Lecture Notes in Electrical Engineering 480

Shahram Montaser Kouhsari *Editor*

Fundamental Research in Electrical Engineering

The Selected Papers of The First International Conference on Fundamental Research in Electrical Engineering

Springer

Lecture Notes in Electrical Engineering

Volume 480

Shahram Montaser Kouhsari
Editor

Fundamental Research in Electrical Engineering

The Selected Papers of The First International Conference on Fundamental Research in Electrical Engineering

Springer

Editor
Shahram Montaser Kouhsari
Department of Electrical Engineering
Amirkabir University of Technology
Tehran
Iran

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-10-8671-7 ISBN 978-981-10-8672-4 (eBook)
https://doi.org/10.1007/978-981-10-8672-4

Library of Congress Control Number: 2018941969

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

The present volume collects the selected papers of the First International Conference on Electrical Engineering (Tehran, Iran, 2017). The proceedings are aimed at addressing problems and topics of concern in all the subbranches of Electrical Engineering by bringing the recent advancements in the field to the attention of the experts; such a general conference in the field can also make the possibility of developing multidisciplinary collaborations and approaches. It is a suitable platform to share the recent findings without making any restriction on the topics. Hope that this proceeding can benefit graduate students, and also researchers in the field.

The first part of the present proceedings volume collects the selected papers on Biomedical Engineering. Topics like contrast enhancement of ultrasound images, mammography, wireless sensor networks, speech recognition, and disease diagnosis have been covered in the first part. The second part is on Control Engineering that presents topics like vibration control, circuit design for controlling automatic gain, nonlinear predictive control, and manipulators controlling in robots. The third part of this volume has been devoted to Electronics Engineering—this section covers optofluidic materials, time series prediction, robot speech control, ionization vacuum gauges with COMSOL, acetone sensing, LUT design, etc. The fourth part is about Power Engineering, and includes the papers that cover topics like photovoltaic solar cells, pumped-storage power stations, optimal capacitors in distribution networks, wind turbines, phase balancing in distribution networks, microelectromechanical switches in smart grids, axial-flux permanent-magnet machines, voltage stability enhancement, etc. Then the present volume ends with the selected papers on Telecommunication that covers topics like cloud environment, node clustering in wireless systems, electrostatics MEMS switches, microstrip antenna, distribution network reconfiguration, machine learning algorithms, security of Internet of Things, data reduction, q-learning, networks' deadlock detection methods, etc.

Tehran, Iran Shahram Montaser Kouhsari

Contents

Part II Control Engineering

Part III Electronic Engineering

Part IV Power Engineering

Part V Telecommunication Engineering

Part I
Biomedical Engineering

Bioelectrical Signals: A Novel Approach Towards Human Authentication

Hamed Aghili

Abstract Human authentication based on electrical bio-signals, or bioelectrical signals, is a rapidly growing research area due to increasing demand for establishing the identity of a person, with high confidence, in a number of applications in our vastly interconnected society. Studies show that bioelectrical signals can be not only employed for diagnostic purposes in medicine, but also used in human authentication since they have unique features among individuals. This article reviews examples of up-to-date researches that have applied bioelectrical signals like Electrocardiogram (ECG), Electroencephalogram (EEG) and Electrooculogram (EOG) in human authentication. Utilizing bioelectrical signals provides a novel approach to user authentication that contains all the crucial attributes of previous traditional authentication. The most significant reasons for deployment of electrical bio-signals in user authentication include their measurability, uniqueness, universality and resistance to spoofing, while other conventional biometrics like face shape, hand shape, fingerprint and voice can be artificially generated.

Keywords Human authentication · Biometrics · Bioelectrical signals
Electroencephalogram signal · Electrocardiogram signal · Electrooculogram signal

1 Introduction

Authentication is carried out in a wide range of areas of different levels of security and importance. Not having a comprehensive understanding of the requirements for authentication according to different circumstances, we use the same traditional authentication, either through an object for example an ID card or via knowledge like passwords, for every situation. This is while new authentication methods have advanced even beyond using conventional biometrics, and are applying

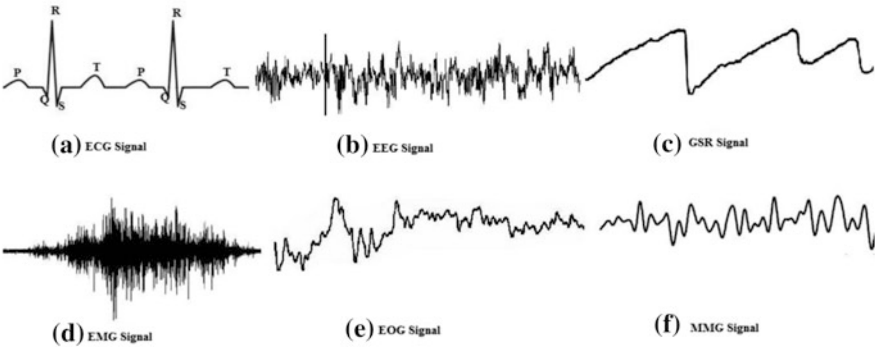
H. Aghili (✉)
Department of Electrical Engineering (Robotic Engineering),
Payame Noor University (PNU), Tehran, Iran
e-mail: engineer.aghili@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_1

bio-electrical signals for authentication purposes. The recent studies have shown that bio-signals can provide human authentication with the resistance to fraudulent attacks since they have specific features that are unique among individuals. In this article we introduce bioelectrical signals and mention their advantage over other conventional biometrics. After that we review some researches that have been carried out in the field of applying Electrocardiogram, Electroencephalogram and Electrooculogram signals for human authentication.

2 What Are Bioelectrical Signals?

Bio-signals are records of a biological event such as a beating heart or a contracting muscle. The electrical, chemical, and mechanical activity that occurs during these biological events often produces signals that can be measured and analyzed [1]. Bio-signals are divided into six groups according to their physiological origin: bioelectrical signals, bio-magnetic signals, bio-chemical signals, bio-mechanical signals, bio-aquatic signals and bio-optical signals. The bio-signal of our interest in this article is bioelectrical signals. Bioelectrical signals are those that are generated by the summation of electrical potential differences across an organ [2]. Via surface electrodes attached or close to the body surface, signals from a broad range of sources can be recorded [3] precisely, if a nerve or muscle cell is stimulated, it will generate an action potential that can be transmitted from one cell to adjacent cells via its axon. When many cells become activated, an electric field is generated. These changes in potential can be measured on the surface of the tissue or organism by using surface electrodes [1]. Bioelectrical signals are very low amplitude and low frequency electrical signals [4]. These signals are generally used for medical diagnosis, but research findings confirm that since they have unique features among individuals, they can also be used for human authentication. The examples of bioelectrical signals are Electrocardiogram, Electroencephalogram, Galvanic skin response and Electrooculogram "Fig. 1".



(a) ECG Signal (b) EEG Signal (c) GSR Signal

(d) EMG Signal (e) EOG Signal (f) MMG Signal

Fig. 1 Bioelectrical signals [2]

3 The Advantage of Bioelectrical Signals Over Conventional Biometrics

Biometric authentication systems use a variety of physical or behavioural characteristics including fingerprint, face, hand geometry, iris and voice pattern of an individual to establish identity. By using biometrics it is possible to establish an identity based on who you are, rather than by what you possess, such as an ID card, or what you remember, such as a password [5]. Although this conventional biometrics is unique identifiers, they are not confidential and neither secret to an individual since people put biometric traces anywhere. So, the original biometric can be easily obtained without the permission of the owner of that biometric. For example, in case of fingerprints, an artificial finger, known as a gummy finger, can be made by pressing a live finger to plastic material, and then mould an artificial finger with it or by capturing a fingerprint image from a residual fingerprint with a digital microscope, and then make a mould to produce an artificial finger [6]. In addition, thanks to the recent advancement in digital cameras and digital recording technologies, the acquisition and processing of high quality images and voice recordings has become a trivial task. Therefore, Iris scanners can be spoofed with a high resolution photograph of an iris held over a person's face [7]. The vulnerability of conventional biometrics to spoof has caused considerable concern especially in those fields that require high reliable user authentication. This heightened concern leads to great interest in assessing the probability and efficiency of using bioelectrical signals in authentication systems. Using bioelectrical signals as biometrics offers several advantages. In addition to their uniqueness, bioelectrical signals are confidential and secure to an individual. They are difficult to mimic and hard to be copied. To be more precise, the biological information of a person is genetically governed from deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) proteins. Eventually, the proteins are responsible for the uniqueness in the certain body parts. Similarly, the organs like heart and brain are composed of protein tissues called myocardium and glial cells, respectively. Therefore, the electrical signals evoked from these organs show uniqueness among individuals [4]. So, by using bioelectrical signals as biometrics we can benefit from sufficiently invulnerable authentication systems.

4 The Electroencephalogram Signal as a Biometric

As mentioned above the electroencephalogram (EEG) signal is one of the bioelectrical signals generated by brain activity, and can be recorded by positioning voltage sensitive electrodes on the surface of the scalp "Fig. 2". Typically, from 11 to 256 electrodes are placed on the scalp, each provides a time series sampled at 5.5–1.5 kHz, and generated hundreds of megabytes of data that must be analyzed in order to extract useful information. The feature space of EEG data is very large

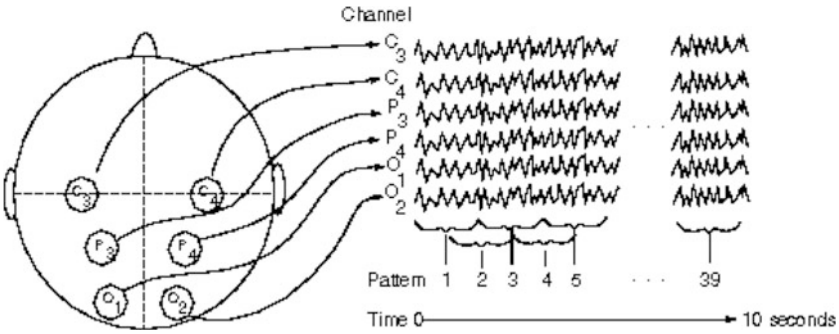


Fig. 2 Signal acquisition (www.cs.colostate.edu)

coming from the fact that information is usually accumulated throughout parallel (across every single electrode) as well as considering the human brain is really an extremely complex dynamical system [1]. The EEG can reflect both the spontaneous activity of the brain with no specific task assigned to it, and the evoked potentials, which are the potentials evoked by the brain as a result of sensory stimulus [8]. EEG-based authentication has been studied nowadays and researches have demonstrated that the EEG brainwave signals could be used for individual authentication. These researches can be categorized into three groups based on the type of signal acquisition protocol used in authentication task and the mental state of the subject during signal acquisition [9]; EEG recordings while relaxation with closed or open eye; EEG recordings while being exposed to visual simulation; EEG recordings while performing mental tasks. The example of each category is explained in the following:

Gui et al. [10] have presented an EEG-based biometric security framework. The data flow of authentication framework contained four steps. The first step was to collect raw EEG signals. 1.1 s of raw EEG signals was recorded from 6 midline electrode sites from 32 adult participants. Since it is argued that the brain activities are very focused during the visual stimulus process, the participants were asked to silently read an unconnected list of texts which included 75 words. In the next part, the noise level of raw EEG signals was reduced through ensemble averaging and low-pass filter. Ensemble averaging is a very effective and efficient technique in reducing noise because the standard deviation of noise after average is reduced by the square root of the number of measurements. After ensemble averaging, a 65 Hz low-pass filter was followed to remove the noise out of the major range of the EEG signals. In the third part, frequency features were extracted using wavelet packet decomposition. A wavelet is a mathematical function which can be used to divide a continuous-time signal into different scale component. A 4 level wavelet decomposition of the EEG signal after low pass filtering with 65 Hz was used to get the 5 EEG sub-bands, namely delta band (5–4 Hz), theta band (4–1 Hz), alpha band (1–15 Hz), beta band (15–35 Hz), and gamma band. Since the energy distributions of the frequency components are quite different for each individual, it was possible to adopt those frequency components as the features to represent the EEG signals. The mean, standard deviation and entropy were also calculated to form the feature vectors. So, there were 3 × 5 = 15 features for each subject. Finally, in classification part, the input feature vector was compared to the feature vectors that have been stored in dataset to authenticate the identity of the subject.

