws & regulations, government policies) <p>[16] [25] [4]</p>
<i>Intangible Resources</i>		
Data-driven culture	<ul style="list-style-type: none"> • Data-driven culture 	<ul style="list-style-type: none"> • Decision-making culture <p>[16] [18] [22] [17] [23]</p>
Intensity of organizational learning	<ul style="list-style-type: none"> • Intensity of organizational learning 	<p>[16]</p>
Perceived benefits	<ul style="list-style-type: none"> • Perceived benefit of external data usage 	<ul style="list-style-type: none"> • Perceived benefit of internal data usage <p>[20]</p>

B. Big Data Analytics Resources

Based on RBV theory, some studies have categorized BDA resources into tangible and intangible resources. Table II listed factors and variables for tangible and intangible resources discussed in BDA studies. The big data analytic capability model developed by Gupta and George (2016) [16] equally relies on tangibles resources such as data, technology, basic resources, and intangibles resources such as data-driven culture and intensity of organizational learning. Furthermore, intangible resources also include perceived benefits of external and internal data usage [20].

C. Big Data Analytics Quality

BDA is associated with transforming and analyzing raw data into valuable information as well as knowledge in creating business values. In relation to this, quality of data and information are critical for organizational impact which will facilitate top management in decision-making. Earlier studies have shown that BDA quality factors are consisting of data quality, information quality and system quality as shown in Table III. An empirical study in BDA environment by Ji-fan Ren et al. (2016) [21] identified system quality and information quality is very crucial in enhancing business value and firms' performance.

TABLE III. BIG DATA ANALYTICS QUALITY FACTORS AND VARIABLES

Factors	Variables	Sources
Data Quality	• Data consistency • Data completeness	[20]
Information Quality	• Completeness • Currency • Format • Accuracy • Security and integrity	[5] [21]
System Quality	• Reliability • Adaptability • Integration • Accessibility • System response time • System privacy	[21]

III. METHODOLOGY

The development of the conceptual model in assessing the BDA implementation was carried out based on multiple steps as illustrated in Fig. 1. The review processes was based on the systematic literature review (SLR) approach [26][27], and continued with content analysis before the relevant theories were identified. The review exercise has investigated 15 articles on the implementation of BDA in various domains, and categorized them into two types of research analysis namely empirical and case study.

The initial investigation begins by formulating a research question such as ‘What are the factors and elements to be considered in developing BDA implementation assessment (BDAIA) model?’ Then, the process continues by searching information from electronic journal databases such as scopus, science direct, google scholar and snowballing technique, and then 15 relevant articles were selected. The relevant articles were then analyzed using matrix table (Table IV). This study has listed 4 factors and 15 elements to be considered in developing the assessment model, and the findings are discussed in the following section (Section 4). Finally, the design of the conceptual model for BDAIA with related factors and elements is presented in the conclusion.

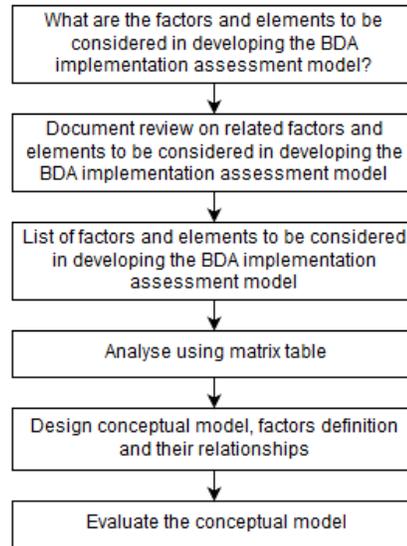


Fig. 1. The process of developing BDAIA conceptual model.

IV. FINDINGS AND DISCUSSION

Table IV presents the related BDA factors discussed in other studies which includes BDA capabilities, tangible resources, intangible resources and BDA quality, together with their elements that affects BDA implementation. The frequency of each element were shown in Fig. 2. Technology element in both BDA capability and tangible resources, and talent or human element were shown to be the two most frequent elements highlighted in BDA implementation. This is followed by organizational capability, data-driven culture, and information processing capability elements. Subsequently, elements such as data and basic resources were the least discussed in other studies. BDA elements such as management capability, other capability, BDA quality and perceived benefits were the least discussed elements in the empirical studies as they were considered still in their infancy stage in the big data environment.

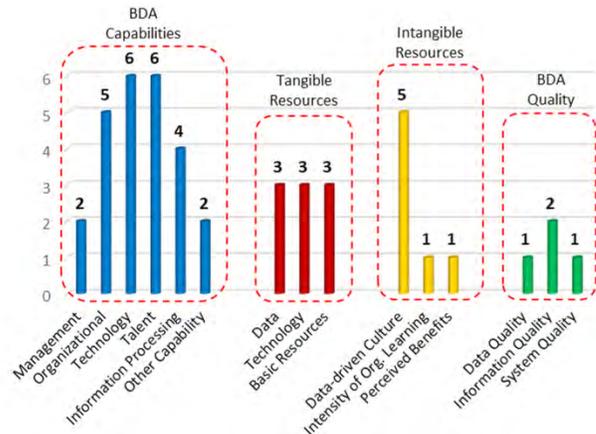


Fig. 2. The frequency of BDA elements.

Drawing on RBV theory and ISSM, big data analytics implementation assessment is conceptualized and being determined by three dimensions: organization, people and technology. Performing data strategy factor is an important criteria in organization dimension. It refers to the organizations' assurance in performing strategic analytics alignment, managerial commitment and resources management. Meanwhile, collaborative knowledge worker factor is determined by the people dimension which refers to the analytics personnel skills,

TABLE IV. FACTORS AND ELEMENTS AFFECTING THE BDA IMPLEMENTATION

Factors/ Elements	Authors	Koronios et al., 2014 [5]	Akter et al., 2016 [7]	Wamba et al., 2017 [15]	Janssen et al., 2017 [22]	Popović et al., 2016 [4]	Dutta & Bose, 2015 [23]	Chen et al., 2015 [24]	Wang & Byrd, 2017 [17]	Gupta & George, 2016 [16]	Cao & Duan, 2014 [18]	Wang & Hajji, 2017 [19]	Wang et al., 2018 [2]	Kim & Park, 2016 [25]	Kwon et al., 2014 [20]	Ji-fan et al., 2016 [21]	Frequency
BDA Capabilities																	
Management			X	X													2
Organizational		X			X	X	X	X									5
Technology		X	X	X	X		X		X								6
Talent/Human		X	X	X		X	X			X							6
Information Processing							X				X	X	X				4
Others				X					X								2
Tangible Resources																	
Data						X				X				X			3
Technology								X		X				X			3
Basic Resources						X				X				X			3
Intangible Resources																	
Data-driven Culture					X		X		X	X	X						5
Organization Learning										X							1
Perceived Benefits															X		1
BDA Quality																	
Data Quality															X		1
Information Quality		X														X	2
System Quality																X	1

TABLE V. FACTORS, ELEMENTS AND DEFINITIONS

Factors	Elements	Definition	Sources
Perform Data Strategy	• Strategic Analytics Alignment	• The alignment of BDA goals to support the business and IT strategy, and their goals.	[4] [23]
	• Managerial Commitment	• The organization leaders and top management commitment to support the BDA implementation and provide sufficient resources.	[16]
	• Resources Management	• The management of BDA resources include technology, people and competency development.	[15] [16]
Collaborate Knowledge Worker	• Analytics Skills	• The capability of analytics personnel to analyze big data effectively with the specific capability, level of knowledge and skills.	[5][7][15]
	• Organizational Relationship	• The collaboration and interaction between analytics personnel and domain experts in achieving BDA goals.	[7][22]
	• Analytics Culture	• The practice of organization leaders and decision-makers using statistical data and analytics information in decision-making.	[18][22]
Execute Data Analytics	• Infrastructures	• The technology of IT equipment's and application systems used to operationalize the BDA implementation.	[6][7]
	• Information Processing	• The capability of IT platforms and analytics applications to process raw data and transform to the valuable information.	[5][25]
	• Data Governance	• The governance of big data includes formulating data policy, data integration and data management.	[4][2]
	• Data Quality Management	• The management of data and information quality in processing big data to provide complete, accurate and quality reports.	[20][21]

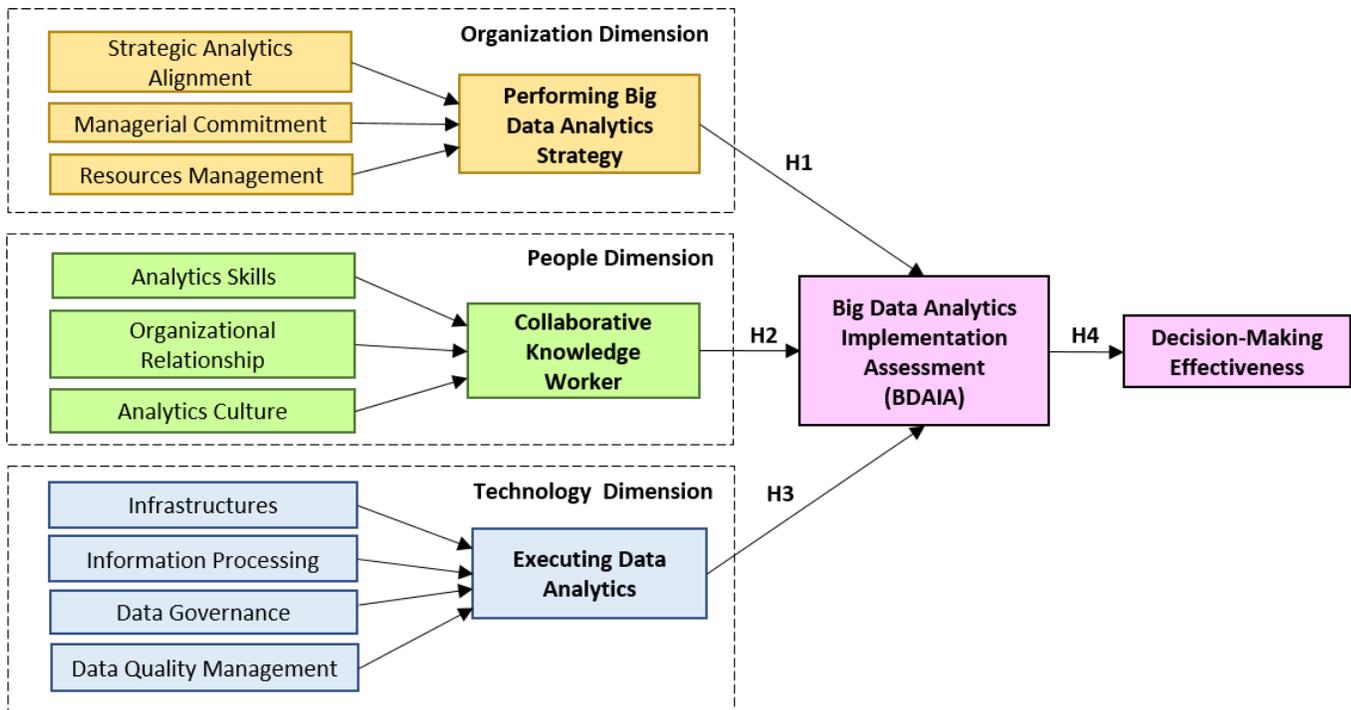


Fig. 3. Conceptual model BDA implementation assessment effect on decision-making.

organizational relationship and analytics culture. Subsequently, in technology dimension, the elements of executing data analytics factor including IT infrastructures, information processing, data governance and data quality management. Table V presents the factors with related elements and its definition that are crucial for creating the conceptual model.

Fig. 3 illustrated the conceptual model for big data analytics implementation assessment (BDAIA). The effect of BDAIA consequently influences decision-making effectiveness, as well as the indirect effect of performing BDA strategy, collaborative knowledge worker and executing data analytics factors. In this line, the study hypothesis is:

H1: Performing BDA strategy has a significant positive effect on BDAIA

H2: Collaborative knowledge worker has a significant positive effects on BDAIA.

H3: Executing data analytics has a significant positive effect on BDAIA.

H4: BDAIA has positive effects on decision-making effectiveness.

V. CONCLUSION

A conceptual model was found the most useful for assessing the big data analytics implementation. The model is capable of measuring the relationship of organization, people and technology dimensions that affect the assessment of BDA implementation and decision-making. The follow-up research activity will be developing a survey instrument using questionnaires. The proposed conceptual model and questionnaires will be then to be verified by experts from academics and industry. In relation to this, a pilot study will be conducted and to be followed by the actual study. The model will be validated then using statistical tools, and the results gained will provide an insight for business leaders to plan, sustain and enhance the capability of decision-making based on data-driven by assessing the relevant resources.