Nakanishi et al. [11] are also other researchers who have proposed new feature of EEG signals for authentication. They have used the concavity and convexity of spectral distribution in the alpha band of EEG signal in authentication to reduce the computational load for feature extraction, and authentication was done based on a linear combination of these features. They applied a consumer-use electroencephalograph that had only one electrode (single-channel) and was more convenient and practical compared to multi conventional channel measurements which increase the number of processing data, and require subjects to set a number of electrodes on the scalp. The single electrode was set on the frontal region of a head by using a head-band and subjects were asked to sit on a chair at rest with eye closed in quiet room that was the most suitable circumstances under which alpha wave can be detected. They adopted the spectrum analysis based on fast Fourier transform because it makes it easy to filter the spectrum in the alpha band and the concavity as well as the convexity of spectral distribution was used for distinguishing individuals. The concavity of spectral distribution was defined by detecting the maximum of the power spectrum and then calculating its tenth part and adopting it as a criterion. Then, frequencies of which power spectral values that were under the criterion were squared and summed. In addition to the concavity, the convexity of spectral distribution was another important feature. To define the convexity of spectral distribution the power spectral values in the alpha band were ranked and then the values and the frequencies of the top three were averaged. Next, the spectral values, which were greater than the averaged power spectrum, were summed. These three obtained features were as features which represent the convexity in spectral distribution. Finally, the subject authentication was done according to some calculation on combination of these obtained features.

Another research has been carried out by Liu et al. [12]. They recruited twenty right-handed subjects with normal or corrected-to-normal visual acuity and 64-channels EEG signals were recorded continuously by electrodes that were placed on the scalp. Two hundred and sixty color pictures were presented to the subject on a computer monitor located 1 m away from him. Stimulus duration of each picture was 3 s and all pictures were common and meaningful, identified and named easily. To find out suitable EEG features, several methods were employed to extract the EEG biometric features, including AR model, one of the most popular algorithms of feature extraction in which the series are estimated by a linear difference equation in time domain, power spectrum of the time-domain analysis that provides basic information of how the power distributes as a function of time, power spectrum of the frequency-domain analysis that provides basic information of how the power distributes as a function of frequency and phase-locking value which is a method to describe the synchronism between two signals. Then, all of the above-mentioned features were given to a support vector machine for classification respectively.

5 The Electrocardiogram as a Biometric

The heart makes use of electrical activity to activate the muscles required to pump blood through the circulatory system. By laying sensitive recording electrodes at certain regions around the heart, the signals can be recognized. The signals generated by the heart beat forms a regular pattern that records the electrical activity of the heart [1]. This signal is known as Electrocardiogram and can be used in human authentication. Recent works in the ECG biometric recognition field can be categorized as either fiducial point dependent or independent. Fiducials are specific points of interest on the ECG heart beat, namely, P, QRS and T waves that are shown in “Fig. 3”. By using these features a reference vector is produced to use for authentication. Israel et al. [13] have shown that ECG attributes are unique to each individual and can be used in human authentication. In their experimentation, data were collected at high temporal resolution from twenty nine individuals. At first step, a filter was designed and used to extract ideal data from raw ECG data and to locate fiducial positions by removing non-signal artifacts. The raw data contained both low and high frequency noise components associated with changes in baseline electrical potential of the device and the digitization of the analog potential signal respectively. After applying filtering, the ECG trace fiducial positions were located. For human identification, attributes were extracted from the P, R, and T complexes and four additional fiducial points which were named L′, P′, S′ and T′. Physically, the L′ and P′ fiducials indicate the start and end of the atrial depolarization and S′ and T′ positions indicate the start and end of ventricular depolarization “Fig. 4”.

Attributes that show the unique physiology of an individual were extracted by calculating the distance among the ECG fiducials. Classification was performed on heartbeats using standard linear discriminate analysis. A conversion was required to link the performance of the heartbeat classification to human identification. Standard, majority and voting were used to assign individuals to heartbeat data. The conversion was performed using contingency matrix analysis. Steven A. Israel et al. also demonstrated that the extracted features are independent of sensor location by

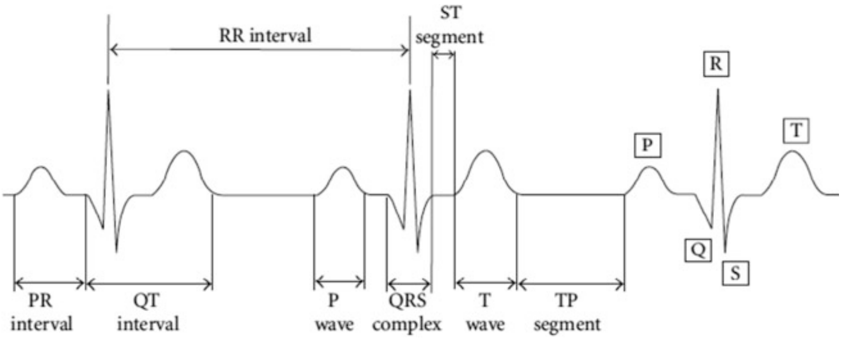


Fig. 3 A typical ECG signal that includes three heartbeats [4]

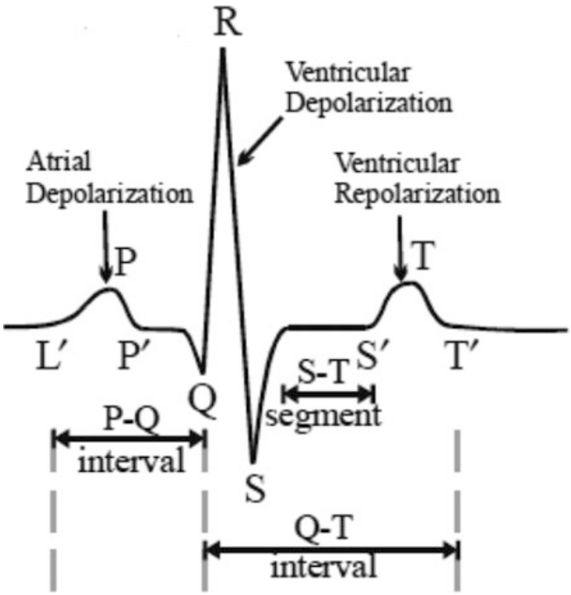


Fig. 4 Fiducial points' physical positions [13]

collecting ECG data at two electrode placements, one at the base of the neck and another one at fifth intercostals spacing. After testing they found a strong agreement between neck and chest ECG data which proved that the extracted ECG attributes are independent of sensor location. In addition, they proved that ECG attributes invariant to the individual's state of anxiety. Dey et al. [9] also used ECG as a biometric feature to authenticate a person. They generated an ECG feature matrix by using the features extracted from ECG, namely the time durations for the R-R, S-S, Q-Q, T-T, P-R, Q-T, and QRS intervals. Then, an inner product was performed between this feature matrix and a constant matrix. The product is then compared with a previously set threshold. If the result lied above the threshold, a binary value of 1 was assigned to it; otherwise 5. The combination of 1 and 5 produced the ECG-Hash code. After that, another ECG-Hash code was generated by using the original feature matrices and constant matrices in the same way as mentioned above. A matching was performed between these two ECG-Hash codes. On the event of a match, the individual was authenticated. Else, the authentication procedure failed.

Matos et al. [14] are other researchers that applied ECG as a biometric for human authentication by using the "the off-the-person approach". In this approach, as opposed to common ECG-based biometric systems that collects date by placing sensors on chest area, the ECG were acquired at the fingers with dry Ag/AgCl electrodes, and using a custom ECG sensor which consists of a differential sensor design with virtual ground when subjects were at resting situation. Then features were extracted based on a frequency approach and was based on Odinaka algorithm in which a single heart beat was divided into 64 ms windows, the analysis was performed in the frequency domain, computing the short time Fourier transform for each window. Finally a matching was performed on extracted features to do authentication.

6 The Electrooculogram as a Biometric

There are different types of eye movements like saccade and smooth pursuit which comprise enough information to human authentication, and among them saccade is the most popular and simplest for biometric authentication. According to measurement methods, eye movement signals can be divided into two groups: electrooculographical and videooculographical [2]. In Electrooculography the cornea-retinal potential that exists between the front and the back of the human eye is measured by placing electrodes left and right or top and above eye, and in video oculography the horizontal, vertical and torsional position components of the movements of both eyes are recorded by small cameras. Compared to other bioelectrical signals, fewer researches have been carried out in the field of applying eye oriented bioelectrical signals in human authentication. One of these few researches has been carried out by Abo-Zahhed et al. [15]. They have proposed a new biometric authentication based on the eye blinking waveform and used the Neurosky Mindwave wireless headset to collect the raw eye blinking signal of 25 healthy subjects. The headset is actually for recording EEG signals, but by placing the armed sensor which is made of dry electrode on forehead above the eye; it can be used to measuring EOG signals. Each subject was asked not to do any eye movement, and to make 1–12 eye blinks when signal recording was performing in quiet and normal temperature environment at daylight. The first step was isolating EOG signal from EEG signal through the technique of Empirical Mode Decomposition. Precisely, the raw EEG signal was decomposed into Intrinsic Mode Functions and after analyzing them, it was found that the first two IMFs belonged to EEG and others were related to EOG signals. After this step, eye blinking signal was extracted from EOG signal with the help of its largest amplitude in EOG signal. Then, a certain threshold was adopted to detect the positive and negative peaks of the eye blink. The next step was feature extraction and four groups of features were extracted based on time delineation of the eye blinking waveform and its derivatives “Fig. 5”.

Amplitude of positive peak of eye blink, area under positive pulse of eye blink, slope at the onset of positive pulse and position of positive peak of first derivative of eye blinking signal are one sample of each group. To evaluate the performance of system, the proposed system was tested under each four group of features, and based on achieving results, Abo-Zahhed et al. came to conclusion that the group of feature which was including area under positive pulse of eye blink, area under negative pulse of eye blink, energy of the positive pulse of eye blink, energy of the

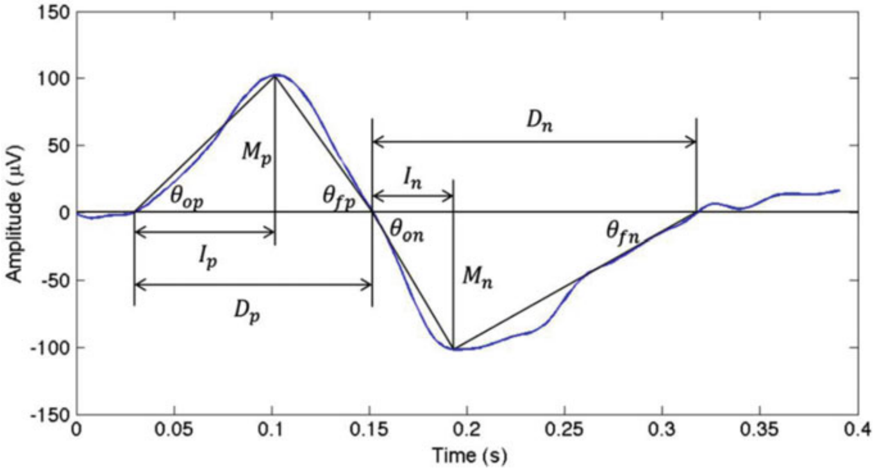


Fig. 5 Features extracted from eye blinking [11]

negative pulse of eye blink, average value of positive pulse of eye blink and average value of negative pulse of eye blink was the best for authentication of the subjects.

Juhola et al. [10] also have introduced a method in which a subject's saccade was applied to authentication. From their point of view, saccades are easy to stimulate and natural while reading or looking at the surroundings all the time. They decreased data for authentication process by using only the saccades parts of eye movements' signals. They asked each subject to sit down at a computer and the computer system had to verify him or her to be or not to be the authenticated subject. The system consisted of a device able to detect a subject's saccades and a program that computed features from saccades. They employed two small video cameras, one for each eye, to follow the pupils of a subject's eyes. Every subject was seated in chair at a fixed location and with the same distance from the stimulation device and was due to look at a small, horizontally jumping target and his or her eye movements were recorded for the authentication purpose. Signals given by this video-oculography system could be typically measured with a low sampling frequency, in this case with 35 Hz. After the recognition of every valid saccade, its amplitude, accuracy, latency and maximum velocity were computed to be used in authentication process "Fig. 6".

Latency is the time difference between the beginnings of the stimulus movement and response, accuracy is equal to the difference of the amplitudes of the stimulation and saccade and to compute the maximum angular velocity, the first derivative was approximated by differentiating an eye movement signal numerically and searching for the maximum velocity during the eye movement. They took these four particularly after having observed how clearly they varied between individuals. In addition, they applied EOG signal to user authentication and although the VOG signals contained less noise than the EOG signals, in most situations the EOG

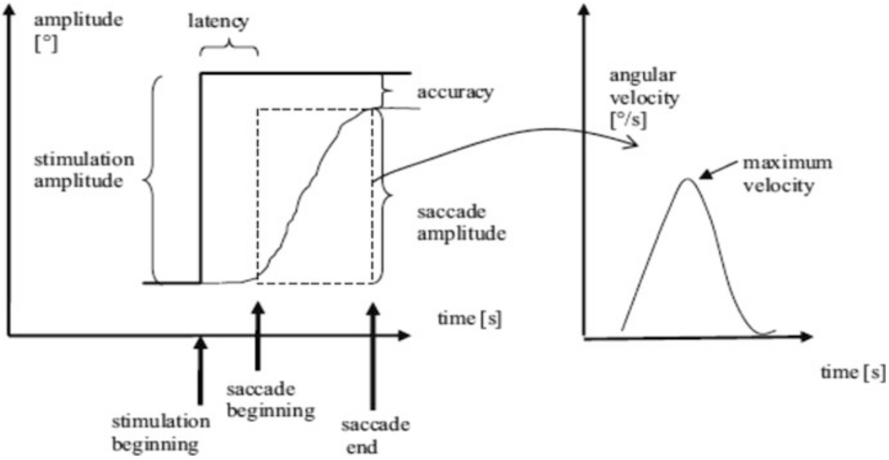


Fig. 6 An ideal saccade as a response to stimulation [11]

measurements achieved better results on the average than the VOG measurements. They supposed that the higher original sampling frequency of the EOG signals leads to better authentication results.

7 Conclusion and Discussion

This article has presented some of researches that have been carried out in the field of applying bioelectrical signals in human authentication. All of these researches agree that each bioelectrical signal has its own confidential physiological features which cannot be stolen and mimic. So, through these highly secured features, bioelectrical signals offer more advantage compared with conventional biometrics like fingerprint or iris for human authentication. But there are some issues and challenges involved in applying bioelectrical signals as biometrics. Firstly, all of mentioned researches have been done under laboratory condition with limited subjects. Therefore, the performance of bioelectrical -signal based authentication system might decline in practical real condition with more subjects secondly, the data acquisition of bioelectrical chest or EEG signals can be recorded by placing some electrodes over the scalp and the placement of electrodes to right position may cause distortion in the recorded signal. So, the data acquisition of bioelectrical signals could be an obstacle in applying these signals to human authentication in non-laboratory condition. Lastly, it should be considered that bioelectrical signals might be dependent to the mental and emotional state of subject. For example, fatigue, alcohol and aging could affect EOG signals, or EEG and ECG signals might vary with stress and anxiety.

References

1. Enderle JD, Bronzino JD (2012) Introduction to biomedical engineering. Academic press
2. Pal A, Gautam AK, Singh YN (2015) Evaluation of bioelectric signals for human recognition. Procedia Comput Sci 41:747–753
3. Van Den Broek EL, Spitters M (2013) Physiological signals: the next generation authentication and identification methods?. In: 2013 European intelligence and security informatics conference (EISIC). IEEE, pp 159–162
4. Singh YN, Singh SK, Ray AK (2012) Bioelectrical signals as emerging biometrics: issues and challenges. ISRN Sig Process 2012
5. Jain AK, Ross AA, Nandakumar K (2011) Introduction to biometrics. Springer Science & Business Media
6. Matsumoto T, Matsumoto H, Yamada K, Hoshino S (2002) Impact of artificial gummy fingers on fingerprint systems. In: electronic imaging 2002. International Society for Optics and Photonics, pp 275–219
7. Roberts C (2007) Biometric attack vectors and defences. Comput Secur 26(1):14–25
8. Hadjileontiadis LJ (2006) Biosignals and compression standards. In: M-Health. Springer US, pp 277–292
9. Dey M, Dey N, Mahata SK, Chakraborty S, Acharjee S, Das A (2014) Electrocardiogram feature based inter-human biometric authentication system. In: 2014 international conference on electronic systems, signal processing and computing technologies (ICESC). IEEE, pp 355–354
10. Gui Q, Jin Z, Xu W (2014) Exploring EEG-based biometrics for user identification and authentication. In: 2014 IEEE signal processing in medicine and biology symposium (SPMB). IEEE, pp 1–6
11. Nakanishi I, Baba S, Miyamoto C (2009) EEG based biometric authentication using new spectral features. In: International symposium on intelligent signal processing and communication systems, 2009. ISPACS 2009. IEEE, pp 651–654
12. Liu S, Bai Y, Liu J, Qi H, Li P, Zhao X, … Li Q (2014) Individual feature extraction and identification on EEG signals in relax and visual evoked tasks. In: Biomedical informatics and technology. Springer, Berlin, Heidelberg, pp 355–311
13. Israel SA, Irvine JM, Cheng A, Wiederhold MD, Wiederhold BK (2000) ECG to identify individuals. Pattern Recogn 31(1):133–142
14. Matos A C, Lourenço A, Nascimento J (2014) Embedded system for individual recognition based on ECG biometrics. Procedia Technol 17:265–272
15. Abo-Zahhad M, Ahmed SM, Abbas SN (2015) A novel biometric approach for human identification and verification using eye blinking signal. Signal Process Lett IEEE 22(7): 176–115

Recognition of Speech Isolated Words Based on Pyramid Phonetic Bag of Words Model Display and Kernel-Based Support Vector Machine Classifier Model

Sodabeh Salehi Rekavandi, Hamidreza Ghaffary and Maryam Davodpour

Abstract This study aimed to improve the classification of individual (isolated) words, and specifically, the numbers from one to twenty. In this study, a strong model was suggested to gain a unified view of voice. It is based on the idea of phonetic bag for voice that has been developed into a pyramid state. The pyramid idea can model temporal relationships. One of the problems of Support Vector Machine to classify words is its inability to model temporal relationships unlike hidden Markov models. Using the BOW-based pyramid idea in the extraction of the display containing temporal information of voice, the SVM can be given the capability of considering the time relationships of speech frames. One of the main advantages of Support Vector Machine model is its fewer parameters than the hidden Markov model. As the experiments' results have shown, it has much higher accuracy than the hidden Markov model in applications such as the recognition of single words, where the data set volume is limited. Using the pyramid BOW idea, the accuracy of SVM-based method can be increased as 20% compared to previous methods.

Keywords Speech recognition · Isolated words recognition · Classification of speech introduction · Display of phonetic bag of words · Support vector machine

S. S. Rekavandi (✉) · H. Ghaffary · M. Davodpour
Department of Computer Engineering, Islamic Azad University, Ferdows, Iran
e-mail: s_salehi19@yahoo.com

H. Ghaffary
e-mail: hamidghaffary53@yahoo.com

M. Davodpour
e-mail: Maryam.Davodpour2@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_2

1 Introduction

In this study, an efficient method based on pyramid bag of words (BOW) model and the SVM classifier model were provided to recognize isolated words. The provided BOW method has the ability to describe and model the temporal relationships in the speech, and by using kernel-based nonlinear support vector machine model can be used as an efficient technique used in recognition applications of isolated words.

Hedges et al. [1] studied the isolated words recognitions word using the support vector machine. In this method, first, the voice framed, and the Mel Frequency Cepstral Coefficients (MFCC) features extracted from each frame.

This stage is common in the most speech processing studies, and it indeed models a descriptive frequency of the frame. In fact, we expect the corresponding frames to have a MFCC feature vector similar to a particular part of a phoneme (e.g., frames related to explosion part of the explosive phoneme "b"). In other words, the difference is expected to be negligible. In this study, this stage as a conventional tool in describing a frame is constant in all discussing suggested methods. In their approach [1], the MFCC characteristics of each frame of a word (sound) is given to the Support Vector Machine (SVM) Classifier with the label of that word. For example, suppose a sound with the tag of "Five" includes 100 frames in 32 ms (with taking into account the overlap). Of these 100-frame, we calculate100 MFCC feature vector. Each of these 100 vectors (39-next) are labeled as "Five" and insert into the classifier. The same process is repeated during testing the training model with these difference that 100 labels predict by the SVM model. To obtain the label, the majority vote is considered among the 100 obtained predictions. This strategy has two major problems which we resolve them in this study.

To understand the first problem, consider this example that the phoneme "I" exists in both words of "Five" and "Nine". Thus, this method gives the frames related to this phoneme to the classifier with two different labels. Regardless of the classifier model, this strategy will disrupt the learning process of the model. In this research, we have resolved this problem by generating a unified display of speech based on bag of word (BOW) techniques. The second problem is the lack of modeling of temporal relationships in recognizing the words. In this study, using the pyramid-making idea of displaying BOW (Pyramid BOW), which has been highly regarded in recent years in the processing of images for modeling the spatial relationships, we provide a pyramid display model for voice (sound) that can model the temporal relationships (transposition of frame).

Models such as hidden Markov model inherently model the temporal relationships in the sound. However, in this study, we have used support vector machine as the classifier model.

The disadvantage of HMM models is their failure to have sufficient efficiency in small applications and recognizing isolated words. As a result, we would require massive datasets for their training. In fact, the number of HMM model parameters is very high, and in order to prevent the model overfitting, we need a lot of data. In HMM model, we need only to train a HMM model per word with a sufficient number of modes (for example, 6 modes). In each of these modes, we need to estimate the conditional probability of all observations. Suppose that the observations are possible for 50 MFCC models. Each of these patterns is related to different passes of one of the phonemes (e.g., the explosive section of “B”).

Thus, we need to estimate 6 * 50 conditional probabilities for each word. For 20 words, this number is 6000 parameters, which is a large figure compared to the number of data. However, the parameters can be somewhat reduced by techniques such as modeling at the phoneme level (each HMM models a phoneme). Of course, using such techniques requires providing the label at the level of the phonemes, which is a very time-consuming process; and at the same time, even if we consider two states for each phoneme, we should estimate 100 parameters, and to estimate the probabilities, we should have a high number of phonemes which do not practically make a significant change in the applications such as recognizing isolated words, but it can be used for continuous speech recognition. In methods like Support Vector Machine, using techniques such as reducing dimension, the number of model parameters can be controlled, and the overfitting of model can be prevented. Thus, the dimensionality reduction technique of principal component analysis (PCA) is raised. Therefore, this method is used to reduce the BOW-based feature vectors.

The results show the effectiveness of proposed methods to classify the isolated words.

2 Prior Research

In this section, first, the stages of extracting common characteristics of the sound signal are described. Then, the background of works related to displaying the bag of words and classification are described. In the next section, this method has been developed to classify speech isolated words.

3 Pre-processing and Feature Extraction

Several stages of recognition system are performed in the preprocessing phase. First, the speech is segmented into frames. Usually in speaker recognition applications, for better performance, the noise parts and the speech silence are eliminated. In this study, we have applied this stage as well.

In all branches of speech processing (speech recognition, word finding, speaker identification, etc.), the second phase is to extract feature from speech frames. Different feature vectors have been used for speech, including linear prediction coefficients, Mel Frequency Cepstral Coefficients (MFCC), wavelet coefficients and so on. In this study, the best and most effective ones, the MFCC has been used.

The Mel Frequency Cepstral Coefficients (MFCC) have been known as the most common and most widely used feature vector in processing of voice (audio) signal. After obtaining the filters bank energy, the feature vector of MFCC will be achieved by using discrete sine-cosine transform. In this section, the stages of feature extraction are explained below. The output of this stage is a feature vector sequence that each has been extracted from one of the input speech frames.