REFERENCES

- [1] R. Sharma, S. Mithas, and A. Kankanhalli, "Transforming Decision-Making Processes: A Research Agenda for Understanding the Impact of Business Analytics on Organisations," *European Journal of Information Systems*, vol. 23, no. 4, pp. 433–441 (2014).
- [2] Y. Wang, L. Kung, and T. A. Byrd, "Big Data Analytics: Understanding Its Capabilities and Potential Benefits for Healthcare Organizations," *Technological Forecasting and Social Change*, vol. 126, pp. 3–13 (2018).
- [3] Y. Wang, L. Kung, W. Y. C. Wang, and C. G. Cegielski, "An Integrated Big Data Analytics-enabled Transformation Model: Application to Health Care," *Information and Management*, vol. 55, pp. 64–79 (2018).
- [4] A. Popović, R. Hackney, R. Tassabehji, and M. Castelli, "The Impact of Big Data Analytics on Firms' High Value Business Performance," *Information Systems Frontiers*, pp. 1–14 (2016).
- [5] A. Koronios, J. Gao, and S. Selle, "Big Data Project Success - A Meta Analysis," in 18th Pacific Asia Conference on Information Systems, PACIS Proceedings, pp. 376–390 (2014).
- [6] S. Fosso Wamba, S. Akter, A. Edwards, G. Chopin, and D. Gnanzou, "How 'big data' can make big impact: Findings from A Systematic Review and A Longitudinal Case Study," *International Journal of Production Economics*, vol. 165, pp. 234–246 (2015).
- [7] S. Akter, S. F. Wamba, A. Gunasekaran, R. Dubey, and S. J. Childe, "How to improve firm performance using big data analytics capability and business strategy alignment?," *International Journal of Production Economics*, vol. 182, pp. 113–131 (2016).
- [8] C. Adrian, Rusli Abdullah, R. Atan, and Yusmadi Yah Jusoh, "Factors Influencing to the Implementation Success of Big Data Analytics : A Systematic Literature Review," in 2017 International Conference on Research and Innovation in Information Systems (ICRIIS), pp. 1–6 (2017).
- [9] S. Lavalley, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, "Big Data, Analytics and the Path from Insights to Value," *MIT Sloan Management Review*, vol. 52, no. 2, pp. 21–34 (2011).
- [10] A. Abbasi, S. Sarker, and R. Chiang, "Big Data Research in Information Systems: Toward an Inclusive Research Agenda," *Journal of the Association for Information Systems*, vol. 17, no. 2 (2016).
- [11] M. Mach-Król, "A Survey and Assessment of Maturity Models for Big Data Adoption," in 11th International Conference on Strategic Management and Its Support by Information Systems (SMSIS), pp. 391–399 (2015).
- [12] M. Wade and J. Hulland, "Review: The Resource-Based View and

Information Systems Research: Review, Extension, and Suggestions for Future Research,” MIS Quarterly, vol. 28, no. 1, pp. 107–142 (2004).

- [13] W. H. DeLone and E. R. McLean, “The DeLone and McLean Model of Information Systems Success: A Ten-Year Update,” Journal of Management Information Systems, vol. 19, no. 4, pp. 9–30 (2003).
- [14] J. B. Barney, “Firm Resources and Sustained Competitive Advantage,” Journal of Management, vol. 17, no. 1, pp. 99–120 (1991).
- [15] S. F. Wamba, A. Gunasekaran, S. Akter, S. J. Ren, R. Dubey, and S. J. Childe, “Big Data Analytics and Firm Performance: Effects of Dynamic Capabilities,” Journal of Business Research, vol. 70, pp. 356–365 (2017).
- [16] M. Gupta and J. F. George, “Toward the Development of a Big Data Analytics Capability,” Information and Management, vol. 53, no. 8, pp. 1046–1064 (2016).
- [17] Y. Wang and T. Byrd, “Business Analytics-Enabled Decision Making Effectiveness through Knowledge Absorptive Capacity in Health Care,” Journal of Knowledge Management, vol. 21, no. 3, pp. 517–539 (2017).
- [18] G. Cao and Y. Duan, “Gaining Competitive Advantage from Analytics through the Mediation of Decision-Making Effectiveness: An Empirical Study of UK Manufacturing Companies,” in PACIS Proceedings, p. 377 (2014).
- [19] Y. Wang and N. Hajli, “Exploring the Path to Big Data Analytics Success in Healthcare,” Journal of Business Research, vol. 70, pp. 287–299 (2017).
- [20] O. Kwon, N. Lee, and B. Shin, “Data Quality Management, Data Usage Experience and Acquisition Intention of Big Data Analytics,” International Journal of Information Management, vol. 34, pp. 387–394 (2014).
- [21] S. Ji-fan Ren, S. Fosso Wamba, S. Akter, R. Dubey, and S. J. Childe, “Modelling Quality Dynamics, Business Value and Firm Performance in a Big Data Analytics Environment,” International Journal of Production Research, pp. 1–16 (2016).
- [22] M. Janssen, H. Van Der Voort, and A. Wahyudi, “Factors Influencing Big Data Decision-Making Quality,” Journal of Business Research, vol. 70, pp. 338–345 (2017).
- [23] D. Dutta and I. Bose, “Managing a Big Data project: The case of Ramco Cements Limited,” International Journal of Production Economics, vol. 165, pp. 293–306 (2015).
- [24] D. Q. Chen, D. S. Preston, and M. Swink, “How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management,” Journal of Management Information Systems, vol. 32, no. 4, pp. 4–39 (2015).
- [25] M.-K. Kim and J.-H. Park, “Identifying and Prioritizing Critical Factors for promoting the Implementation and Usage of Big Data in Healthcare,” Information Development, pp. 1–13 (2016).
- [26] B. Kitchenham and S. Charters, “Guidelines for performing Systematic Literature Reviews in Software Engineering,” EBSE-2007-01 Technical Report (2007).
- [27] C. Okoli and K. Schabram, “A Guide to Conducting a Systematic Literature Review of Information Systems Research,” Sprouts: Working Papers on Information Systems, vol. 10(26), pp. 1–51 (2010).



Rodziah Atan

Rodziah Atan is currently Associate Professor in Faculty of Computer Science and Information Technology, and as Head for Halal Management and Policy Laboratory at Universiti Putra Malaysia. She obtained her PhD in Software Engineering from Universiti Putra Malaysia in 2006. She specializes in the area of Software Engineering, Software Process Modelling and Cloud Computing Services. She also has published academic books, chapter in books, more than 100 journal publications and conference papers.



Yusmadi Yah Jusoh

Yusmadi Yah Jusoh is a senior lecturer at Faculty of Computer Science and Information Technology, Universiti Putra Malaysia. She holds PhD in Information Technology (2008) and Master in Information Technology (1998) from National University of Malaysia. She obtained Bachelor of Economic, major in Economic Analysis and Public Policy (1997) from National University of Malaysia. Her research interests include Management Information System, Information Systems, Information Technology Strategic Planning, and Software Project Management.



Cecilia Adrian

Cecilia Adrian is PhD candidate in the Department of Software Engineering and Information Systems at Universiti Putra Malaysia. She has received Bachelor Engineering in Electronics Computer (1999) from Universiti Putra Malaysia and Master of Science in Computer Security and Resilience (2008) from Newcastle University, United Kingdom. Her interests lie in Information Systems and Big

Data Analytics implementation.



Rusli Abdullah

Rusli Abdullah is currently a Professor and Leader of Applied Informatics Research Group (AIRG) in the Department of Software Engineering and Information Systems, Universiti Putra Malaysia. He holds a PhD from Universiti Teknologi Malaysia (2005), Master of Science in Computer Science (1996) and Bachelor in Computer Science (1988) from Universiti Putra Malaysia. He is also an active member of Association for Information Systems (AIS). His major research interests lie in Knowledge Management and Information Systems. He has authored and co-authored over 80 journals and 63 prestigious conferences.

MSA for Optimal Reconfiguration and Capacitor Allocation in Radial/Ring Distribution Networks

Emad A. Mohamed^{1,2} *, Al-Attar Ali Mohamed², Yasunori Mitani¹

¹ Department of Electrical Engineering, Kyushu Institute of Technology, 1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka 804-8550 (Japan)

² Department of Electrical Engineering, Aswan University, 81542 Aswan (Egypt)

Received 30 January 2018 | Accepted 13 April 2018 | Published 1 June 2018



ABSTRACT

This work presents a hybrid heuristic search algorithm called Moth Swarm Algorithm (MSA) in the context of power loss minimization of radial distribution networks (RDN) through optimal allocation and rating of shunt capacitors for enhancing the performance of distribution networks. With MSA, different optimization operators are used to mimic a set of behavioral patterns of moths in nature, which allows for flexible and powerful optimizer. Hence, a new dynamic selection strategy of crossover points is proposed based on population diversity to handle the difference vectors Lévy-mutation to force MSA jump out of stagnation and enhance its exploration ability. In addition, a spiral motion, adaptive Gaussian walks, and a novel associative learning mechanism with immediate memory are implemented to exploit the promising areas in the search space. In this article, the MSA is tested to adapt the objective function to reduce the system power losses, reduce total system cost and consequently increase the annual net saving with inequity constraints on capacitor size and voltage limits. The validation of the proposed algorithm has been tested and verified through small, medium and large scales of standard RDN of IEEE (33, 69, 85-bus) systems and also on ring main systems of 33 and 69-bus. In addition, the obtained results are compared with other algorithms to highlight the advantages of the proposed approach. Numerical results stated that the MSA can achieve optimal solutions for losses reduction and capacitor locations with finest performance compared with many existing algorithms.

KEYWORDS

Radial Distribution System, Optimal Capacitor Location, Loss Reduction, Moth Swarm Algorithm.

DOI: 10.9781/ijimai.2018.05.002

I. INTRODUCTION

MOST of electrical distribution networks feed inductive loads at low voltage levels. This effect leads to higher currents and power losses accompanied by voltage drop whereas about 13% of the total power generation has been considered as line losses [1]. Therefore, these losses must be diminished to improve the power system stability and reliability, power factor and voltage profile. Connecting shunt capacitors is considered as one of the basic methods which has been used in distribution systems to solve such problems [2, 3]. However, the random locating of capacitors can cause more voltage drop and higher power losses. Moreover, the capacitor allocation problem has a combinatorial nature because capacitor locations and sizes are discrete variables [4]. Therefore, several optimization algorithms have been proposed in recent years to solve the optimal shunt capacitor placement and sizing problems in radial and ring distribution systems for maximizing their benefits. Flower pollination algorithm (FPA) [5], particle swarm optimization (PSO) [6, 7], discrete particle swarm optimization (DPSO) [8], genetic algorithm (GA) [9], teaching-learning-based optimization (TLBO) [10], artificial bee colony (ABC) [11], cuckoo search algorithm (CSA) [12], gravitational search algorithm (GSA) [13], modified monkey search (MMS) [14], whale

optimization algorithm (WOA) [15], improved harmony algorithm (IHA) [16], fuzzy-GA [17], direct search algorithm (DSA) [18], differential evolution algorithm (DEA) [19], simulated annealing (SA) [20], plant growth simulation algorithm (PGSA) [21], fuzzy reasoning (FRB) [22], Analytical IP [23], improved binary particle swarm optimization (IBPSO) [24], Mixed-integer nonlinear programming (MINLP) [25] and fuzzy real coded genetic algorithm (FRCGA) [26] have been proposed to solve the capacitor allocation problem. However, some of these algorithms are not highly effective as the power losses still have high values [8, 9]. Other algorithms appear to be effective, but they may not achieve the optimal cost value [5, 10].

Al-Attar et. al [27] has proposed a new optimization technique called the moth swarm algorithm (MSA) which is inspired from the orientation of moths towards moonlight. This algorithm is developed based on the conventional moth flame algorithm by enhancing its exploitation and exploration by applying adaptive cross over levy mutation with associative learning mechanism. It is clear from the literature review that the MSA technique has not been applied to solve the problem of optimal capacitor location in the RDN. Hence, the authors propose to use the MSA method for dealing with the mention problem.

In this paper, MSA is presented to minimize the system power losses, decrease the total cost and maintain the voltage profile for various electrical distribution systems. It is tested on multiple IEEE standard distribution systems i.e., (33 and 69-bus). Furthermore, it is tested on the mesh distribution systems which have two ways

* Corresponding author.

E-mail address: q595506a@mail.kyutech.jp

between generation and consumers and this is more complicated in design and requires complex protection schemes which includes higher investment than RDN. In addition, the obtained results from the proposed approach are compared with those obtained from other algorithms to confirm its superiority. The rest of this work is organized as follows; section.2 provides the objective function formulation. MSA algorithm is represented in section 3. In section.4, the implementing of MSA algorithm for solving the capacitor allocation problem has been presented. Section 5 shows the numerical results of the proposed technique applied on multiple IEEE standard systems. The last section concludes the results and advantages of the proposed method.

II. PROBLEM FORMULATION

A. Load Flow Calculation

RDN creates some negative conditions such as radial meshed networks, unbalanced operation, high R/X ratios and distributed generation. Due to these problems, the Newton Raphson, Gauss Siedel and other conventional load flow algorithms are not effective to solve the load flow calculation of the distribution systems [28]. Therefore, the modern algorithm called backward/forward sweep [28] is used in this work to analyze the power flow in the tested IEEE distribution systems. The line current I_k is calculated from (1) as follows:

$$|I_k| = \left[\frac{P_k + Q_k}{|V_k|} \right] \quad (1)$$

The active power flow (P_{k+1}) and reactive power flow (Q_{k+1}) in RDN are calculated by (2) and (3) derived from single-line diagram as shown in Fig. 1.

$$P_{k+1} = P_k - P_{L(k+1)} - R_k |I_k|^2 \quad (2)$$

$$Q_{k+1} = Q_k - Q_{L(k+1)} - X_k |I_k|^2 \quad (3)$$

where k is the sending end and $k+1$ is the receiving end. Voltages of a transmission line and real power losses in the line can be calculated from (4), (5), and (6) respectively:

$$|V_{k+1}|^2 = |V_k|^2 - 2(R_k P_k + X_k Q_k) + (R_k^2 + X_k^2) |I_k|^2 \quad (4)$$

$$P_{loss(k,k+1)} = R_k |I_k|^2 \quad (5)$$

$$Q_{loss(k,k+1)} = X_k |I_k|^2 \quad (6)$$

The total system loss is calculated by summing all line losses in the system as shown in (7):

$$P_{Tloss} = \sum_{k=1}^{n-1} P_{loss(k,k+1)} \quad (7)$$

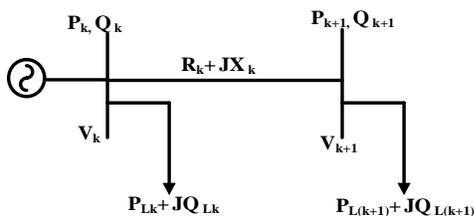


Fig. 1. Simple radial distribution system.