3.1 First Stage: Removing Silence from the Beginning and End of Words

In this research, for better efficiency, the silence at the beginning and end of words has been deleted using the method presented in [2]. This method has been implemented in MATLAB software at high speed[1]. This implementation is used in this study. The output of this method includes segments containing speech activity that the word's part of speech can be achieved by incorporating them. Voice Activity Detection (VAD), which is also called speech activity detection or speech recognition, is a process in the area of speech processing in which the presence or absence of human speech is recognized. Although the main use of this technique is in speech encoding and speech recognition, but it is also used in some other activities, such as speaker recognition. The goal in this method is to separate speech parts from silence and non-speech parts. The voice active areas usually refer to areas that are not related to environmental noise or silence. VAD methods extract parameters such as Linear Predictive Coding (LPC) distance, energy and zero crossing rate and compare these parameters with a set of threshold values to detect intervals including speech. Since these threshold values are estimated by analysis of silence periods, the classification accuracy of these methods highly reduces under unfavorable acoustic conditions. Normally, there is only noise in areas of the signal with silence. Through this measure with the ability to detect pure noise, it is possible to detect silence in the signal. The VAD problem is usually challenging in terms of low signal-to-noise (low SNR). Low SNR along with unstable noise signal can greatly reduce the precision of a VAD system. The basic methods for VAD detecting are based on signal energy. However, this measure does not work well when the SNR is low, since the energy of parts with sound activity is almost identical to noisy areas, and even in the unstable noise of energy, a measure becomes quite useless. The algorithm of method used to remove the silence at [2] is fully described.

[1] http://www.mathworks.com/matlabcentral/fileexchange/28826-silence-removal-in-speech-signals/

3.2 Second Stage: MFCC Feature Extraction Method

In this section, the detailed steps of MFCC method used in this study are described.

Suppose that $s_1, \ldots, s_{512}$ are examples of the studied frame. The stages of MFCC method used for each frame are as follows:

- Frame energy calculation: The mean square of frame samples

$$E = \frac{\sum_{i=1}^{512} s_i^2}{512}. \tag{1}$$

Applying 512-point Hamming window on $s_1, \ldots, s_{512}$

$$s_{w-1}, \ldots, s_{w-512} = h_1 s_1, \ldots, h_{512} s_{512} \tag{2}$$

- Calculating the FFT of windowed frame

$$f_1, \ldots, f_{512} = fft(s_{w-1}, \ldots, s_{w-512}) \tag{3}$$

- Calculating the result of 12 Mel filters (12 channels) on$f_1, \ldots, f_{512}$. At this stage, 12 $Z_1, \ldots, Z_{12}$ are obtained.
- Calculating 13 features by using the following equation:

$$\begin{aligned} &mfcc_1, \ldots, mfcc_{12} = DCT(\log(Z_1, \ldots, Z_{12})) \\ &mfcc_{13} = \log(E) \end{aligned} \tag{4}$$

Calculating 13 features by using the derivative of $mfcc_1, \ldots, mfcc_{13}$:

These features are called Delta. To calculate the derivative, every two consecutive numbers are subtracted (The first number is subtracted from the last number). The feature obtained in this step are called as $d_1, \ldots, d_{13}$.

- Calculating 13 features by using the derivative of $d_1, \ldots, d_{13}$: These features are called Delta-Delta. At this stage, the $dd_1, \ldots, dd_{13}$ is obtained.
- The final feature vector is as follows (including 39 real number):

$$F = [mfcc_1, \ldots, mfcc_{13}, d_1, \ldots, d_{13}, dd_1, \ldots, dd_{13}]. \tag{5}$$

Therefore, for each studied voice frame, a 39-item feature vector is extracted.

4 Display of Bag of Words

The display of bag of words (BOW) has been primarily inspired in the field of image processing from the field of text processing [3]. As the number of each word can be easily counted within a text, our goal here is to count the patterns in an image or a sound. In using BOW-based methods in image, initially, the possible patterns in a dictionary are learned. For example, an eye pattern can be one of the patterns available in the dictionary. This idea has been widely used in image processing [4–6]. In audio processing tasks, this method has been sometimes introduced as Bag of Acoustics [7]. This method has been regarded in recent years in the issue of sense detection [7] and recognition voice from event [8].

5 Dimensionality Reduction of Principal Component Analysis

In dimensionality reduction methods, a multi-dimensional space are mapped to a space of lower dimension. With reducing the dimensions of the original space, the number of model parameters would reduce, and thus, the probability of model overfitting will decrease. PCA dimensionality reduction is as such to maintain information as much as possible. In addition to this feature, the PCA method finds the direction of highest changes and depict the data in those directions. Therefore, it is a useful feature transfer method that is used in most applications of pattern recognition [9–11]. In this study, after extraction simple and pyramid BOW display provided, this method has been used to reduce the dimensions.

6 Suggested Method

In this section, first, the proposed method for finding a BOW-based display of input speech is described. Then, the idea has been developed to model temporal relationships in the speech into a pyramid way. Finally, the diagram block of the proposed method is provided.

7 Display of Bag of Acoustics

Figure 1 shows the proposed method to obtain a BOW display of an acoustic signal.

As can be seen in Fig. 1, a dictionary including K patterns (templates) is provided (The dictionary learning method is described in the next section). Each input

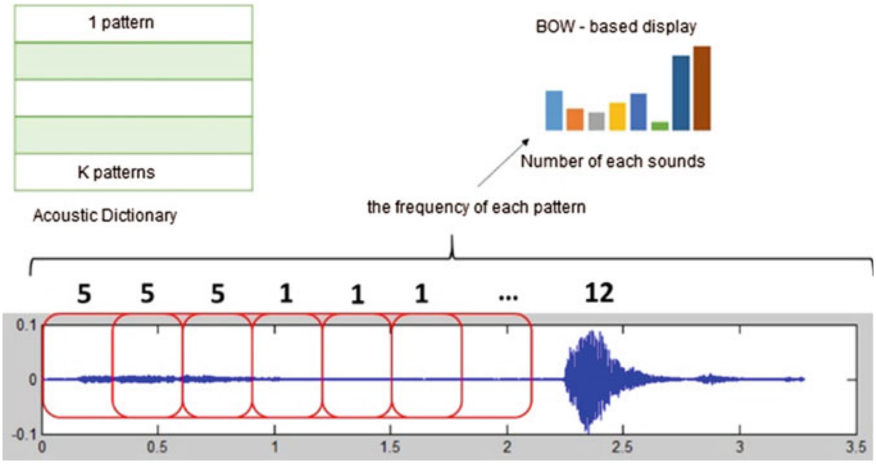


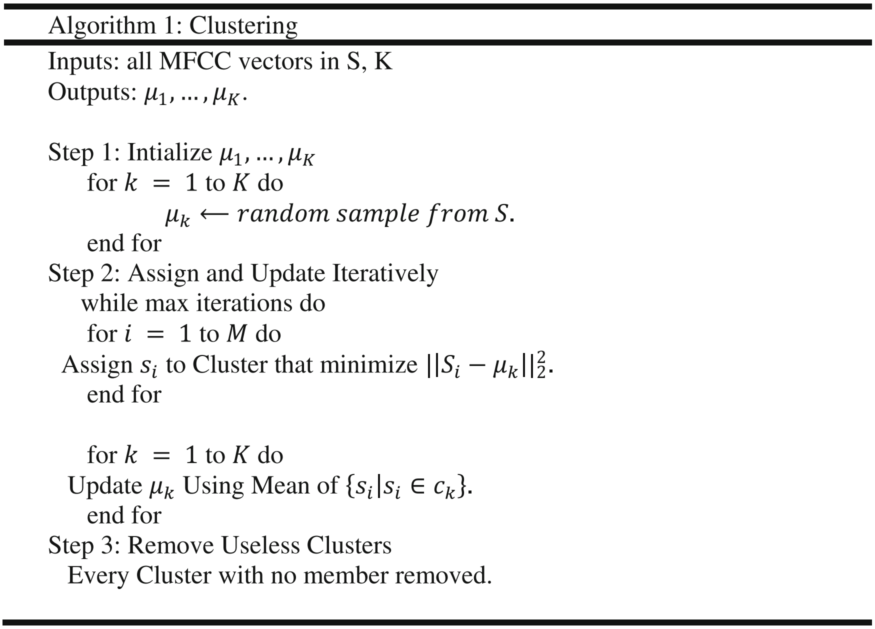
Fig. 1 Extraction of BOW display from a sound (first suggested method)

sound is divided into consecutive frames with overlapping. The MFC features are extracted from each of these frames. For each frame, the closest MFCC model in the dictionary is found. After this stage, the number of each model can be counted. Therefore, we have a display resulting from the frequency of K patterns in the sound. In this study, we will solve one of the fundamental problems of the basic method by using the BOW method, in which each frame is given to the classifier separately. However, there is another problem in this method. Although we have considered many different acoustic patterns in the display of sound, but no information has been modeled about their order. This problem has been solved by using the idea of pyramid-making of BOW display [12], which has been highly regarded in recent years in images processing for modeling spatial relationships [13, 14].

8 Learning of Phonetic (Acoustic) Dictionary

To learn phonetic (acoustic), dictionary any clustering method can be used. In this study, we have used the known k-means method. First, the 39-item MFCC vectors are extracted from all frames of total sounds in the training data set. The goal is to learn K cluster centers (phonetic pattern) of these vectors in such a way that the quantization error is so small. Quantization error refers to the difference of each vector with the nearest cluster, i.e. a cluster that belongs to it.

Therefore, at this point, it is assumed that M MFCC vectors have been selected as $S = \{s_1, s_2, \ldots, s_M\}$. Now, it is just enough to train K patterns of the vectors within the S. to this end, the S vectors must be grouped into K clusters $\{S_k, k = 1, 2, \ldots, K\}$, while clusters have patterns different from each other. For this



```
Algorithm 1: Clustering
Inputs: all MFCC vectors in S, K
Outputs: $\mu_1, \dots, \mu_K$.

Step 1: Intialize $\mu_1, \dots, \mu_K$
    for $k = 1$ to $K$ do
            $\mu_k \longleftarrow random\ sample\ from\ S$.
    end for
Step 2: Assign and Update Iteratively
    while max iterations do
    for $i = 1$ to $M$ do
  Assign $s_i$ to Cluster that minimize $||S_i - \mu_k||_2^2$.
    end for

    for $k = 1$ to $K$ do
  Update $\mu_k$ Using Mean of $\{s_i | s_i \in c_k\}$.
    end for
Step 3: Remove Useless Clusters
  Every Cluster with no member removed.
```

Fig. 2 Clustering algorithm to learn the phonetic dictionary

reason, it is enough to do the clustering based on MFCC features, since it is proportional to the human auditory system.

If the k-means clustering algorithm is applied on these vectors, the vectors in the S are divided into K clusters, $C_1, \ldots, C_K$, and the μ_k phrase is chosen as the center of cluster C_k. *The* K-means clustering algorithm is as follows: (Fig. 2).

In the first phase of algorithm 1, the centers are initialized. The second phase varies repeatedly between the two stages of attributing to the cluster centers and updating the centers until reaching the desired number of repetitions. The third step removes all clusters with no members. After applying the k-means algorithm, the cluster centers show the intended phonetic dictionary in the MFCC space.

9 Display of Pyramid Bow to Model Temporal Relationships in Speech

The pyramid-making idea to fix the problem of BOW display in modeling spatial relations in the image was raised for the first time in [9], and has been of great concern in the field of image processing so far. Models such as hidden Markov models inherently model temporal relationships in the sound. But as noted, in this study, we have used the support vector machine as the classifier model.

The disadvantage of HMM models, which causes their inefficient use in small applications and recognizing individual words, is the need to massive dataset for training them. In fact, the number of HMM model parameters is very high, and a lot of data is required to prevent the model overfitting. In methods such as support vector machines, using techniques such as dimensionality reduction, the number of model parameters can be controlled, and the model overfitting can be prevented. The idea of pyramid-making of BOW relies on fragmentation of the image to the required level and calculating the frequency of patterns in each slice (Fig. 3). For example, if we tell someone that there are two models of eyes and a nose pattern in an image, it cannot be expected that the person can guess where on the image the patterns occur. But with pyramid-making of BOW display, this problem goes away.

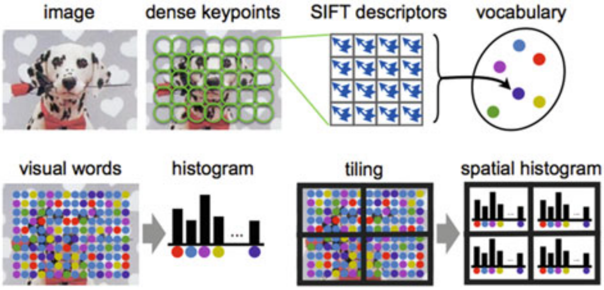
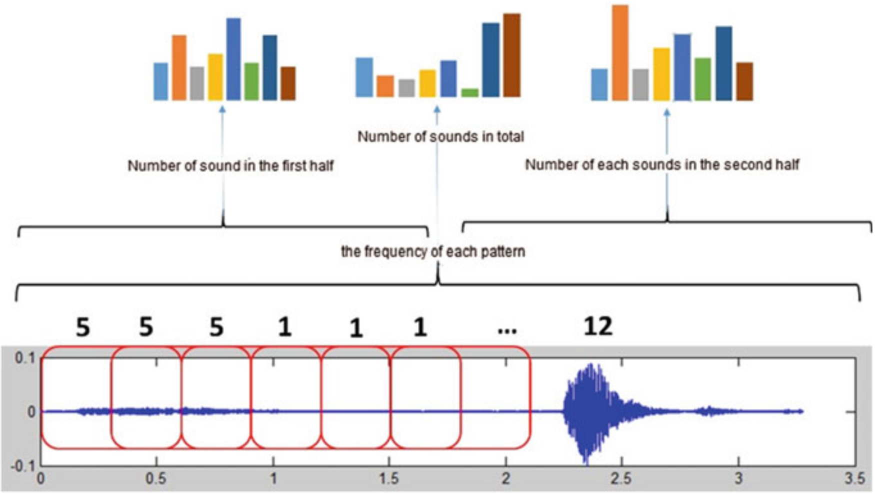


Fig. 3 BOW pyramid display in the image



Number of sound in the first half
Number of sounds in total
Number of each sounds in the second half
the frequency of each pattern
5
5
5
1
1
1
...
12
0.1
0
-0.1
0
0.5
1
1.5
2
2.5
3
3.5

Fig. 4 Pyramid display of BOW in the sound (second alternative method)

In this study, we have used the idea of pyramid-making for modeling temporal relationships in the sound. Figure 4 shows the proposed approach for pyramid making of display in the sound.

Two levels are used in Fig. 4. If needed, the number of levels can be increased. In the next level, four areas are achieved. If the words are long and the number of phonemes of each word is high, higher number of levels would more appropriate.

10 Diagram Block of the Proposed Method

In previous sections, the proposed methods based on simple and pyramid BOW display were described. In this section, the steps of the proposed method are summarized. Figure 5 shows the diagram block of the model teaching stage.

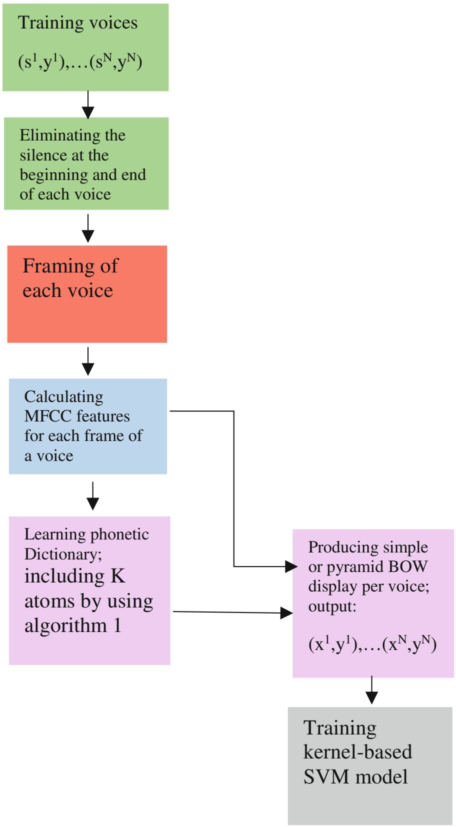

Training voices
$(s^1,y^1),\ldots(s^N,y^N)$
Eliminating the silence at the beginning and end of each voice
Framing of each voice
Calculating MFCC features for each frame of a voice
Learning phonetic Dictionary; including K atoms by using algorithm 1
Producing simple or pyramid BOW display per voice; output:
$(x^1,y^1),\ldots(x^N,y^N)$
Training kernel-based SVM model

Fig. 5 Diagram block of learning algorithm of the classifier model

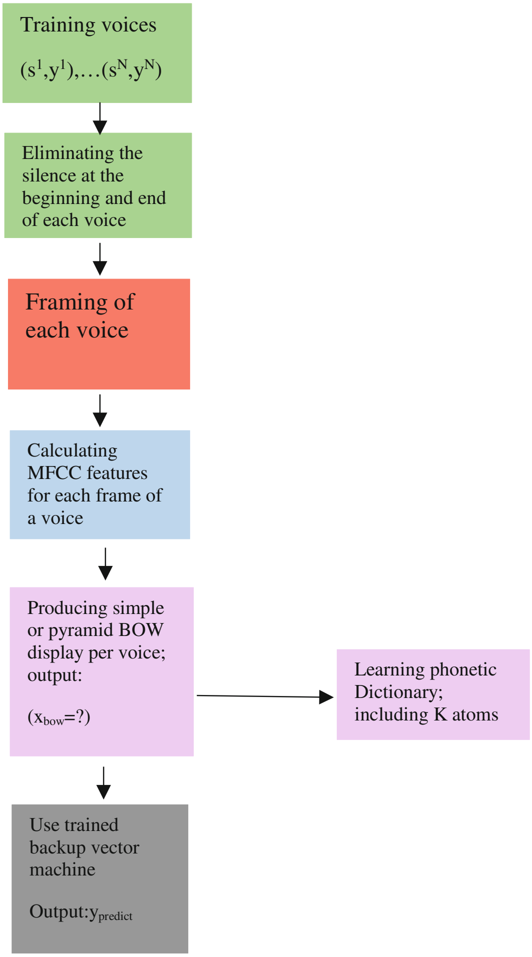

Training voices
$(s^1,y^1),\ldots(s^N,y^N)$
Eliminating the silence at the beginning and end of each voice
Framing of each voice
Calculating MFCC features for each frame of a voice
Producing simple or pyramid BOW display per voice; output:
$(x_{bow}=?)$
Learning phonetic Dictionary; including K atoms
Use trained backup vector machine
Output:$y_{predict}$

Fig. 6 Block diagram of classifying a new data

After teaching the phonetic dictionary and the classifier model, they can be used to classify a new data. The diagram block is the use of the trained model to predict the label of a data as Fig. 6.

- Training voices
- Eliminating the silence at the beginning and end of each voice
- Framing of each voice
- Calculating MFCC features for each frame of a voice
- Learning phonetic Dictionary; including K atoms by using algorithm 1
- Producing simple or pyramid BOW display per voice; output

- Training kernel-based SVM model
- Training voices
- Eliminating the silence at the beginning and end of each voice
- Framing of each voice
- Calculating MFCC features for each frame of a voice
- Learning phonetic Dictionary; including K atoms by using algorithm 1
- Producing simple or pyramid BOW display per voice; output
- Training Kernel-based SVM model.

11 Experiments

In this section, we test the basic techniques and the described method. First, the training data set is described. Then, the evaluation criteria are described. Finally, the settings related to experiments and the test results are given.

12 Training Data Set

Training data set includes 10 different speakers. Each of these speakers have uttered words 1–20 once, the words with sampling frequency of 16 kHz have been recorded. The data related to 7 speakers have been selected as training data, and 3 speakers as the test data.

13 Noisy Dataset

A noisy data set has been also made to evaluate the effectiveness of models in the presence of ambient noise. This data set has four samples per voice (in the previous section):

- Original sound version
- Signal-to-noise: 30 dB
- Signal-to-noise: 20 dB
- Signal-to-noise: 10 dB

Therefore, this dataset contains 28 samples per word in the training data set and 12 samples per word in the test data set.

14 Feature Extraction

The feature is segmented to 32 ms (ms) frames with an overlap of 16 ms. For better performance, as mentioned in previous sections, the noise and silence parts have been removed before framing. At this stage, by the number of frames per voice, the MFCC 39-fold vectors are obtained.

15 Classification Tests Results

Table 1 shows the test results of basic models and suggested methods. The HMM method is indeed discrete HMM method. The number of optimal HMM states for words was obtained as 6. The number of clusters in the HMM method was equal to 40 optimized ones, while the number of clusters in BOW-based and pyramid BOW methods was equal to 100 optimized clusters. In BOW and pyramid BOW methods, the PCA method was used to reduce the dimensions.

16 Classification Analysis

16.1 Number of Optimal Clusters

The results of Table 1 show that HMM method has low accuracy in the recognition of isolated words, and it was expected due to the high number of HMM model parameters compared to the number of data. The number of optimal clusters in HMM method (40 clusters) was lower compared to methods based on SVM (100 clusters). A total of 40 clusters, especially in noisy data, cannot model a variety of phonetic patterns. However, by increasing the number of clusters instead of increasing the precision, we would have accuracy reduction as well. The reason for this phenomenon is that by increasing the number of clusters, the number of model parameters will extremely increase and there is no way to control the number of

Table 1 Classification accuracy of isolated words

Method name	Accuracy of the normal data set		Accuracy of the noisy data set	
Words	20-1	5-1	20-1	5-1
Random	5%	20%	5%	20%
HMM	28%	53.3%	15.3%	41.6%
Basic method [1]	31%	59.4%	18.2%	46.6%
Display + BOW SVM	*50.3%*	**73.3%**	*45.8%*	**63.3%**
Display BOW + Pyramid SVM	*72%*	**96%**	*68.6%*	**84.6%**

parameters. The number of parameters of SVM model is inherently lower than the HMM method. In addition, the use of PCA method can control the number of parameters.

17 Comparing the SVM-Based Method with Bow SVM

The results show that the proposed method in [1, 15], which has been reported as the basic method in Table 1, has also a less accuracy than the proposed conventional BOW method. In basic method [1], the MFCC features of each frame from every word (sound) are given tagged with that word to the support vector machine classifier. For example, suppose a sound with the tag of “Five” includes 100 frames in 32 ms (with taking into account the overlap). Of these 100 frames, 100 MFCC feature vectors are achieved. Each of these obtained 100 vectors (39-next) are labeled “Five” and given to the classifier. The same process is repeated when testing the training model; the difference is that 100 labels are predicted by the SVM model. To obtain the label, the majority vote is taken among the 100 predictions. This strategy has two major problems, which are addressed in this study. To understand the first problem, consider this example that the phoneme “I” exists in both words of “Five” and “Nine”. Thus, this method gives the frames related to this phoneme to the classifier with two different labels. Regardless of the classifier model, this strategy will disrupt the learning process of the model.

18 Comparison of Bow Method with Pyramid Bow Approach

The results show a significant increase in the accuracy of the pyramid display compared to the typical BOW. As described, the pyramid display model the temporal information in the sound and extracts more information from the sound.

19 Methods Resistance Against Noise

Among the methods, BOW and BOW pyramid methods, which are the proposed methods, have a high resistance to noise. The reason is that the noise in the MFCC features partly disappears in the quantization phase (in clustering) as clustering error, but in the basic method [1], this noise gives itself as a part of the feature vector to the SVM model. Clustering methods inherently eliminate the noise partly as quantization error.

20 Conclusion

This paper aimed to classify the isolated speech words. First, a pyramid model proposed based on BOW to display the voice which is able to increase the prediction power of support vector machine model. This model could model the temporal relationships in the sound. Using this model, the accuracy of support vector machine classifier based methods significantly improved. In this study, we demonstrated that the dimension reduction techniques are useful for increasing the accuracy of Support Vector Machine.