B. Objective Functions

The main aim of the objective function of the optimal capacitor placement problem is to minimize the total cost per year by reducing the real power losses and the cost of installing capacitors subjected to voltage and reactive power limits. This paper uses the weighted sum method to evaluate the effectiveness of the proposed approach to find the benefits of optimal allocation and rating of shunt capacitors. The weighted sum method allows the multi-objective to be cast as a single-objective mathematical optimization problem resulting in only one solution, in addition to its lower computational cost (CPU-time). These advantages are more proper for real world problems. Hence, the multi-objective functions have been performed by using the following mathematical statement:

$$f = \text{Min}(F_1 + F_2) \quad (8)$$

where F_1 and F_2 are described as: $F_1 = \text{min}(P_{Tloss})$, $F_2 = \text{min}(Cost)$

where the cost function is defined as:

$$Cost = K_e \sum_{j=1}^l T_j P_j + \sum_{k=1}^n K_{fc} Q_{fc} \quad (9)$$

C. Constraint Conditions

The objective function is subjected to:

1) Voltage Constraint

The buses voltages are the inequality constraints. The bus voltage magnitude of each bus must be maintained within the following range:

$$V_{\min} \leq |V_k| \leq V_{\max} \quad (10)$$

where V_{\max} and V_{\min} are the maximum and minimum values of bus (k) voltages. The lower and upper values are taken as 0.9 and 1.05 Pu, respectively.

2) Total Reactive Power Constraint

The total injected reactive power, which represents the equality constraints must be limited by:

$$Q_{fc} \leq \sum_{k=1}^n Q_{Lk} \quad (11)$$

$$Q_{\text{sys}} + Q_{\text{cap}} = Q_d + Q_{Tloss} \quad (12)$$

Power-flow equations, equality restrictions (2) and (3), can be satisfied during the process of power-flow calculation. In the encoding period, the inequality restrictions (10)–(12) can be satisfied through adding penalty function into the objective function in such a way that it penalizes any violation of the constraints. Consequently, the constrained optimization problem is then converted into an unconstrained form.

III. OVERVIEW OF MSA

The moth swarm algorithm has been presented in 2017 by Al-Attar et. al [27]. It is inspired from the orientation of moths towards moonlight. The available solution of any optimization problem using MSA is performed by the light source position, and its fitness is the luminescence intensity of the light source. Furthermore, the proposed method consists of three main groups, the first one is called pathfinders which is considered a small group of moths (n_p) over the available space of the optimization. The main target of this group is to guide the locomotion of the main swarm by discriminating the best positions as light sources. Prospectors group is the second one which have a

tendency to expatiate in a non-uniform spiral path within the section of the light sources determined by the pathfinders. The last one is the onlookers, this group of moths move directly to the global solution which has been acquired by the prospectors.

The steps of the MSA technique are discussed as follows:

A. Initialization

Initially, the positions of moths are randomly created for dimensional (d) and population number (n) as seen in (13).

$$x_{ij} = rand[0,1] \cdot (x_j^{\max} - x_j^{\min}) + x_j^{\min} \forall i \in \{1,2,\dots,n\}, \\ j \in \{1,2,\dots,d\} \quad (13)$$

where, x_j^{\max} and x_j^{\min} are the upper and lower limits, respectively.

Afterwards, the type of each moth is selected based on the determined fitness. Consequently, the best moths are elected as light sources and the following groups of moths (i.e., the best and worse) will be dealing as prospectors and onlookers, respectively.

B. Reconnaissance Phase

The moths may be concentrated in the regions, which seem to be a good performance. Therefore, the swarm quality for reconnaissance may be decreased during the process of the optimization and this process may lead to a stagnation case. To avoid the early convergence and enhance the solution diversity, a part of the swarm is compelled to determine the less congested area. The moths, which perform this role, update their positions by interacting with each other.

A new strategy for the diversity of solutions is presented to choose the crossover points. Firstly, the normalized dispersal degree σ_j^t of the individuals is measured as follows:

$$\sigma_j^t = \frac{\sqrt{\frac{1}{n_p} \sum_{i=1}^{n_p} (x_{ij}^t - \bar{x}_j^t)^2}}{\bar{x}_j^t} \quad (14)$$

$$\bar{x}_j^t = \frac{1}{n_p} \sum_{i=1}^{n_p} x_{ij}^t$$

where,

Then, the coefficient of variation, which measures the relative dispersion, is calculated as:

$$\mu^t = \frac{1}{d} \sum_{j=1}^d \sigma_j^t \quad (15)$$

Any element of the pathfinder moths exposed to a low dispersal degree will be taken in the group of crossover points C_p , as described below: $j \in C_p$ if $\sigma_j^t \leq \mu^t$

To complete the full trail solution, each host vector (i.e., pathfinder solution) will update the position through the crossover processes by integrating the modified variables of the sub-trail solution into the analogical variables. The full trial solution V_{pj}^t can be defined as:

$$V_{pj}^t = \begin{cases} v_{pj}^t & \text{if } j \in C_p \\ x_{pj}^t & \text{if } j \notin C_p \end{cases} \quad (16)$$

C. Lévy Flights

Lévy flights/motions are random processes based on α -stable distribution with ability to travel over large scale distances using different size of steps. Lévy α -stable distribution strongly linked with heavy-tailed probability density function (PDF), fractal statistics, and anomalous diffusion. The PDF of the individual jumps $\lambda(q) \approx |q|^{-1-\alpha}$ decaying at large generated variable q . The stability/tail index

$\alpha \in [0, 2]$ or so called the characteristic exponent describes the shape of the distribution taper [27]. There are a few special cases that have a close form for the density of the general Lévy distribution, and can be defined as:

- Gaussian or normal distribution, $q \approx N(\mu, \sigma_G^2)$ if density is:

$$f(q) = \frac{1}{\sqrt{2\pi\sigma_G}} \exp\left(-\frac{(q-\mu)^2}{2\sigma_G^2}\right) \quad -\infty < q < \infty \quad (17)$$

- Cauchy distribution, $q \approx cauchy(\sigma, \mu)$ if density is:

$$f(q) = \frac{1}{\pi(\sigma^2 + (q-\mu)^2)} \quad -\infty < q < \infty \quad (18)$$

- A simple version of Lévy distribution, $q \approx Levy(\sigma, \mu)$ if density is:

$$f(q) = \sqrt{\frac{\gamma}{2\pi}} \frac{1}{(q-\mu)^{3/2}} \exp\left(-\frac{\sigma}{2(q-\mu)}\right) \quad 0 < \mu < q < \infty \quad (19)$$

Mantegna's algorithm [27] is used to emulate the α -stable distribution by generating random samples L_i that have the same behavior of the Lévy-flights, as follows:

$$L_i \approx step \oplus Levy(\alpha) \approx 0.01 \frac{u}{|y|^{1/\alpha}} \quad (20)$$

where, step is the scaling size related to the scales of the interest problem, \oplus is the entrywise multiplications, $u = N(0, \sigma_u^2)$ and $y = N(0, \sigma_y^2)$ are two normal stochastic distributions with

$$\sigma_u = \left[\frac{r(1+\alpha)\sin\pi\alpha/2}{r((1+\alpha)/2\alpha^{(\alpha-1)/2})} \right], \quad \sigma_y = 1.$$

D. Difference Vectors Lévy-Mutation

For $n_c \in C_p$ crossover operations points, the proposed algorithm creates the sub-trial vector $v_p = [v_{p1}, v_{p2}, \dots, v_{pn}]$ by perturbing the selected components of the host vector $x_p = [x_{p1}, x_{p2}, \dots, x_{pn}]$, with related components in the donor vectors (e.g. $x_{r1} = [x_{r11}, x_{r12}, \dots, x_{r1n}]$). The Mutation strategy may be used for synthesis such a sub-trial vector, as follows:

$$\vec{v}_p^t = \vec{x}_{r1}^t + L_{p1}(\vec{x}_{r2}^t - \vec{x}_{r3}^t) + L_{p2}(\vec{x}_{r4}^t - \vec{x}_{r5}^t) \\ \forall r^1 \neq r^2 \neq r^3 \neq r^4 \neq r^5 \neq p \in \{1,2,\dots,n_p\} \quad (21)$$

where, L_{p1} and L_{p2} are two independent identical variables used as the mutation scaling factor and generated by a heavy tail Lévy-flights using ($L_p \sim random(n_c) \ominus Levy(\alpha)$). The set of mutually indices ($r1, r2, r3, r4, r5$, and p) are exclusively selected from the pathfinder solutions.

E. Selection Strategy

The fitness value of the full trail solution is determined after finishing the last procedure, and then it is compared with its corresponding host solution. The suitable solutions are selected to continue for the next generation, which is used for minimization problems as follows:

$$\vec{x}_p^{t+1} = \begin{cases} \vec{x}_{pj}^t & \text{if } f(\vec{V}_p^t) \geq f(\vec{x}_p^t) \\ \vec{v}_{pj}^t & \text{if } f(\vec{V}_p^t) < f(\vec{x}_p^t) \end{cases} \quad (22)$$

The probability P_p which is proportional to luminescence intensity f_{ip} can be calculated from (23) and f_{ip} is estimated from the objective function value f_p with minimization problems from (24).

$$P_p = \frac{fit_p}{\sum_{p=1}^{n_p} fit_p} \quad (23)$$

$$fit_p = \begin{cases} \frac{1}{1+f_p} & \text{for } f_p \geq 0 \\ 1+|f_p| & \text{for } f_p < 0 \end{cases} \quad (24)$$

F. Transverse Orientation

The prospector moths are the next best luminescence intensity group of moths. The number of prospectors n_f is proposed to decrease through all iterations T as follows:

$$n_f = \text{round}\left((n - n_p) \times \left(1 - \frac{t}{T}\right)\right) \quad (25)$$

After the pathfinders have finished their search, the information about luminescence intensity is shared with prospectors, which attempt to update its positions in order to discover new light sources. Each prospector moth x_i is soared into the logarithmic spiral path as shown in Fig. 2(a) to make a deep search around the artificial light source x_p , which is chosen on the basis of the probability P_p using (23). The new position of i th prospector moth can be expressed mathematically as follows:

$$x_i^{t+1} = |x_i^t - x_p^t| e^{\theta} \cdot \cos 2\pi\theta + x_p^t \quad \forall p \in \{1, 2, \dots, n_p\} \\ ; \forall i \in \{n_p + 1, n_p + 2, \dots, n_f\} \quad (26)$$

where, $\theta \in [r, 1]$ is a random number to define the spiral shape and $r = 1 - t/T$. Although the same formula has been used in Moth-flame Optimization (MFO) [27] algorithm, the MSA is dealing with each variable as an integrated unit. In the MSA model, the moths are changed dynamically. Therefore, any prospector moth uplifts to become pathfinder moth if it discovers a solution with luminescence more than the existing light sources. That means the new lighting sources and moonlight will be presented at the end of this stage.

G. Emphyreal Navigation

The diminishing of the number of prospectors during the optimization process increases the onlookers number ($n_o = n - n_f - n_p$). This may lead to an increase in the speed of the convergence rate of MSA towards the global solution. The onlookers are the moths that have the lowest luminescent sources in the swarm. Their main aim for traveling directly to the moon is the most shining solution as shown in Fig. 2(b). In the MSA, the onlookers are forced to search for the hot spots of the prospectors effectively. These onlookers are divided into the two following parts:

The first part, with the size of $n_G = \text{round}(n_o / 2)$, walks according to Gaussian distributions using (5). The new onlooker moth in this subgroup x_i^{t+1} moves with series steps of Gaussian walks, which can be described as follows:

$$x_i^{t+1} = x_i^t + \varepsilon_1 + [\varepsilon_2 \times best_g^t - \varepsilon_3 \times x_i^t] \\ \forall i \in \{1, 2, \dots, n_G\} \quad (27)$$

$$\varepsilon_1 \sim \text{random}(\text{size}(d)) \oplus N\left(best_g^t, \frac{\log t}{t} \times (x_i^t - best_g^t)\right) \quad (28)$$

Where, ε_j is a random number generated from Gaussian distribution,

ε_2 and ε_3 are random samples drawn from a uniform distribution within the interval $[0, 1]$, $best_g^t$ is the global best solution (moonlight) obtained in the transverse orientation phase. Based on many optimization algorithms, there is a memory to transfer information from the current generation to the next generation. However, the moths may fall into the fire in the real world due to the lack of an evolutionary memory. The performance of moths is intensely affected by the short-term memory and the associative learning [27]. The associative learning has an important role in connection among moths. Therefore, the second part of onlooker moths $n_A = n_o - n_G$ will sweep towards the moon light depending on the associative learning operators with an instantaneous memory to imitate the actual behavior of moths in nature. The instantaneous memory is initialized from the continuous uniform of Gaussian distribution on the range from $x_i^t - x_i^{\min}$ to $x_i^t - x_i^{\max}$. The updating equation of this type can be completed in form:

$$x_i^{t+1} = x_i^t + 0.001 \cdot G[x_i^t - x_i^{\min}, x_i^{\max} - x_i^t] + \left(1 - \frac{g}{G}\right) \cdot r_1 \\ \cdot (best_p^t - x_i^t) + 2g / G \cdot r_2 \cdot (best_g^t - x_i^t) \quad \forall i \in \{1, 2, \dots, n_A\} \quad (29)$$

where, r_1 and r_2 are random number within the interval $[0, 1]$, $2g/G$ is the social factor, $1-g/G$ is the cognitive factor and $best_p^t$ is a light source selected from the modified swarm based on the probability P_f . It is worth mentioning that the constraints are checked and satisfied after each fitness evaluation in the flowchart of MSA (see Fig. 3).

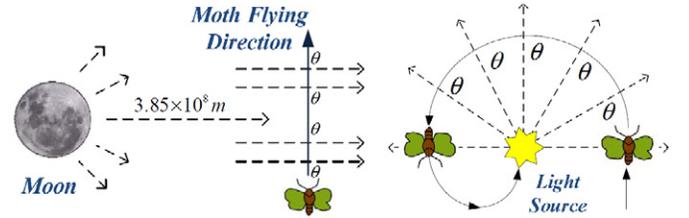


Fig. 2 Orientation behavior of moth swarm: (a) Moth flying in a spiral path into nearby light source (b) Moth flying in a fixed angle relative to moonlight.

IV. NUMERICAL EXPERIMENTS OF MSA

In order to tune the parameters of the proposed MSA and evaluate its performance in terms of exploitation, exploration, convergence behavior and solution quality, a set of 23 benchmark functions commonly used in literature were tested. The details of these functions are given in [29]. In this section, a swarm of 50 moth with seven pathfinders has been employed over 50 independent runs with a 1000 maximum number of function evaluations for $f_1 - f_{13}$, and 500 iteration for $f_{14} - f_{23}$. MSA is compared with four metaheuristics algorithms, including MPSO [30], Modified Differential Evolution (MDE) [31] approach, MFO [32], and Flower pollination algorithm (FPA) [5], respectively. To maintain comparison consistency, these algorithms are tested with 50-population size under the same conditions and using their standard control-parameters setting as given in Table I. The mean and the standard deviation are used in order to assess the robustness of the algorithms under study.

A. Determination Control Parameters in MSA

In nature, light can be dangerous and a large number of artificial lights will decrease the flight activity of moths. A statistical study has been used to specify the required number of pathfinders, and the obtained results for a swarm of 50 moths at different values of n_p are illustrated at Table II appendix (A). Judging from Table II, it can be seen that, the best required number of pathfinders is approximately 13% of the total populations.

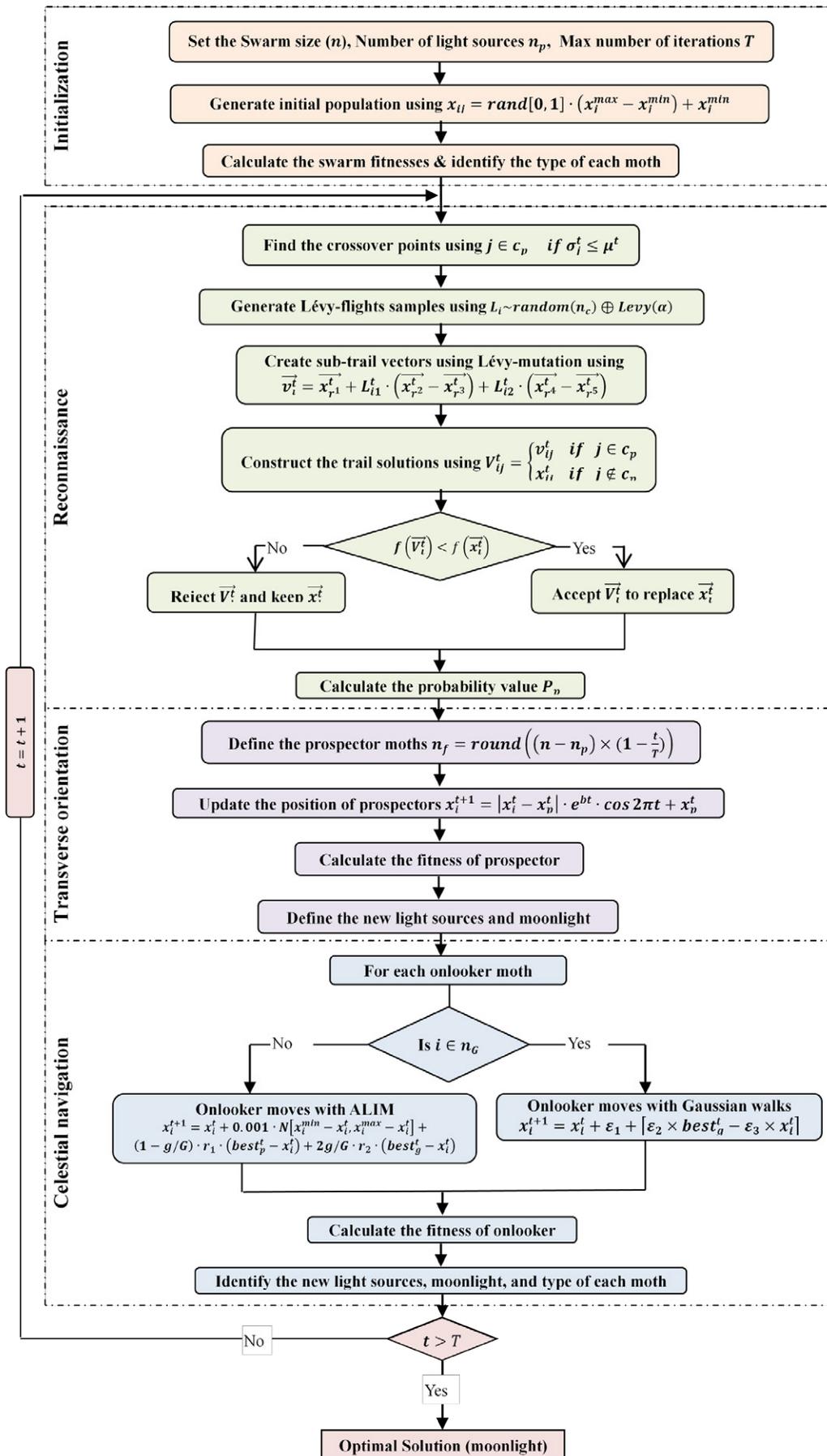


Fig. 3. Flowchart of the proposed MSA.

B. Exploitation Analysis Based on Unimodal Benchmark Functions

The first set of experiments aimed to benchmark the exploitation ability of the proposed MSA. The unimodal function ($f_1 - f_7$), are designed to compare the convergence rate of the search algorithms. In the MSA, pathfinders and prospectors primarily carry out the exploration (global search). The mean and the standard deviation (noted as StDev) are performed as reported in Table III in appendix A. According to the overall rank, although the MSA and MPSO are satisfied the condition of convergence rate and significantly better than other metaheuristic algorithms, the MSA is stronger than MPSO in fine tuning around the global optimum due to its better global search ability. On the other hand, MFO is mainly searched in a small local neighborhood. In addition, the widespread step of the FPA is not a guarantee for obtaining the advanced order.

C. Exploration Analysis Based on Multimodal Benchmark Functions

A test suite has been employed to compare MSA performance with other algorithms at the high-dimensional multimodal functions ($f_8 - f_{13}$), and the final results are summarized in Table IV in appendix (A). It is obvious that, MSA and MDE are clearly escaped from the poor local optimum, and the GMSA approaches the neighborhood of the global optimum at f_8 and hits the exact optimum every time at ($f_9 - f_{11}$). On the other hand, the Lévy-flights updating strategy of FPA maintains a small protection against the premature convergence; whereas the MFO has a low probability to make such a long jumps, which may be the reason for its poor average best fitness.

The experimental study for the low-dimensional multimodal functions ($f_{14} - f_{23}$), given in Table V appendix (A), shows that the MSA and MDE have the best results compared to the rest of the algorithms, while MPSO has difficulties with functions of this kind. Although f_{18} is an easy problem, the GMSA has failed to find the global optimum solution as other algorithms. In the three Shekel functions ($f_{21} - f_{23}$), FPA obtains a better average performance than the other optimizers. In sum, the algorithms achieve a similar performance ranking for both multimodal categories, where MSA is ranked 1st followed by MDE, MPSO, MFO, FPA, respectively. To validate the comparative study, the pairwise Wilcoxon's rank-sum test, a nonparametric statistical test, is carried out at 0.05 significance level to judge whether the results of the GMSA differ from the other algorithms in a statistical method. The p -values of the Wilcoxon's rank-sum, based on outcomes of Tables III-V in appendix (A), are displayed in Table VI appendix (A). In this table, the p -values that are less than 0.05 proved a sufficient evidence against the null hypothesis.

In order to verify the solution quality and further assess the robustness of the proposed algorithms, the graphical analysis of the Analysis of Variance (ANOVA) test for functions f_4, f_8 , and f_{21} are used, as depicted in Fig. 4. The boxplots confirm that MSA achieves, on average, superiority in comparison with the rest of the algorithms.

D. Analysis of the Convergence Behavior

The algorithms under study have been executed on 50 independent runs in order to assess their robustness through the mean and the standard deviation. To investigate the convergence behavior of the best evolution curves for the proposed methods are seen in Fig. 5. Generally, MSA and MPSO have smooth curves with a faster convergence rate more than the other algorithms. Whereas, the MPSO suffers from a premature convergence, caused by particles stagnating around local optima, when handling nonlinear functions. MFO has linear characteristics, meanwhile suffers from excessively slow rate as in f_{10} . In other hand, the MDE and FPA have non-smooth convergence characteristics. We can say that, the developed optimization algorithm

is a deep-PSO, fast-MFO and linear convergence of MDE and FPA. Seven benchmark functions are used to assess the convergence speed of the MSA against the four techniques under study and the well-known PSO [33], group search optimizer (GSO) [33] and modified group search optimizer (MGSO) [34] as shown in Table VII appendix (A). It is clear that MFO gives less computational time cost than the other algorithms. This is because MFO owns one updating equation even though it applies to each component of variables. Although, the MSA contains a number of strategies, but each of them apply to a certain group of the population. Except the small group pathfinders, all moths are dealing with each variable as an integrated unit. In addition, no longer need to store the velocities and personal best solution for each onlooker moth, as in the basic PSO. These properties made the MSA give acceptable computational cost results. The MPSO has a higher cpu-time than the other methods, which may be attributed for the application of the two modifications on all particles. It is important to point out that, the MPSO and MSA have the highest convergence speed and therefore the quickest answer to the problem at hand.

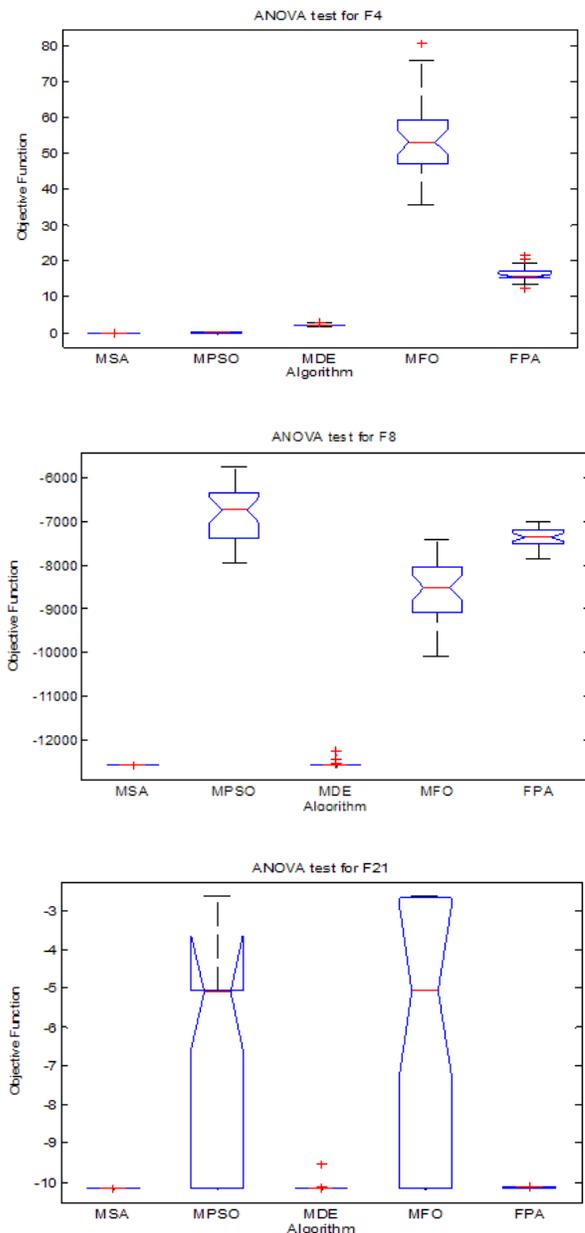


Fig. 4. ANOVA tests for different algorithms.