References

1. Hegde S, Achary KK, Shetty S (2012) Isolated word recognition for Kannada language using support vector machine. In: Wireless networks and computational intelligence. Springer, Berlin Heidelberg, pp 262–269
2. Giannakopoulos T (2009) A method for silence removal and segmentation of speech signals, implemented in Matlab. Department of Informatics and Telecommunications, University of Athens, Greece, Computational Intelligence Laboratory (CIL), Insititute of Informatics and Telecommunications (IIT), NCSR DEMOKRITOS, Greece
3. Gabriella C, Dance C, Fan L, Willamowski J, Bray C (2004) Visual categorization with bags of keypoints. In: Workshop on statistical learning in computer vision, ECCV, vol 1, no 1–22, pp 1–2
4. Yang, Jun, Yu-Gang Jiang, Alexander G. Hauptmann, and Chong-Wah Ngo. "Evaluating bag-of-visual-words representations in scene classification." In: *Proceedings of the international workshop on Workshop on multimedia information retrieval*, pp. 197–206. ACM, 2007
5. Ramesh B, Xiang C, Lee TH (2015) Shape classification using invariant features and contextual information in the bag-of-words model. Pattern Recogn 48(3):894–906
6. Yang Y-B, Zhu Q-H, Mao X-J, Pan L-Y (2015) Visual feature coding for image classification integrating dictionary structure. Pattern Recogn
7. Pokorny FB, Graf F, Pernkopf F, Schuller BW (2015) Detection of negative emotions in speech signals using bags-of-audio-words. In: 2015 International conference on affective computing and intelligent interaction (ACII). IEEE, pp 879–884
8. Grzeszick R, Plinge A, Fink GA (2015) Temporal acoustic words for online acoustic event detection. In: Pattern recognition. Springer International Publishing, pp 142–153
9. Wu P, Hoi SCH, Xia H, Zhao P, Wang D, Miao C (2013) Online multimodal deep similarity learning with application to image retrieval. In: Proceedings of the 21st ACM international conference on Multimedia. ACM, pp 153–162
10. Quan C, Wan D, Zhang B, Ren F (2013) Reduce the dimensions of emotional features by principal component analysis for speech emotion recognition. In: 2013 IEEE/SICE International symposium on system integration (SII). IEEE, pp 222–226
11. Chiou B-C, Chen C-P (2013) Feature space dimension reduction in speech emotion recognition using support vector machine. In: Signal and information processing association annual summit and conference (APSIPA), 2013 Asia-Pacific. IEEE, pp 1–6
12. Lazebnik S, Schmid C, Ponce J (2006) Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: 2006 IEEE computer society conference on computer vision and pattern recognition, vol 2. IEEE, pp 2169–2178

13. Kim SY, Sohn K-A (2015) Mobile phone spam image detection based on graph partitioning with pyramid histogram of visual words image descriptor. In: 2015 IEEE/ACIS 14th international conference on computer and information science (ICIS). IEEE, pp 209–214
14. Lan Z, Hauptmann AG (2015) Beyond spatial pyramid matching: space-time extended descriptor for action recognition. *arXiv preprint* arXiv:1510.04565
15. Seryasat OR, Aliyari-shoorehdeli M, Honarvar F (2010) Multi-fault diagnosis of ball bearing based on features extracted from time-domain and multi-class support vector machine (MSVM). In: IEEE international conference on systems man and cybernetics (SMC), pp 4300–4303

A Novel Improved Method of RMSHE-Based Technique for Mammography Images Enhancement

Younes Mousania and Salman Karimi

Abstract Contrast improvement is one of the most important steps in medical image enhancement procedures such as mammography. In this paper, a combination of best features related to direct and indirect histogram equalization techniques is proposed in a two dimensional workspace. Using different advantages of these methods, while the proposed algorithm is able to improve the contrast and brightness of mammography images, it could decrease different effects of noises, too. On the other hand, in order to reduce undesirable effects of traditional histogram equalization techniques, an improvement of recursive mean-separate histogram equalization using a fusion of contrast-limited adaptive histogram equalization is proposed, too. Evaluation results using four effective measurement techniques e.g. peak signal-to-noise ratio, **mean squared error**, **absolute mean brightness error** and effective measure of enhancement, shows that the suggested method has significant results in contrast enhancement.

Keywords Contrast-limited adaptive histogram equalization · Recursive mean-separate histogram equalization · Effective measurement enhancement Peak signal-to-noise ratio · Mammographic image · Mean squared error Absolute mean brightness error

1 Introduction

Breast cancer is the second main disease after lung cancer that causes death in women. Breast cancer and fibroids are among the masses that are common among women and if detected on time, the process of recovery and treatment will increase

Y. Mousa nia (✉) · S. Karimi
Department of Electrical and Electronic Engineering,
Lorestan University, Khorram-abad, Lorestan, Iran
e-mail: Mousania.yo@fe.lu.ac.ir

S. Karimi
e-mail: Karimi.salman@lu.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_3

substantially [1, 2]. In patients with symptoms that are suspected to have cancer or associated with cysts and other organs, the physician performs mammography. In this method, sound waves are used to create images of various parts of the body, including the breast and X-rays do not play any role. In cases where the breast tissue is very dense or the age is less than 30 years old, the doctor prescribes ultrasound. It should not be forgotten that ultrasound is not a substitute for mammography, but is an adjunct to it.

Ultrasound is currently the best way to diagnose breast cysts, which is similar in appearance to your full masses [3]. Most of women's problems are related to chest pain and swelling in the breasts which leads to inability to do daily works. Breast cancer is the result of an out-of-body growth in abnormal breast cells. In both benign and malignant tumors, there is a rapid and high growth of the cells. The process of increasing cells in benign tumors stops at a definite stage [4]. In malignant tumors, this growth continues unabashedly to an extent that, in the absence of treatment, affects all parts of the body and fails to work. Throw away The most common type of breast cancer is cancer of the origin of ducts, and since this type of tissue is found to be in the upper and lower quarters of the breast, about half of the breast cancers are found in the upper and outer quarters.

It should be noted that in all tumors there is a rapid and high growth of cancer. What is important and the main difference between these two types of tumors is that the process of increasing cells in benign tumors stops at a definite stage, but continues in the non-inhibiting tumors of the malignant tumor. The cell growth in the malignant tumors continues to some extent, which, if not treated, affects all parts of the body and abilities. While this does not happen in benign tumors. No matter how much breast cancer is diagnosed earlier, treatment is easier and more successful. For this reason, women need to know the facts about the disease in order to protect their health.

Mammography is the only sure-tier method by which one can reveal a mass in the chest before being detectable by touch. Mammograms are divided into two main categories according to which direction they are coming from: the craniocaudal taken from the top to the bottom and the mediolateral axis that is taken in half-fold and perverted [5]. The purpose of this work is to examine the chest in different ways in order to better detect lesions.

Micro-calcifications one of the symptoms that are used to detect early breast cancer. Each micro-calcification appears as a bright grain, which has several pixels in digital images and these pixels are brighter with respect to their adjacent pixels [6].

Detection of suspicious areas in a mammogram that includes micro-calcification clusters is usually performed by the radiologist, but it is difficult to determine if a particular cluster is associated with a benign or malignant process [7]. However, because the size and shape of each micro-calcification is different, and also the texture of the mammogram context is heterogeneous, the grains cannot be easily identified individually. In other words, due to the low contrast of the mammographic images, the precise diagnosis of the cancer symptoms such as masses and calcification is difficult for the radiologist. Over time, radiologists have empirically discovered the rules that decide on the appearance of micro-calcification, their dispersal, and other features such as the benign or malignant micro-calcification cluster.

All local features of the original image are extracted by the radiologist's vision system, which is usually not done accurately.

Normally, if the patient's mammogram is suspected of having a micro-calcification cluster, it will be introduced for biopsy (tissue sampling). Different evaluations show that out of all four biopsy surgeries, only one of them is successful. In terms of the factors that endanger the health of the patients, biopsy is not a suitable surgery and it is preferred to avoid this as much as possible [8]. In this way, finding a technique to differentiate between benign and malignant samples in the most accurate way is very helpful in preventing unnecessary biopsies. Hence, relying on image processing techniques in this field is seen necessary to diagnose micro-calcification tissues as the best way as possible. The techniques of the digital image processing are used to enhance the quality of digital mammogram images as well as to increase the detection accuracy of micro calcifications [8, 9]. Improvement of the image contrast is one of the most important requirements used in image processing and vision system applications. In general, methods of the contrast improvement are divided into two major categories: direct method and indirect method [10].

2 Direct Contrast Optimal Methods

In the direct methods, while defining a criterion for measuring the image contrast, attempts are made to improve image contrast by improving this criterion. Creating an appropriate measurement criterion for image contrast is an important stage to improving the image directly. The direct contrast approach considers both the general and local information of the image, hence it can outperform in many applications. In this regard various approaches have been proposed that are based on the phase entropy principle, which transmits the image to the phase domain, and the phase entropy is calculated, and in this manner the local contrast is measured [11].

3 Indirect Contrast Optimal Methods

Improving contrast with the indirect method involves modifying the histogram of the image. In indirect method, the dynamic range of the gray levels of the image is increased to improve contrast. Indirect methods which have been paid more attention in recent years due to direct and knowledge-based representation are categorized into four categories:

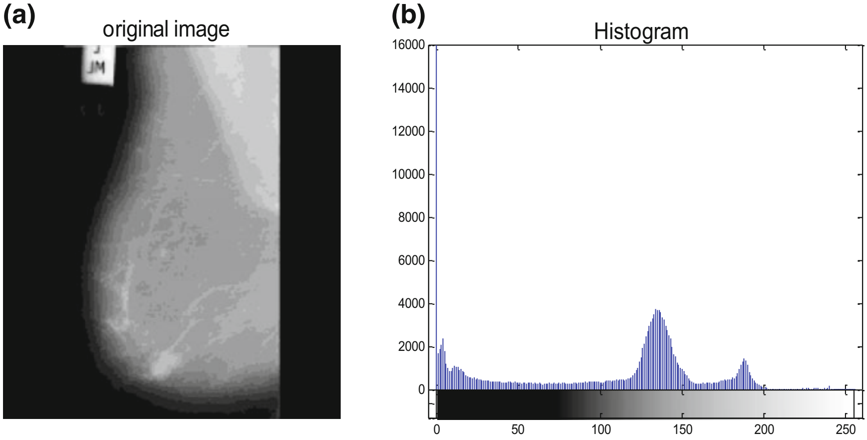


Fig. 1 **a** Original mammography image. **b** Histogram image

- Methods that modify the up and down frequency components of the image [12]
- Methods based on Conversion [12, 13]
- Methods based on histogram modification [14, 15]
- Methods based on Soft calculation [16].

The proposed algorithm and techniques presented in this paper are based on histogram correction methods. In Fig. 1, a mammography image with its histogram is displayed.

3.1 Histogram Equalization (HE)

The main idea of HE is mapping of the values of the input image intensity to the new intensity values through a transformation function created for the cumulative density function (CDF). First, HE converts the histogram of the original image to a plane histogram using an average value that is the average range of gray levels [17]. Therefore, the histogram of the image is divided into two parts based on its average gray level, and the HE algorithm is separately applied on each divided section of the histogram. Secondly, histogram equalization performs the improvement action based on the overall content of the image.

HE is powerful in highlighting the boundaries and edges between different objects, but it may change the local details in these objects, particularly smooth and small areas. The other problem of HE is an abnormal increase and saturation effects of intensity and also it is not appropriate to maintain the brightness of the original image due to the changes in the brightness of the image [18].

3.2 Contrast-Limited Adaptive Histogram Equalization (CLAHE)

CLAHE is a kind of adaptive equalization of the histogram. This method divides the original image into several sub-images without overlapping [19]. The secondary histogram of the images is limited to the value of the improvement per each pixel and then equalization is performed. Details of the image are evidently revealed with respect to the background [20]. At the same time, the contrast of the image is improved equally, which results in an output contrast image with high quality [9]. In this paper, using an adaptive filtering procedure, the histogram of different parts of the partitioned image is calculated and then the histogram balancing is utilized to rearrange the brightness values of the total image. So our proposed method is different from the smoothing of the fundamental histogram, since in this method, as a traditional equation technique, only one histogram is used for the whole image [21].

Consequently, for the purpose of improving the localized image contrast and extracting more details from the image, while significant noise would be generated, the contrasting histogram is equalized.