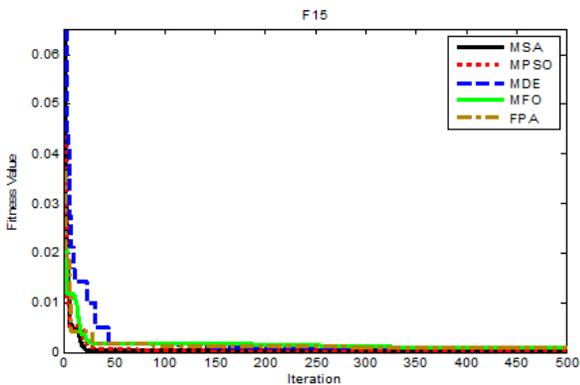
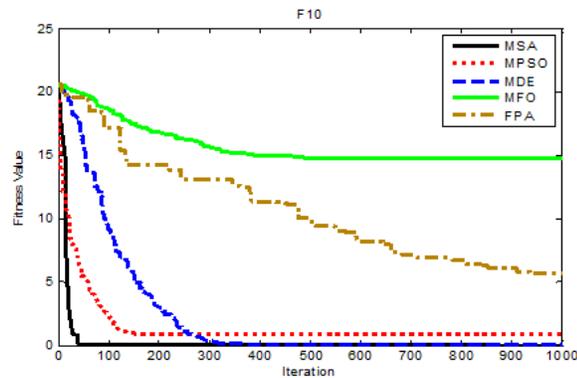
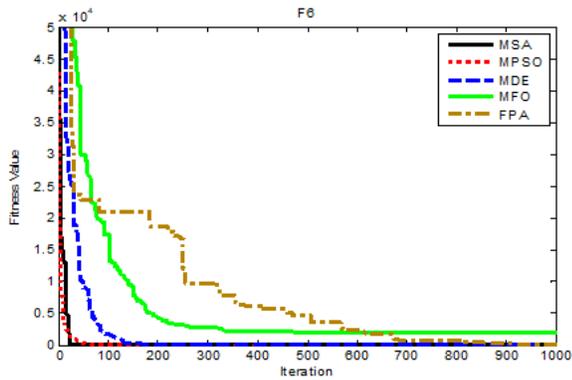
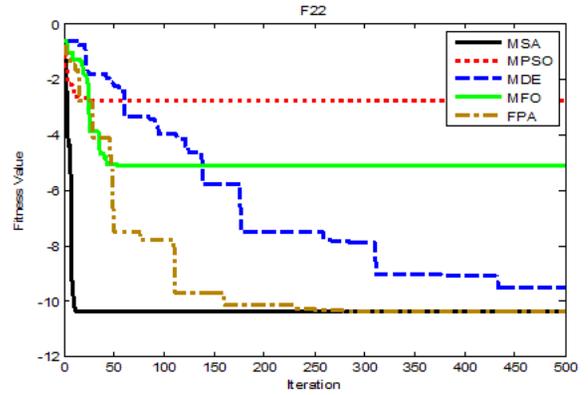
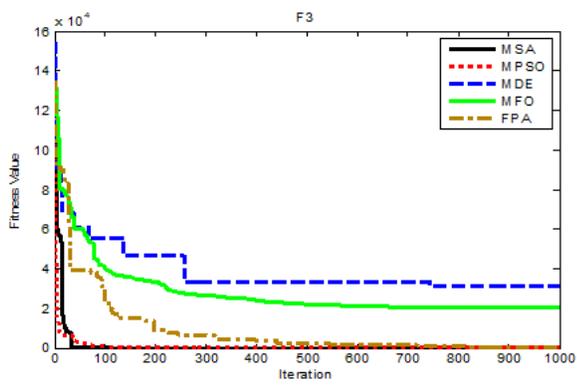


Fig. 5. Comparison of convergence curves of GMSA and literature algorithms obtained in some of the benchmark problems.

V. RESULTS AND DISCUSSION OF RADIAL DISTRIBUTION SYSTEM

To evaluate the efficiency of the proposed MSA method against power loss and energy cost minimization, The IEEE radial distribution systems of 33 and 69-bus have been applied for this simulation. The MATLAB is used to implement the MSA technique for the optimal capacitor placement problem.

This study includes the annual cost of real power loss and the total capacitor banks. The obtained results are compared with other conventional algorithms over 50 independent runs described as follows.

A. IEEE 33-Bus Test System

To evaluate the impact of the proposed MSA on the medium scale of distribution system, the IEEE 33-bus system has been tested. Fig. 6 shows the single line diagram of this system. The system rated voltage is 12.66 kV. The load and line data are given in [13]. Load flow calculation is run before compensation, the minimum bus voltage is registered as 0.9036 p.u at bus 18 and the total active power loss is 210.98 kW with the annual energy losses cost of 35442.96 \$. Using the proposed MSA method, only three capacitors are allocated at optimal locations at buses 12, 24 and 33 with the size of 450, 600 and 900 kVAR, respectively. As a result, the real power loss is diminished to 137.227 kW as 35.02% of the base case. Furthermore, as seen from Table VIII, MSA reduced the total losses by 34.55 kW compared with Analytical IP [23], 14.52 kW compared with SA [20], 6.81 kW compared with GA [9] and 4.01 kW compared with FRCGA [26]. Also, The annual energy cost is diminished from 35442.96\$ to 23446.98\$ with a net saving of 33.85%, which is the best percentage value compared with 30.29% of GA [9], 31.66% of FRCGA [26], 26.94% of SA [20] and 17.17% of analytical IP [23] as seen in Table VIII in Appendix (A).

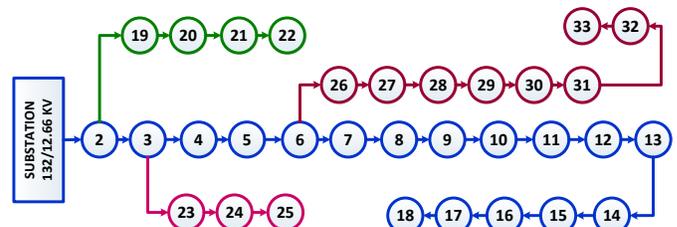


Fig. 6. Single line diagram of IEEE 33-bus RDN [13].

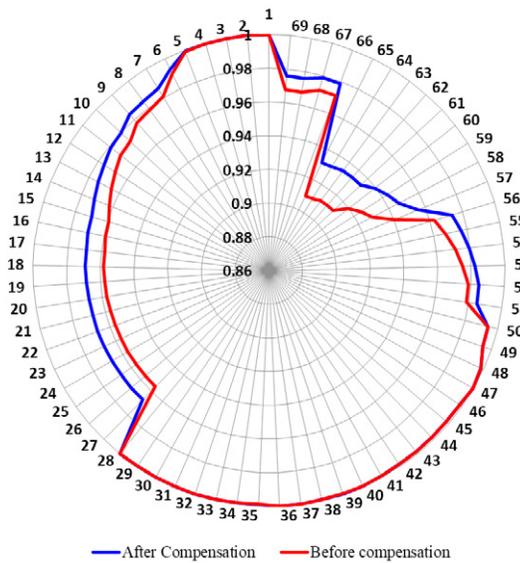


Fig. 7. System Voltage Profile for 33-bus system.

In addition, the system voltage profile is improved and the worst bus voltage is enhanced to 0.9329 PU as shown in Fig. 7. These results have validated the performance and effectiveness of the proposed MSA method. Furthermore, the minimization of active power loss and total cost is stabilized with the fast and smooth convergence as shown in Figs. 8 and 9. It is shown that the proposed MSA is more effective than the conventional algorithms under the medium scale of distribution system.

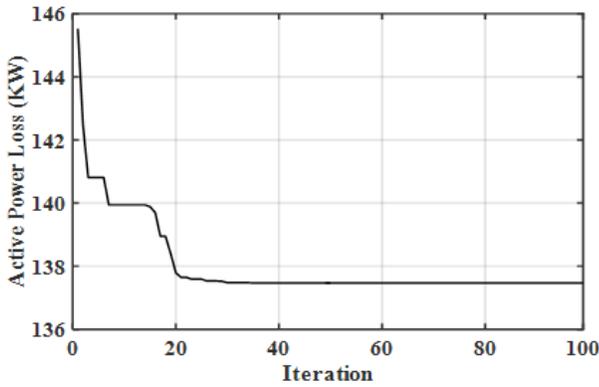


Fig. 8. Power loss convergence of 33-bus system with MSA.

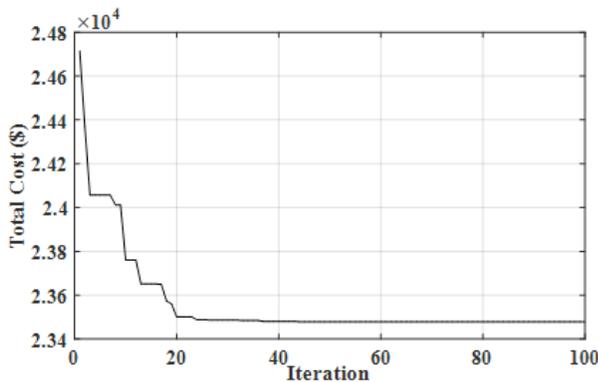


Fig. 9. Total cost convergence of 33-bus system with MSA.

B. IEEE 69-Bus Test System

The proposed MSA is further applied on the IEEE 69-bus system which consists of 69 buses and 68 branches as shown in Fig. 10. The rated line voltage is 12.66 kV and total system load is (1.896MW+j1.347MVAR). The details load and line data are reported in [13]. After running the power flow calculation and before placing the capacitors banks in the RDN, the power loss is obtained at 224.975 kW with the lowest bus voltage at bus 65 is (0.9092 p.u.) and the total energy losses cost is 37800 \$ per year. When applying the proposed method on this RDN, the best active power loss reduction is at 145.404 kW which increased the percentage of loss reduction to 35.37%. This result shows the best value when it is compared with 30.37% of GA [8], 32.22% of PSO [6], 34.66% of DSA [17], 34.95% of TLBO [9], 34.23% of CSA [11], 35.14% of GSA [12] and 35.2% of FPA [5]. This result is considered as the greatest value compared with other algorithms in Table IX Appendix (A). It is found that only three capacitor banks with optimum ratings of 450, 150 and 1200 kVAR have been installed at buses 12, 21 and 61, respectively. Furthermore, the MSA has minimized the total cost per year to 24820.84\$ instead of 37800\$ before compensation. The annual net saving is increased to 34.34% as shown in Table IX in appendix (A). This table displays the statistical performance of the proposed MSA with the best, worst and average values of the total cost for 50 independent runs. Moreover, Fig. 11 confirms the effectivity of the proposed technique by showing the improvement in system voltages. The minimum voltage has been improved to 0.9324 p.u. which is compatible with the voltage constrains. In addition, the fast and effective response of the MSA appears in the convergence curves of total real power loss and total cost in Fig. 12 and 13. The best result obtained from the 50 independent runs for the radial distribution systems were shown in Fig. 8, 9, 12, and 13.

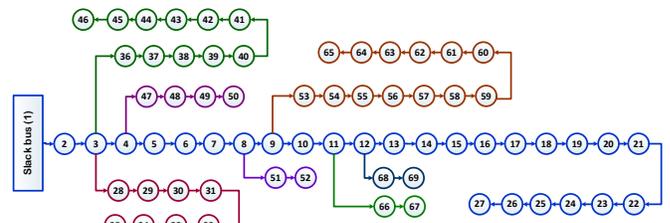


Fig. 10. Single line diagram of IEEE 69-bus RDN [13].

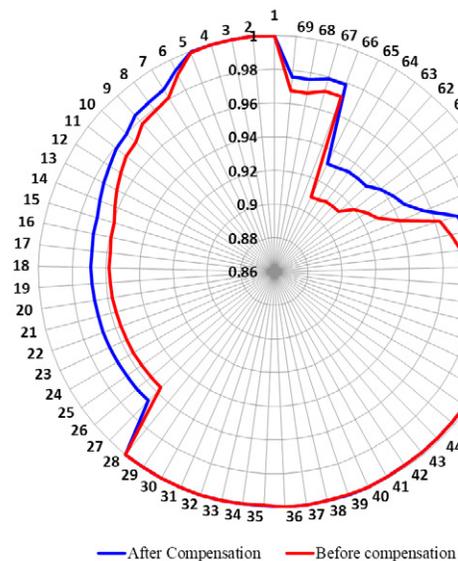


Fig. 11. System Voltage Profile for 69-bus system.

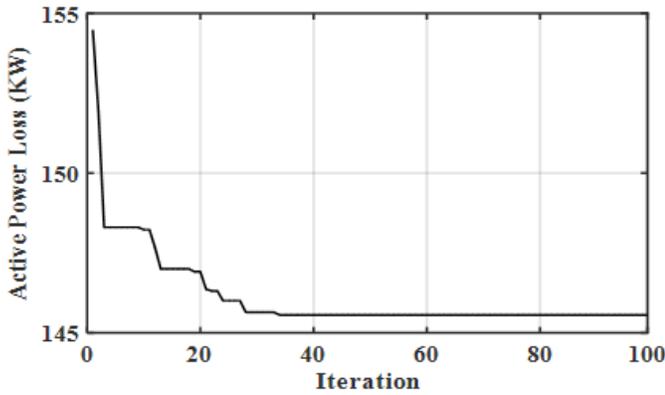


Fig. 12. Power loss convergence rate of 69-bus system with MSA.