In order to suppress these deficiencies, a generalization of Adaptive Histogram Equalization (AHE) of a contrast-limited, or concise, which is called CLAHE, is used.

This technique is designed to overcome the problem of noise exacerbation. CLAHE does not deal with the entire image, but deals with pieces that are in small areas of the image [22]. The contrast of each area is improved in such a way that the histogram of the output region corresponds to approximately the histogram expressed by the distribution parameter.

Neighbor sections are combined to eliminate abnormal induced boundaries by using bidirectional interpolations [22]. Utilizing contrast in homogeneous regions, it is possible to avoid any exacerbation of any unwanted noise that may be present in the low contrast image. Besides user friendly, simple calculation and good output in local areas are of the advantages of CLAHE. Additionally, CLAHE has less noise and can maintain the light saturation which normally occurs in the histogram equalization procedures [23–25].

3.3 Recursive Mean-Separate Histogram Equalization (RMSHE)

One of the first suggestions to overcome the drawbacks of the HE method is brightness preserving of the equalized bi-histogram (BBHE).The method preserves the effective amount of image brightness while improving the contrast. Moreover, it divides the histogram into two sub-histograms based on the average amount of the brightness and equalizes each part individually. If X_m denote the mean of the image X and assume that $X_m \in \{X_0, X_1 \ldots X_{L-1}\}$. Based on the mean X_m the input image is divided into two sub level images X_L and X_U. The transform functions for the sub images are defined as

$$F_L(X) = X_0 + (X_m - X_0)C_L(X) \tag{1}$$

$$F_u(X) = X_{m+1} + (X_{L-1} - X_{m+1})C_u(X) \tag{2}$$

According to the above equations, $C_L(X)$ and $C_U(X)$ is the respective cumulative density functions for X_L and X_U.

The output image (Y) of BBHE, is expressed as

$$Y = F_L(X_L) \cup F_u(X_u) \tag{3}$$

Now we introduce a better technique called RMSHE, which in fact performs the same BBHE algorithm as a recursive one. In aforementioned techniques the input image histograms were divided into two parts. However, in this method, instead of dividing the input image one time, the input image divides to 2^n sub-histograms using an optional criterion called n. Then, each of these sub-histograms is equalized in dependently. When $n = 0$, it means that no sub-image is created, which is the same as the HE method [26]. Using calculations, it is claimed that with increasing n, the brightness of the output image is preserved more efficiently.

$$E(Y) = X_m + \left[\frac{XG - X_m}{2^n}\right] \tag{4}$$

In the above relation XG is the average of gray level and $\mathrm{X_m}$ is the average of efficiency. When the return level n increases E(Y) suddenly converts to an average of efficiency that is obvious from recent equality.

While RMSHE is a recursive method, it also maintains the scalability of image brightness, which is a very important parameter in image processing. The main advantage of the RMSHE method is to improve brightness with a recursive level assigned to a low contrast image.

4 Proposed Algorithm

In the optimal contrast improvement techniques mentioned in this study, histogram of input image is divided to two or more sub-histogram using different methods and then the histogram equalization (HE) method is performed on each of these sub-histograms independently. Evaluation of medical image's contrast improvement techniques, specially on mammography, shows that RMSHE and CLAHE have the best performance on contrast improvement and brightness reservation. Using these methods on MIAS database shows good developments on EME, PSNR, MSE and AMBE parameters. Also, the RMSHE technique brings the best brightness preservation to the images. Using these results leads us to utilize CLAHE in the equalization of sub histograms. Empirical results show significant improvements on contrast restorations.

However in this paper Effective Measure of Enhancement (EME) and Peak Signal to Noise Ratio (PSNR) are used to evaluate the performance of the algorithms. PSNR is a measure of the deviation of the current image from the original image with respect to the peak value of the gray level. The EME is a quantitative measure of image enhancement.

It is obtained by splitting the image into a number of Blocks and using the equation:

$$\mathrm{EME} = \frac{1}{K_1 K_2} \sum_{L=1}^{K_2} \sum_{k=1}^{K_1} 20\mathrm{Log}\left(\frac{\mathrm{I}_{\max}(K,L)}{\mathrm{I}_{\min}(K,L)}\right) \tag{5}$$

In the above equation, K_1 and K_2 are the numbers of horizontal and vertical blocks of the image and $\mathrm{I}_{\max}$ (k, L) and $\mathrm{I}_{\min}$ (k, L) are the maximum and minimum pixel values in a given block, respectively.

Besides EME, in order to improve the confidence of the evaluation results, we use another factor named Absolute Mean Brightness Error (AMBE), which is defined to rate the performance of preserving the original brightness. Smaller values of this parameter are related to the better preservation of image brightness. AMBE is calculated as the absolute difference between original and enhanced images and is given as:

$$AMBE = \left| I(i,j) - \hat{I}(\mathrm{i},\mathrm{j}) \right| \tag{6}$$

In this equation, I(i, j) and $\hat{I}(\mathrm{i},\mathrm{j})$ are average intensity of input and enhanced images, respectively which is defined between 0 and ∞.

Besides these factors, MSE as the Mean Square Error between the original (i.e. s) and the enhanced (i.e. ŝ) images is used as illustrated in Eq. (7):

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{J=1}^{N} \left[I(\mathrm{i},\mathrm{j}) - \hat{I}(\mathrm{i},\mathrm{j}) \right] \tag{7}$$

In the following, the results of the indirect actions of contrast enhancement techniques introduced in this paper, based on the example of mammographic image are displayed (Figs. 2, 3, 4 and 5).

In Tables 1, 2, 3 and 4, the results of the Effective Measure of Enhancement (EME) and peak signal-to-noise ratio (PSNR), mean squared error (MSE) and absolute mean brightness error (AMBE) are presented which have been obtained by applying the indirect contrast enhancement techniques introduced in this paper are based on several examples of mammogram images extracted from the MIAS (Mammography Image Analysis Society) database.

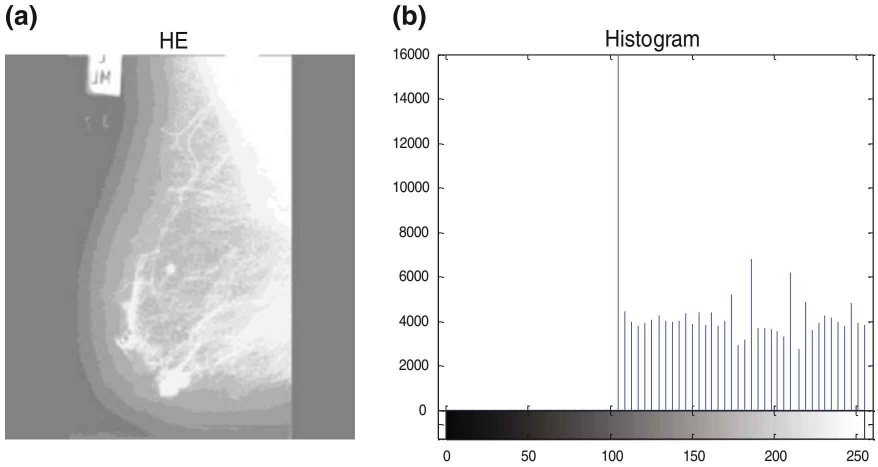


Fig. 2 **a** Contrast enhancement with histogram equalization (HE) technique. **b** Histogram image

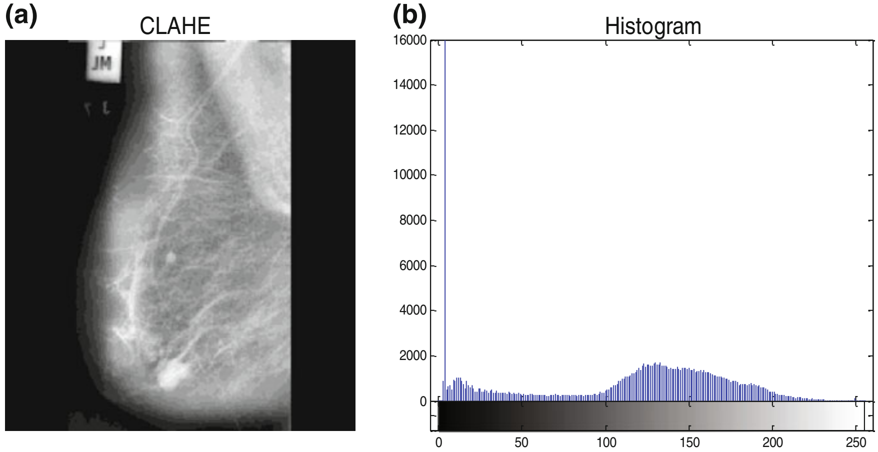


Fig. 3 **a** Contrast enhancement with contrast-limited adaptive histogram equalization (CLAHE) technique. **b** Histogram image

5 Conclusions

In this study, the well-known techniques for improving the image indirect contrast, including HE, CLAHE and RMSHE with their application in low contrast mammographic images were investigated. The traditional HE method significantly changes the image brightness; therefore the details of the image cannot be evidently

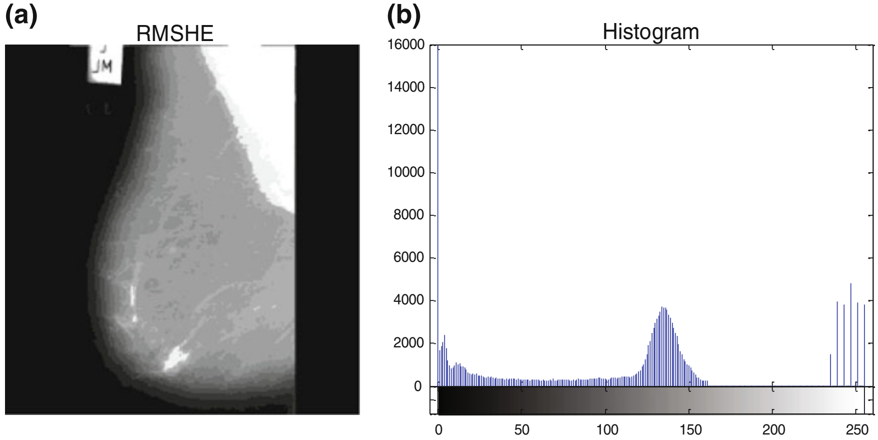


Fig. 4 **a** Contrast enhancement with recursive mean-separate histogram equalization (RMSHE) technique. **b** Histogram image

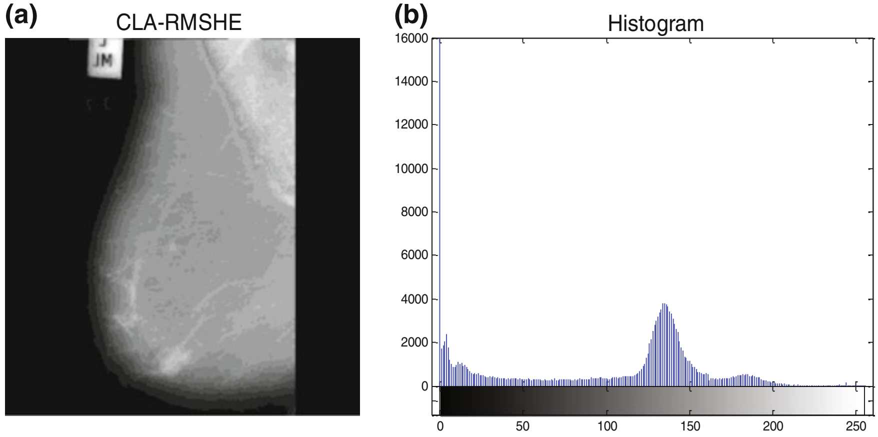


Fig. 5 **a** Contrast enhancement with suggested technique. **b** Histogram image

Table 1 EME values for different contrast enhancement techniques

Image	HE	CLAHE	RMSHE	CLA-RMSHE
mdb009	1.1380	5.1268	7.2849	7.7172
mdb035	0.2818	2.7043	3.8918	4.3630
mdb043	0.2632	2.9431	5.7116	6.1585
mdb057	0.6388	3.3798	4.2820	4.9475
Mdb107	0.5406	3.2968	3.8278	4.5474
Mdb137	0.3705	3.5250	4.5075	5.0512
Mdb145	1.3865	4.6744	6.3660	7.3599
Mdb163	0.5090	3.6073	4.8155	5.2449