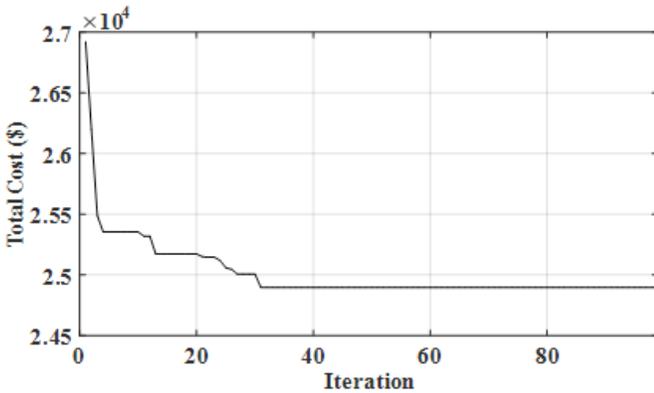


Fig. 13. Total cost convergence rate of 69-bus system.

VI. RESULTS FOR RING DISTRIBUTION SYSTEMS

In this section, the proposed MSA optimization method is tested on the complex power systems known as the ring distribution systems, which are more sensitive to variations and uncertainties. Moreover, the ring main system is considered more complex than the radial system in terms of load flow problems and improper coordination problems. The ring distribution systems are built by modifying the standard IEEE 33 and 69-bus.

A. IEEE 33-Bus System

The radial IEEE 33-bus system is reconfigured to the ring main system as shown in Fig. 14. It consists of sectionalized switches from 1 to 32 and tie-switches from 33 to 37. In case of converting the system from radial to ring by using tie-lines (33 to 37), the power loss reduced to 202.68 kW and minimum bus voltage is 0.913 p.u. With optimal reconfiguration (7-11-14-32-37), it can reduce the active power loss by 41.2% and improve the minimum voltage to 0.938 p.u. On the other hand, applying the MSA method on the ring main system for determining the optimal locations and sizes of capacitor banks needs to minimize the power loss with optimal reconfiguration. The result-based MSA shows a good performance as only three locations have been selected at buses 6, 24 and 33 with total reactive power of 1500 kVAR. Furthermore, the total active power loss is reduced with 69.1% from the base case which is better than the other methods as seen in Table X appendix (A). This table summarizes a detail comparison between the MSA, BGSA [35], HSFLA [36], PSO [36], IPSO [36] and ACO [36] for active power loss, minimum voltage and reduction percentage. Moreover, for all bus voltages-based the MSA method are maintained within desirable values and higher than 0.964 p.u as

shown in Fig. 15. In addition, Fig. 16 shows the effective performance of MSA as total power loss converges smoothly to its minimum values without fluctuations.

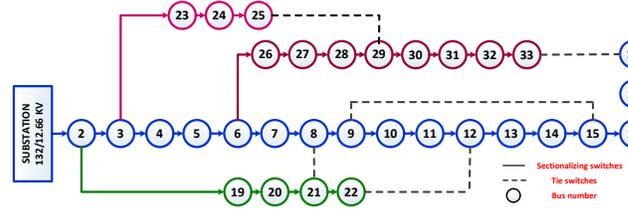


Fig. 14. Single line diagram of IEEE 33-bus Ring system (Tie switches = 21-8; 9-15; 12-22; 18-33; 25-29).

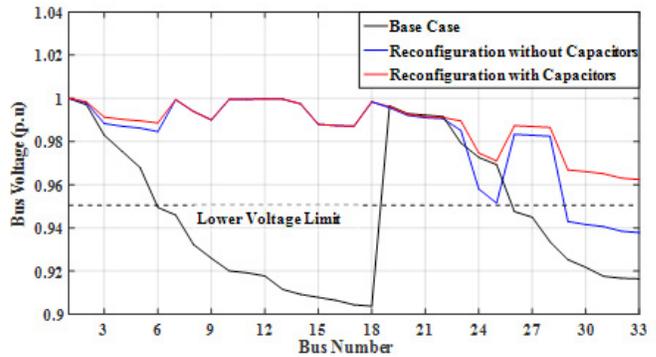


Fig. 15. Bus Voltage Profile of 33-bus Ring system.

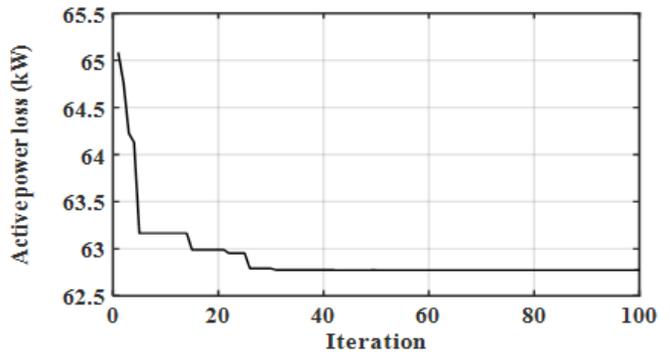


Fig. 16. convergence rate of 33-bus ring system with MSA.

B. IEEE 69-Bus System

The MSA method is further implemented on the large IEEE 69-bus test system after converted to a ring main system as shown in Fig. 17. It consists of sectionalized switches from 1 to 68 (normally closed) and tie-switches from 69 to 73 (normally open). In the base case with tie switches (69-70-71-72-73), the total power loss and the minimum bus voltage are at 224.97 kW and 0.909 p.u, respectively. In case of optimal reconfiguration, it can reduce the active power loss by 59.17% and increase the minimum bus voltage to 0.9877 p.u. as shown in Fig. 18. On the other hand, when installing capacitor banks in the ring main system and implementing the MSA technique for optimizing their locations and sizes gives the great results. Only three locations at buses 11, 50 and 61 have been selected to install capacitors with total reactive power 1500 kVAR. In addition, the total active power loss is diminished by 93.98% from the base case. This result is considered the best comparing with other techniques such as BFO [37], TSA [38], BA [39] and WOA [40] as seen in Table XI in appendix (A). Moreover, the minimum bus voltage increased to 0.99 p.u, which is considered a very

good value as seen in Fig. 18. Furthermore, the effectiveness of the proposed MSA is seen in Fig. 19, which shows the fast convergence of the total active power loss. The best results obtained from the 50 independent runs for the ring distribution systems were shown in Fig. 16 and 19.

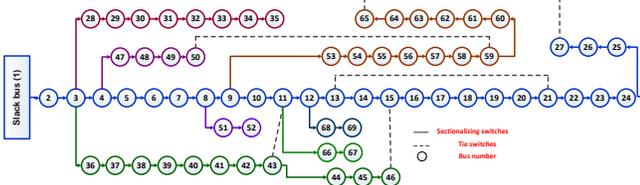


Fig. 17. Single line diagram of IEEE 69-bus Ring system (Tie switches = 11-43; 13-21; 15-46; 50-59; 27-65).

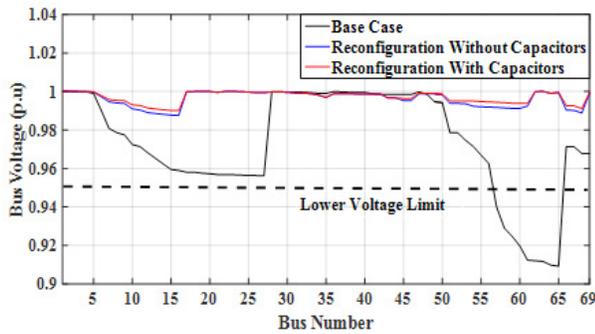


Fig. 18. Bus Voltage Profile of 69-bus Ring system.

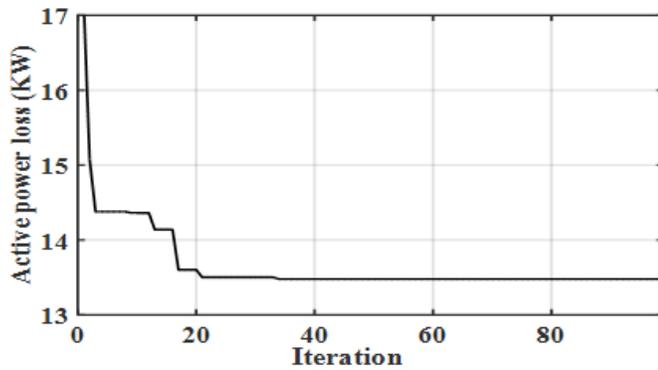


Fig. 19. Convergence rate of 69-bus ring system with MSA.

VII. CONCLUSION

In this article, a novel MSA paradigm has been presented with two new optimization operators of adaptive crossover based on population diversity and associative learning mechanism with immediate memory, which may be appropriate to hybrid with other algorithms in the future. Twenty-three commonly used benchmarks under different statistical metrics are employed to verify the effectiveness of proper hybridization in terms of convergence, local optima avoidance, robustness, computational cost, exploration, and exploitation. From the obtained results, the final algorithm can be considered a hybrid of algorithms of the PSO, DE, MFO, and PFA in line with the natural characteristics of the moth swarm, and suitable for solving the complex problems. The comparative study with several metaheuristic search techniques, confirms the primacy of the proposed paradigm and its potential to find accurate, fast and robustness solutions.

MSA approach has been successfully applied on the small, medium and large scale electrical distribution systems to solve the problem of capacitors allocation for minimizing the real power losses and annual energy cost, which is considered as an attractive economic issue. MSA superiority is clarified by testing it on radial/ring IEEE distribution networks (33 and 69-bus systems). Furthermore, the proposed MSA can improve the voltage profile at each bus in the systems. Moreover, overall numerical results obtained from the proposed MSA method such as minimum voltage, active power loss, power loss cost, capacitor cost, annual energy cost, net saving cost and CPU time have been compared with other algorithms. The MSA method presents a desirable and superior performance with stable convergence against the other techniques. The applications of the proposed MSA method can be considered as the most recent optimization algorithms for the network reconfiguration and dealing with the protection coordination system in presence of capacitors banks and distribution generation during grid faults are the future scope for this work.

APPENDIX (A)

NOMENCLATURE

P_k	Real power flow from bus k
Q_k	Reactive power flow from bus k
P_{Lk}	Real power load connected at bus k
Q_{Lk}	Reactive power load connected at bus k
$P_{L(k+1)}$	Real power load connected at bus k+1
$Q_{L(k+1)}$	Reactive power load connected at bus k+1
R_k	Resistance connected between buses k and k+1
X_k	Reactance connected between buses k and k+1
V_k	Voltage at bus k
V_{k+1}	Voltage at bus k+1
K_e	The annual cost co-efficient per unit of power losses [13]
K_{fc}	Cost co-efficient in \$/KVAR (3) [17]
εI	Random samples drawn from Gaussian stochastic distribution
Q_{fc}	Reactive power compensation
V_{min}	Minimum bus voltage value
V_{max}	Maximum bus voltage value
$P_{T loss}$	Tap setting of transformer
n_p	Number of pathfinders moths
μ^t	Variation coefficient
σ_j^t	Dispersal degree
$best_g$	The global best solution
r_1, r_2	Random number within the interval [0, 1]
P_j	The real power loss during jth load level
n	The number of candidate buses
Q_{fc}	The size of the shunt capacitor
$\varepsilon_2, \varepsilon_3$	Random numbers distributed uniformly within the interval [0,1]

TABLE I. CONTROL-PARAMETERS VALUES FOR THE DIFFERENT ALGORITHMS

MPSO	MDE	MFO	FPA
Personal Coefficient: $c1 = 2$	Strategy: best/1/bin	Shape constant: $b = 1$	Probability switch: $P = 0.8$
Global Coefficient: $c2 = 2$	Mutation factor: $MF = 0.5$		standard gamma function: $\lambda = 1.5$
Inertia Damping Ratio: $Wdamp = 0.99$	Crossover rate: $CR = [0.2, 0.8]$		
Inertia Weight: $w = 1$			

TABLE II. PARAMETER SENSITIVITY ANALYSIS OF MSA AT DIFFERENT BENCHMARK FUNCTIONS

function	$n_p = 5$			$n_p = 6$			$n_p = 7$			$n_p = 8$			$n_p = 9$		
	Mean	StdDev	rank												
F1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1
F2	9.983E-214	0	1	1.6943E-205	0	2	3.5779E-200	0	3	1.3595E-191	0	4	7.3457E-177	0	5
F3	0	0	1	0	0	1	1.8741E-314	0	3	2.2892E-307	0	4	6.3922E-285	0	5
F4	2.4672E-203	0	1	5.1475E-195	0	2	6.9176E-188	0	3	8.3451E-180	0	4	1.4808E-165	0	5
F5	26.6778	0.85684	5	26.3878	0.82463	3	26.3446	0.61077	2	26.5347	0.86174	4	26.0668	0.47009	1
F6	0.00012374	9.0761E-005	5	8.1277E-005	6.6125E-005	4	4.7198E-005	2.8043E-005	2	4.588E-005	3.341E-005	1	6.4774E-005	0.00015573	3
F7	0.00019654	0.0002162	5	0.00019951	0.00015317	4	0.00015027	0.00011724	2	0.00014524	0.00015401	1	0.0001865	0.00016547	3
F8	-12561.5104	43.5228	2	-12561.5321	43.5243	3	-12569.4604	0.080905	1	-12442.2964	696.4124	5	-12549.6018	88.9878	4
F9	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1
F10	8.8818E-016	0	1												
F11	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1
F12	0.0034677	0.018933	3	0.0034628	0.018931	1	0.0069211	0.026312	5	0.0034658	0.018938	2	0.0069172	0.026305	4
F13	0.087631	0.19905	5	0.043883	0.053288	3	0.037208	0.061412	1	0.044498	0.076444	4	0.033096	0.032425	2
F14	2.5035	2.945	5	1.6212	1.8611	1	1.8848	2.0157	2	2.3381	2.9756	3	2.4652	3.3714	4
F15	0.0024799	0.0060749	5	0.00049966	0.00040909	2	0.00042689	0.00029262	1	0.00056821	0.00038599	3	0.001143	0.0036408	4
F16	-1.0316	5.8312E-016	1	-1.0316	5.9036E-016	1	-1.0316	5.7578E-016	1	-1.0316	5.9752E-016	1	-1.0316	5.4546E-016	1
F17	0.39789	1.1366E-015	1	0.39789	1.4366E-015	1	0.39789	2.1748E-015	1	0.39789	9.7879E-016	1	0.39789	6.4868E-016	1
F18	4.8	6.8501	5	3.9	4.9295	3	3.9	4.9295	3	3	1.5339E-015	1	3	1.8011E-015	1
F19	-3.8628	2.3456E-015	1	-3.8628	2.2853E-015	1	-3.8628	2.6039E-015	1	-3.8628	2.3164E-015	1	-3.8628	2.2868E-015	1
F20	-3.2625	0.060463	4	-3.2943	0.051146	2	-3.3022	0.045066	1	-3.2507	0.059241	5	-3.2824	0.057005	3
F21	-9.9024	1.3735	3	-10.1532	9.4183E-006	1	-10.1532	4.3637E-015	1	-9.7433	1.562	5	-9.9024	1.3735	3
F22	-9.5125	2.3091	2	-9.0358	2.0357	4	-9.5125	2.3091	2	-9.7351	2.0378	1	-9.0033	2.8545	5
F23	-9.28	2.6391	4	-10.5364	8.1263E-006	1	-9.1527	2.7367	5	-9.4196	2.54	3	-10.313	1.2234	2
Average rank			2.74			1.91						1.91			2.48
Overall rank			5			1						1			3
															2.65
															4

TABLE III. RESULTS OF UNIMODAL BENCHMARK FUNCTIONS FOR DIFFERENT ALGORITHMS

	MSA			MPSO			MDE			MFO			FPA		
	Mean	StDev	rank												
F1	0	0	1	4.0179E-042	1.5634E-041	2	5.1618E-012	2.9283E-012	3	1333.3334	3.4575E+003	5	489.8355	160.965	4
F2	3.5779E-200	0	1	0.01462316	0.0745866	3	4.9543E-008	1.1054E-008	2	41.6668	26.5334	5	25.4815	4.9106	4
F3	1.8741E-314	0	1	0.12362652	0.1646994	2	25213.1977	3871.0339	5	1.4358E+4	9.9687E+003	4	485.7154	177.7229	3
F4	6.9176E-188	0	1	0.09716659	0.06019727	2	2.0372	0.34802158	3	54.4144	11.9746	5	16.2469	1.9700	4
F5	26.3446	0.61077	1	38.7035	29.24275	2	40.931	19.0759	3	2.6773E+006	1.4593E+007	5	3.2311E+004	1.5337E+004	4
F6	4.7198E-005	2.8043E-005	3	3.142E-019	1.6071E-018	1	5.3249E-012	2.6727E-012	2	666.6834	2.5373E+003	5	518.8479	138.4542	4
F7	0.00015027	0.00011724	1	0.00934203	0.0038743	2	0.0231033	0.00569021	3	3.36868	7.9414	5	0.1329	0.0504	4
Average rank			1.29			2			3			4.86			3.86
Overall rank			1			2			3			5			4

TABLE IV. RESULTS OF MULTIMODAL FUNCTIONS WITH MANY LOCAL MINIMA FOR DIFFERENT ALGORITHMS

	MSA			MPSO			MDE			MFO			FPA		
	Mean	StDev	rank	Mean	StDev	rank	Mean	StDev	rank	Mean	StDev	rank	Mean	StDev	rank
F8	-12569.4604	0.080905	1	-6768.5566	652.5301	5	-12544.8116	67.8879381	2	-8608.724	697.2892	3	-7395.7054	243.0420	4
F9	0	0	1	44.507762	11.6937	2	60.117	4.5993	3	155.092	34.3843	5	135.4706	16.0127	4
F10	8.8818E-016	0	1	0.64426853	0.75617291	3	6.0902E-007	1.4946E-007	2	12.0325	9.2836	5	5.3276	0.7895	4
F11	0	0	1	0.018695461	0.047044852	3	1.0117E-010	1.6559E-010	2	30.1244	49.3831	5	5.8495	1.5169	4
F12	0.0069211	0.026312	2	0.048376744	0.070639813	3	6.5335E-013	3.1533E-013	1	0.350619	0.5190	4	9.5593	2.5942	5
F13	0.037208	0.061412	2	0.046969551	0.23055956	3	3.33E-012	1.5799E-012	1	1.3669E+007	7.4867E+007	5	172.9146	205.6159	4
Average rank			1.33			3.17			1.83			4.5			4.17
Overall rank			1			3			2			5			4

TABLE V. RESULTS OF MULTIMODAL FUNCTIONS WITH ONLY A FEW LOCAL MINIMA FOR DIFFERENT ALGORITHMS

	MSA			MPSO			MDE			MFO			FPA		
	Mean	StDev	rank												
F14	1.8848	2.0157	4	2.6112459	2.0596228	5	0.998	0	1	1.32869	0.7515	1	0.998	3.5971E-005	1
F15	0.00042689	0.00029262	2	0.00035039	0.00023489	1	0.0006994	8.195E-005	4	0.00114865	0.0014	5	0.00060941	0.00012646	3
F16	-1.0316285	5.7578E-016	1	-1.0316285	6.7752E-016	1	-1.0316285	6.7752E-016	1	-1.0316285	6.7752E-016	1	-1.0316285	7.3634E-008	1
F17	0.39789	2.1748E-015	1	0.39789	0	1	0.39789	0	1	0.39789	0	1	0.3979	1.6453E-006	1
F18	3.9	4.9295	5	3	1.4981E-015	1	3	1.296E-015	1	3	1.6097E-016	1	3	2.7239E-009	1
F19	-3.8628	2.6039E-015	1	-3.8628	2.7101E-015	1	-3.8628	2.71E-015	1	-3.86278	2.7101E-015	1	-3.8628	2.2729E-008	1
F20	-3.3022	0.045066	2	-3.2863272	0.05541509	4	-3.32198	5.442E-005	1	-3.20983	0.0502	5	-3.2975	0.0147	3
F21	-10.1532	4.3637E-015	1	-6.7161817	3.1908	4	-10.1328	0.10940485	3	-5.72036	3.1242	5	-10.1432	0.0103	2
F22	-9.5125	2.3091	3	-6.3183386	3.4995	5	-10.4025	0.0011901	2	-8.65134	2.9966	4	-10.4374	0.0551	1
F23	-9.1527	2.7367	4	-7.0198746	3.8745	5	-10.5363	0.00059529	1	-9.41713	2.5711	3	-10.4172	0.0900	2
Average rank			2.4			2.8			1.6			2.7			1.6
Overall rank			1			3			2			5			4

TABLE VI. P-VALUES CALCULATED FOR WILCOXON'S RANK-SUM TEST ($p \geq 0.05$ HAVE BEEN UNDERLINED)

	MPSO	MDE	MFO	FPA
F1	<u>0.0742443</u>	0.0418114	0.01228	6.263E-010
F2	0.0044408	0.041814	2.0318E-008	5.7225E-009
F3	4.1048E-005	6.263E-010	6.263E-010	6.263E-010
F4	5.5847E-005	2.9812E-007	6.263E-010	1.6501E-008
F5	5.7467E-009	7.0497E-010	7.0497E-010	6.2928E-010
F6	<u>0.0742443</u>	0.041814	0.021081	6.263E-010
F7	0.00048083	0.00048083	4.7908E-005	0.00048083
F8	1.6501E-008	1.6501E-008	1.6501E-008	1.6501E-008
F9	7.0165E-010	6.263E-010	6.263E-010	6.263E-010
F10	0.0022103	0.041814	3.6027E-007	1.6501E-008
F11	0.0081071	0.041814	3.0022E-005	1.6501E-008
F12	0.031296	0.041819	0.0003429	1.6567E-008
F13	0.03577	0.041814	0.00044952	7.0165E-010
F14	4.4603E-007	4.448E-007	8.304E-007	3.2786E-007
F15	0.0026421	0.00048083	0.00048045	0.00048083
F16	0.0011031	0.0011031	0.0011031	0.015781
F17	8.4451E-006	8.4451E-006	8.4451E-006	4.9194E-006
F18	2.7033E-008	6.9137E-009	2.4213E-008	1.6501E-008
F19	8.4451E-006	8.4451E-006	8.4451E-006	0.00048083
F20	0.00056673	0.0013295	0.0018718	0.0031498
F21	8.3938E-006	2.7552E-006	3.1691E-005	4.9194E-006
F22	6.4387E-005	2.885E-006	3.4065E-006	4.9194E-006
F23	5.8609E-006	1.2815E-007	3.613E-008	3.2786E-007

TABLE VII. COMPARISONS OF AVERAGE CPU TIME (s) FOR SOME OF BENCHMARK FUNCTIONS

	MSA	MPSO	MDE	MFO	FPA	MGSO[34] ^a	PSO[33] ^b	GSO[33] ^b
F1	4.8359	10.6167	8.1457	3.2555	4.6371	8.98	36.3	27.6
F5	4.9062	10.6391	8.8892	3.8525	5.3113	18.4	37.6	27.8
F7	5.8201	10.4509	10.4318	4.3053	5.4247	19.7	41.4	31.4
F8	4.9735	9.9465	8.3939	3.3662	4.8413	20.5	43.0	32.7
F9	5.0244	10.0350	8.3013	3.4761	4.4404	10.5	71.4	50.1
F10	5.5109	11.0692	9.0345	4.0075	4.9402	12.4	44.1	34.4
F11	5.2731	10.5239	9.0721	3.7285	4.7920	34.8	45.4	35.9

a The experiments were carried out on a PC with a 2.60-GHz Intel Processor and 2.0-GB RAM.

b The experiments were carried out on a PC with a 1.80-GHz Intel Processor and 1.0-GB RAM.

TABLE VIII. COMPARISON RESULTS USING DIFFERENT OPTIMIZATION TECHNIQUES FOR (33-BUS SYSTEM)

Methods	V_{\min} (P.U)	Optimal location (Bus no.) Optimal size (kVAR)	P_{loss} (kW)	Cost of P_{loss} (\$)	capacitor cost (\$)	Worst Total cost (\$)	Mean Total cost (\$)	Best Total cost (\$)	Total cost reduction	CPU Time (s)
Base Case	0.903	-	210.98	35442.9					-	-
MSA	0.932	12(450),24(600),30(900)	137.23	23054.1	392.85	23627	23534.6	23446.98	33.85%	5.4
GA [9]	NA	7(850), 29(25), 30(900)	144.04	24198.7	507.15	NA	NA	24705.87	30.29%	NA
FRCGA [21]	NA	28(25), 6(475), 29(300), 8(175), 30(400), 9(350)	141.24	23728.3	492.86	NA	NA	24221.18	31.66%	NA
Analytical IP [23]	0.950	9(450), 29(800), 30(900)	171.78	28859.0	499.35	NA	NA	29358.39	17.17%	NA
SA [20]	0.959	10(450), 30(350), 14(900)	151.75	25494	401.052	NA	NA	25895.05	26.94%	NA

TABLE IX .COMPARISON RESULTS USING DIFFERENT OPTIMIZATION TECHNIQUES FOR (69-BUS SYSTEM)

Methods	V_{min} (p.u)	Optimal location (bus no.) Optimal size (kVAR)	Ploss (kW)	Cost of Ploss (\$)	capacitor cost (\$)	Worst Total cost (\$)	Mean Total cost (\$)	Best Total cost (\$)	% Total cost reduction	CPU Time (s)
Base Case	0.909	-	224.98	37800					-	-
MSA	0.932	21(150), 12 (450),61(1200)	145.41	24427.99	392.85	24987.9	24908.2	24820.8	34.3	11.42
Fuzzy-GA [17]	0.936	59(100), 61(700), 64(800)	156.62	26312.16	424.8	NA	NA	26736.9	29.2	NA
PSO [6]	NA	46(241), 47(365), 50(1015)	152.48	25616.64	433.99	NA	NA	26050.6	31.1	NA
DE [18]	0.931	57(150), 58 (50), 61(1000), 60(150), 59(100)	151.38	25431.84	316.1	NA	NA	25747.9	31.8	NA
TLBO [10]	0.932	22(300), 61(1050), 62(300)	146.35	24586.8	446.4	NA	NA	25033.2	33.7	15.76
CSA [12]	NA	21(250), 62(1200)	147.95	24855.6	291.5	NA	NA	25147.1	33.4	NA
GSA [13]	0.951	26(150), 13(150), 15(1050)	145.9	24511.2	451.5	NA	NA	24962.7	33.9	NA

TABLE X. COMPARISON RESULTS USING DIFFERENT OPTIMIZATION TECHNIQUES FOR 33-BUS RING SYSTEM

Techniques	V_{min} (p.u)	Ploss (kW)	% Loss reduction	Optimal location (bus no.) Optimal size (kVAR)	Tie switches
Base Case	0.913	202.68	-	-	33-34-35-36-37
MSA	0.9642	62.68	69.1	6(150), 24(450),30(900)	7,11,14,32,37
PSO [36]	0.9635	95.38	52.93	9(300), 10(300), 31(300), 6(600), 29(600)	7-10-14-36-37
IPSO [36]	0.9656	98.83	51.23	5(300), 13(300), 32 (300), 28(1200)	11-28-33-34-36
IBPSO [36]	0.9585	93.06	54.08	11(300), 24(300), 32(300), 6(600), 29(600),	7-9-14-32-37
ACO [36]	0.9656	95.79	52.73	20(600), 28(450), 29(600)	7-9-14-32-37
BGSA [35]	0.956	113.56	43.97	5(NA), 27(NA), 28(NA)	10-14-28-32-37
HSFLA [36]	0.9585	92.58	54.32	300(2-4-10-11-18-24-28-29-30)	7-11-14-32-37

TABLE XI. COMPARISON RESULTS USING DIFFERENT OPTIMIZATION TECHNIQUES FOR 69-BUS RING SYSTEM

Techniques	V_{min} (p.u)	P_{loss} (kW)	% Loss reduction	Optimal location (bus no.) Optimal size (kVAR)	Tie switches (bus no.)
Base Case	0.9092	224.97	-	-	69-70-71-72-73
MSA	0.990	13.534	93.98	11(450), 50(150),61(900)	14-45-52-69-70
BFO [37]	0.9534	33.63	85.05	11(300), 49(450), 61(900)	4-56-61-69-70
TSA [38]	0.956	108.94	51.58	24(100), 45(200), 49(300), 61(400)	14-20-52-61
BA [39]	0.9561	88.413	60.7	27(1350), 37(2250), 62(1200)	11-58-69-70-73
WOA [40]	0.99	66.74	70.3	50(350), 61(1050), 64(390)	13-18-56-61-69

REFERENCES

- [1] A. Y. Abdelaziz, Ehab S. Ali & Sahar M. Abdelazim, "Flower Pollination Algorithm for Optimal Capacitor Placement and Sizing in Distribution Systems" *Electrical Power Components and Systems*. Vol. 44, no. 5, pp. 544-555, 2016.
- [2] J.B.V. Subrahmanyam, "Optimal capacitor placement in unbalanced radial distribution networks", *J. Theor. Appl. Inform. Technol.* Vol. 6, no. 1, pp. 106-115, 2009.
- [3] T.S. Chung, Ge Shaoyun, "A recursive LP-based approach for optimal capacitor allocation with cost-benefit consideration", *Electric Power Syst. Res.*, Vol. 39, no. 2, pp. 129-136, Nov. 1996.
- [4] Ahmed Elsheikh, Yahya Helmy and Yasmine Abouelseoud, "Optimal Capacitor placement and Sizing in Radial Electric Power System", *Alex. Engineering Journal*. Vol. 53, pp. 809-816, 2014.
- [5] Almoataz Y. Abdelaziz, Ehab S. Ali, Sahar M. Abdelazim, "flower pollination algorithm for optimal capacitor placement and sizing in distribution systems", *Electrical Power Components Systems*. Vol. 44, no. 5, pp. 544-555, 2016.
- [6] K. Prakash, M. Sydulu, "Particle swarm optimization based capacitor placement on radial distribution system", *Proceedings of IEEE Power Engineering Society General Meeting*. pp. 1-5, 2007.
- [7] T. Kerdpol et. al, "Optimization of a battery energy storage system using particle swarm optimization for stand-alone microgrids" *Electrical Power and Energy Systems*. Vol. 81, pp. 32-39, 2016.
- [8] A. Elsheikh, Y. Helmy, Y. Abouelseoud, A. Elsherif, "Optimal capacitor placement and sizing in radial electric power systems", *Alex. Eng J*, Vol. 53, no. 4, pp. 809-816, 2014.
- [9] Sydulu M, Reddy VVK, "Index and GA based optimal location and sizing of distribution system capacitors", *IEEE power engineering society general meeting*. pp. 1-4, June. 2007.
- [10] Sneha Sultana, Provas Kumar Roy, "Optimal capacitor placement in radial distribution systems using teaching learning based optimization", *Elsevier Int. J. Electr. Power Energy Syst.*, Vol. 54, pp. 387-398, 2014.
- [11] R.V. Rao, V.K. Patel, "Optimization of mechanical draft counter flow wet-cooling tower using artificial bee colony algorithm", *Energy Convers Manage*, Vol. 52, no. 7, pp. 2611-2622, 2011.
- [12] Attia A. El-Fergany, Almoataz Y. Abdelaziz, "Cuckoo Search-based Algorithm for Optimal Shunt Capacitors Allocations in Distribution Networks", *Electrical Power Components and Systems*. Vol. 41, no. 16, PP. 1567-1581, 2016.
- [13] Y. Mohamed Shuaiba, M. S. Kalavathib, C. C. Asir, "Optimal capacitor placement in radial distribution system using Gravitational Search Algorithm", *International Journal of Electrical Power & Energy Systems*, Vol. 64, PP.384-397, 2015.
- [14] Felipe Duque, Leonardo Oliveira, Edimar Oliveira et al., "Allocation of capacitor banks in distribution systems through a modified monkey search optimization technique", *International Journal of Electrical Power & Energy Systems*, Vol. 73, PP. 420-432, 2015.
- [15] D. B. Prakash, C. Lakshminarayana "Optimal siting of capacitors in radial distribution network using whale optimization algorithm" *Alexandria Eng J.*, Vol. 56, no. 4, pp. 499-509, 2016.
- [16] E.S. Alia, S.M. Abd Elazima, A.Y. Abdelazizb, "Improved Harmony Algorithm for optimal locations and sizing of capacitors in radial distribution systems", *International Journal of Electrical Power & Energy Systems*. Vol. 79, pp. 275-284, 2016.
- [17] Ying-Tung Hsiao, Chia-Hong Chen, Cheng-Chih Chien, "Optimal capacitor placement in distribution systems using a combination fuzzy-GA method", *International Journal of Electrical Power & Energy Systems*. Vol. 26, no. 7, PP. 501-508, 2004.
- [18] M. Ramalinga Rajua, K.V.S. Ramachandra Murthyb, K. Ravindraa, "Direct search algorithm for capacitive compensation in radial distribution systems", *International Journal of Electrical Power & Energy Systems*. Vol. 42, no. 1, PP. 24-30, 2012.
- [19] S. Neelim "Optimal capacitor placement in distribution networks for loss reduction using differential evolution incorporating dimension reducing load flow for different load levels" *Energetic*, 2012.
- [20] H.D. Chiang, J. Wang, O. Cockings, "Optimal capacitor placements in distribution systems: part 1: a new formulation and the overall problem", *IEEE Trans. Power Deliv*, Vol. 5, pp. 634-642, 1990.
- [21] R. Srinivasas Raa, S.V.L. Narasimhamb, "Optimal capacitor placement in a radial distribution system using Plant Growth Simulation Algorithm", *International Journal of Electrical Power & Energy Systems*, Vol. 33, no. 5, pp. 1133-1139, 2011.
- [22] Ching-Tzong Su, Guor-Rurng Lii, Chih-Cheng Tsai, "Optimal capacitor allocation using fuzzy reasoning and genetic algorithms for distribution systems", *Mathematical and Computer Modelling*, Vol. 33, no. 6-7, pp. 745-757, 2001.
- [23] Ahmed R. Abul'Wafa, "Optimal capacitor allocation in radial distribution systems for loss reduction: a two stage method", *Elsevier Electric Power Syst. Res.*, Vol. 95, pp. 168-174, 2013.
- [24] Mostafa Sedighzadeh, Marzieh Dakhem, Mohammad Sarvi, Hadi Hosseini, "Optimal reconfiguration and capacitor placement for power loss reduction of distribution system using improved binary particle swarm optimization", *Int J Energy Environ*, Vol.73, no.5, pp. 1-11, 2014.
- [25] Sayyad Nojavan, Mehdi Jalali, Kazem Zare, "Optimal allocation of capacitors in radial/mesh distribution systems using mixed integer nonlinear programming approach", *Elsevier Electric Power Syst. Res*. Vol. 107, pp.119-124, 2014.
- [26] Ahmed R. Abul'Wafa, "Optimal capacitor placement for enhancing voltage stability in distribution systems using analytical algorithm and Fuzzy-Real Coded GA", *International Journal of Electrical Power & Energy Systems*, Vol. 55, pp. 246-252, 2014.
- [27] Al-Attar Ali Mohameda, Yahia S. Mohamedb, Ahmed A.M. El-Gaafaryb, Ashraf M. Hemeida, "Optimal power flow using moth swarm algorithm", *Electric Power Systems Research*, Vol. 142, pp. 190-206, 2017.
- [28] J. A. Michline Ruba and S. Ganesh "Power Flow Analysis for Radial Distribution Systems using Backward/Forward Sweep Method", Vol. 8, no.10, pp. 1621 - 1625, 2014.
- [29] Yao X, Liu Y, Lin G. "Evolutionary programming made faster". *IEEE Trans. Evol. Comput*. Vol. 3, no. 2, pp.82-102, 1999.
- [30] Karami A. "A fuzzy anomaly detection system based on hybrid psokmeans algorithm in content-centric networks", *Neurocomputing*; Vol. 149, pp.1253-1269, 2015.
- [31] L. D. Coelho, V. C. Mariani, J. V. Leite. "Solution of Jiles-Atherton vector hysteresis parameters estimation by modified Differential Evolution approaches", *Expert Syst. Appl.*, Vol. 39, no. 2, pp.2021-2025, 2012.
- [32] Mirjalili S., "Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm", *Knowledge Based Syst.*; Vol. 89, pp.228-249, 2015]
- [33] He S., Wu Q.H., "an optimization algorithm inspired by animal searching behavior", *IEEE Trans. Evol. Comput*. Vol. 13, no. 5, pp. 973-990, 2009.
- [34] Zare K., Haque M. T., Davoodi E. "Solving non-convex economic dispatch problem with valve point effects using modified group search optimizer method", *Electr. Power Syst. Res*. Vol. 84, no. 1, pp.83-89, 2012.
- [35] H. Reza and F. Rohollah. "Distribution system efficiency improvement using network reconfiguration and capacitor allocation" *Elect.Power and Energy Sys*. Vol. 64, pp. 457-468, 2015.
- [36] M. Sedighzadeh and M.M. mahmoodi, "Optimal Reconfiguration and Capacitor Allocation in Radial Distribution Systems Using the Hybrid Shuffled Frog Leaping Algorithm in the Fuzzy Framework" *Journal of Operation and Automation in Power Engineering*, Vol. 3, no. 1, pp. 56-70, 2015.
- [37] M. Mohammadi, A. Mohammadi Rozbahani, S. Bahmanyar. "Power loss reduction of distribution systems using BFO based optimal reconfiguration along with DG and shunt capacitor placement simultaneously in fuzzy framework", *Journal of Central South University*, Vol. 24, no. 1, pp. 90-103, 2017.
- [38] T. Lantharthong, N. Rugthaicharoencheep. "Network Reconfiguration for Load Balancing in Distribution System with Distributed Generation and Capacitor Placement", *Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*. Vol. 6, no. 4, 2012.
- [39] Mohammad-Reza Askari, "A New Optimization Framework To Solve The Optimal Feeder Reconfiguration And Capacitor Placement Problems" *Journal of Scientific & Technology Research*. Vol. 4, no. 7, July 2015.
- [40] Sarfaraz Nawaz, A. Bansal, M.P. Sharma, "An Analytical Approach for DG Placement in Reconfigured Distribution Networks", *Journal of Applied Power Engineering*, Vol. 5, no. 3, pp. 137-143, 2016.



Emad A. Mohamed

Emad A. Mohamed received the B.Sc. and M.Sc. degrees in electrical power engineering from Aswan University, Aswan, Egypt in 2005 and 2013 respectively. He is working as a Demonstrator in the Department of Electrical Engineering, Aswan Faculty of Engineering, Aswan University, Aswan, Egypt. Between November 2007 and August 2013, and he is an assistant lecturer from 2013 to 2015. Now, he is a Ph.D. student in Kitakyushu institute of Technology, Japan. He was in a Master Mobility scholarship at Faculté des Sciences et Technologies - Université de Lorraine, France – 1. The scholarship sponsored by FFEEBB ERASMUS MUNDUS. He was a research student from April to October 2015 in Kyushu University, Japan. His research interests are applications of superconducting fault current limiters (SFCLs) on power systems, power system stability, and protection.



Al-Attar Ali Mohamed

Al-Attar Ali Mohamed received his B.Eng., M.Eng, and Ph.D. from Aswan University, Aswan, Egypt in 2010, 2013 and 2017. Currently, he is an assistant professor sponsored in Faculty of Engineering, Aswan University. His research interests are in the area of analysis and design of power systems, optimization, and controls.



Yasunori Mitani

Yasunori Mitani received B.Sc., M.Sc., and D.Eng. degrees in Electrical Engineering from Osaka University, Japan in 1981, 1983 and 1986, respectively. He was a visiting research associate at the University of California, Berkeley, from 1994 to 1995. He is currently a professor at the department of electrical and electronics engineering, Kyushu Institute of Technology (KIT), Japan. At present, he is the head of environmental management center of KIT and vice dean of a graduate school of engineering, KIT. He is also the President of the Institute of Electrical Engineers of Japan (IEEJ), Power and Energy Society.