Table 2 PSNR values for different contrast enhancement techniques

Image	HE	CLAHE	RMSHE	CLA-RMSHE
mdb009	8.4180	21.2360	18.2825	28.8805
mdb035	4.5880	24.2680	17.7157	29.9343
mdb043	4.2616	25.3344	15.6068	30.7476
mdb057	7.2361	24.5555	19.7264	28.7411
Mdb107	8.4115	21.2360	24.4981	27.7643
Mdb137	6.3530	20.9996	19.8844	24.2115
Mdb145	12.1369	20.9771	34.4241	38.6424
Mdb163	7.6651	22.5691	24.8680	28.9360

Table 3 MSE values for different contrast enhancement techniques

Image	HE	CLAHE	RMSHE	CLA-RMSHE
mdb009	167.3960	74.2731	102.2214	60.1739
mdb035	202.7276	75.7823	105.1597	57.0855
mdb043	206.0635	71.8472	116.8537	54.8107
mdb057	177.5863	74.7005	95.1016	60.5947
Mdb107	167.4504	88.1873	74.9152	63.9247
Mdb137	185.6033	89.2360	94.3532	75.9965
Mdb145	138.9922	89.3364	49.2541	36.9343
Mdb163	173.8180	82.5009	73.5427	63.2929

Table 4 AMBE values for different contrast enhancement techniques

Image	HE	CLAHE	RMSHE	CLA-RMSHE
mdb009	92.8761	6.8134	15.3766	2.4532
mdb035	146.7379	7.4329	12.2321	1.2084
mdb043	153.7188	7.8285	17.2210	0.0738
mdb057	103.9611	4.2711	11.7311	2.7701
Mdb107	88.7582	2.8612	6.8836	4.8524
Mdb137	117.5142	6.0282	11.4282	4.2499
Mdb145	58.1488	4.8354	2.8243	0.4942
Mdb163	96.1106	3.1374	6.9858	3.1002

verified. By comparing the obtained results of several image samples from the MIAS database, two RMSHE and CLAHE techniques perform better in contrast of mammographic images, while the RMSHE technique has the best brightness preservation. Applying the contrast-limited adaptive histogram equalization (CLAHE) to the sub-histograms derived from image decomposition with RMSHE technique, effective improvement results and a better peak signal-to-noise ratio can be achieved for improvement of the image contrast.

References

1. Tang J, Rangayyan RM, Xu J, El Naqa I, Yang Y (2009) Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. IEEE Trans Inf Technol Biomed 13(2): 236–251
2. Akila K, Jayashree LS, Vasuki A (2015) Mammographic image enhancement using indirect contrast enhancement techniques—a comparative study. Procedia Comput Sci 47:255–261
3. Athanasiou A, Aubert E, Vincent Salomon A, Tardivon A (2014) Complex cystic breast masses in ultrasound examination. Diagn Intervent Imaging 95:169–179
4. Shao Y-Z, Liu L-Z, Bie M-J, Li C-c, Wu Y-p, Xie X-m, Li L (2011) Characterizing the clustered microcalcifications on mammograms to predict the pathological classification and grading: a mathematical modeling approach, Published online: 22 April 2011 Society for Imaging Informatics in Medicine
5. Popli MB, Teotia R, Narang M, Krishna H (2014) Breast positioning during mammography: mistakes to be avoided. Breast Cancer: Basic Clin Res 8
6. Zhou Y, Panetta K, Agaian S (2010) Human visual system based mammogram enhancement and analysis. In: 2010 2nd international conference on image processing theory tools and applications (IPTA). IEEE, pp 229–234
7. Suhail Z, Denton ERE, Zwiggelaar R (2017) Tree-based modelling for the classification of mammographic benign and malignant micro-calcification clusters. Multimed Tools Appl. Published with open access at Springer
8. Bilous M (2010) Breast core needle biopsy: issues and controversies. Mod Pathol 23:S36–S45
9. Sivaramakrishna R, Obuchowski NA, Chilcote WA, Cardenosa G, Powell KA (2000) Comparing the performance of mammographic enhancement algorithms: a preference study. Am J Roentgenol 175(1):45–51
10. Polesel A, Ramponi G, Mathews V (2000) Image enhancement via adaptive unsharp masking. IEEE Trans Image Process 9(3):505–510
11. Cheng H-D, Xu H (2000) A novel fuzzy logic approach to contrast enhancement. Pattern Recogn 33(5):809–819
12. Agaian S, Panetta K, Grigoryan A (2001) Transform based image enhancement algorithms with performance measure. IEEE Trans Image Process 10(3):367–382
13. Tang J, Peli E, Acton S (2003) Image enhancement using a contrast measure in the compressed domain. IEEE Sig Process Lett 10(10):289–292
14. Kim YT (1997) Contrast enhancement using brightness preserving bi-histogram equalization. IEEE Trans Consum Electron 43(1):1–8
15. Wang Q, Ward RK (2007) Fast image/video contrast enhancement based on weighted thresholded histogram equalization. IEEE Trans Consum Electron 53(2):757–764
16. Hashemi S, Kiani S, Noroozi N, Moghaddam ME (2010) Contrast enhancement method based on genetic algorithm. Pattern Recogn Lett 31:1816–1824
17. Gonzalez RC, Woods RE (2002) Digital image processing, 2nd edn. Prentice- Hall, Englewood Cliffs
18. Pisano ED, Cole EB, Hemminger BM, Yaffe M, Aylward SR, Maidment ADA, Eugene Johnston R et al (2000) Image processing algorithms for digital mammography: a pictorial essay 1. Radiographics 20(5):1479–1491
19. Al-Ameen Z, Sulong G, Rehman A, Al-Dhelaan A, Saba T (2015) An innovative technique for contrast enhancement of computed tomography images using normalized gamma-corrected contrast-limited adaptive histogram equalization. EURASIP J Adv Sig Process 2015:32. https://doi.org/10.1186/s13634-015-0214-1
20. Pisano ED, Zong S, Hemminger BM, DeLuca M, Johnston RE, Muller K, Patricia Braeuning M, Pizer SM (1998) Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. J Digit Imaging 11(4): 193–200

21. Gupta P, Kumare JS, Singh UP, Singh RK (2017) Histogram based image enhancement techniques: a survey. Int J Comput Sci Eng 5(6). E-ISSN: 2347-2693
22. Jayaraman S, Esakkirajan S, Veerakumar T (2015) Digital image processing. Tata McGraw-Hill Education Pvt. Ltd
23. Kim JY, Kim L, Hwang S (2001) An advanced contrast enhancement using partially overlapped sub-block histogram equalization. IEEE Trans Circ Syst Video Technol: 475–484
24. Haddadnia J, Seryasat OR, Ghayoumi-Zadeh H, Rabiee H (2015) An efficient method for detection of masses in mammogram images. Cumhuriyet Sci J 36(3):2269–2277
25. Seryasat OR, Haddadnia J, Ghayoumi Zadeh H (2016) Assessment of a novel computer aided mass diagnosis system in mammograms. Iran J Breast Dis 9(3):31–41
26. Chen S-D, Ramli AR (2003) Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation. IEEE Trans Consum Electron 49(4): 1301–1309

Contrast Improvement of Ultrasound Images of Focal Liver Lesions Using a New Histogram Equalization

Younes Mousania and Salman Karimi

Abstract Contrast improvement is an important issue in the processing of medical images. Due to the difficulty of detecting liver lesions in conventional ultrasound imaging and the low contrast of these images, we tried to provide an indirect optimization technique on the ultrasound images of Focal Liver Lesions database in the space of two-dimensional histogram to improve the quality and the contrast of these images. To prevent undesirable effects due to the adjustment of the histogram images, two techniques are used: CLAHE and RMSHE. By using four effective measurement techniques metrics of EME, PSNR, MSE and AMBE, shows that the proposed method has significant consequences. Furthermore; results of the study revealed that improved outcomes are obtained when the proposed technique is utilized on other standard ultrasound and medical images like mammography.

Keywords Ultrasound image · Contrast enhancement · Histogram equalization CLAHE · RMSHE · EME · PSNR · MSE · AMBE

1 Introduction

The liver is the largest gland of the body with an important role in metabolism and digestion. In this study, focal hepatic lesions and especially hepatic cysts have been investigated with the aim of improving ultrasound images of these lesions. A wide range of liver lesions is presented in differential diagnosis, but in general, these lesions can be classified into two categories of benign and malignant [1]. In order to determine the nature of these masses, one has to consider issues such as age, sex, and the presence of chronic liver disease in the patient [2].

Y. Mousania (✉) · S. Karimi
Department of Electrical and Electronic Engineering,
Lorestan University, Khorram-abad, Lorestan, Iran
e-mail: Mousania.yo@fe.lu.ac.ir

S. Karimi
e-mail: Karimi.salman@lu.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_4

Liver lesions are accidentally detected and are mostly benign. Benign types originate from the liver tissue, but the malignant type or liver cancers can be of different origins [3]. In addition to pain and swelling in the upper quadrant of the abdomen which is also seen in benign lesions, malignant liver lesions can also cause jaundice, bloody ascites, appetite and weight loss. However, Benign masses, sometimes grow so much that they can cause problems [4]. But in most cases, they do not spread to adjacent tissues and usually do not require treatment, unless the patient has symptoms, in which case removal of the tumor with surgery will improve the symptoms. Although the exact diagnosis of the nature of a liver mass is achieved by sampling and pathologic examination, liver masses can be diagnosed by various diagnostic methods such as clinical, pathological and ultrasonographic methods [5]. Ultrasound imaging has many advantages, the most important of which can be non-invasiveness and the use of non-ionizing radiation, which has led to its wide usage in the diagnosis of various diseases [6].

Benign tumors are usually isolated, but sometimes they may be numerous, such as in the case of liver cysts and multiple liver abscesses. In general, cysts are thin-walled structures that contain liquid. Polycystic liver disease is associated with polycystic kidney disease in half of the cases [7]. Few patients bleed into the cyst, the phenomenon which causes pain in the upper extremity and sudden and severe shoulder pain. Bleeding stops without intervention and the pain declines after a few days. Liver cysts do not interfere with liver function. Cysts are usually found by ultrasound or computerized tomography (CT scan), and the simple type is always benign. Only those who experience symptoms are in need of treatment. Removing the fluid from the cyst with a needle simply is not enough because the cyst will be filled in again within a few days. The best and the easiest treatment is to remove a large part of the cyst's wall. This surgery can usually be done through laparoscopy and has a therapeutic effect in almost all patients [8].

Digital image processing techniques are used to raise the quality level of ultrasound images as well as to increase the accuracy of diagnosis of liver lesions. Image contrast improvement is one of the most important requirements used in image processing and vision system applications. In general, methods of the contrast improvement are divided into two major categories: direct methods and indirect methods [9].

2 Optimal Methods of Direct Contrast

In the direct methods, while defining a criterion for measuring the image contrast, attempts are made to improve image contrast by improving this criterion. Creating an appropriate measurement criterion for image contrast is an important stage to improve the image directly. The direct contrast approach considers both the general and local information of the image, hence it can be improved in many applications.

In this regard various approaches have been proposed that are based on the phase entropy principle, which transmits the image to the phase domain, and the phase entropy is calculated, and this way the local contrast is measured [10].

3 Optimal Methods of Indirect Contrast

Improving contrast with the indirect method involves modifying the histogram of the image. In indirect method, the dynamic range of the gray levels of the image are increased to improve contrast. Indirect methods which have been paid more attention in recent years due to direct and knowledge-based representation, are categorized into four categories:

- Methods that modify the up and down frequency components of the image [11]
- Methods based on Conversion [11, 12]
- Methods based on histogram modification [13, 14]
- Methods based on Soft calculation [15].

The proposed algorithm and techniques presented in this paper are based on histogram correction methods. In Fig. 1, an ultrasound image of Focal Liver Lesions with its histogram is displayed.

3.1 Histogram Equalization (HE)

The main idea of HE is mapping of the values of the input image intensity to the new intensity values through a transformation function created for the cumulative

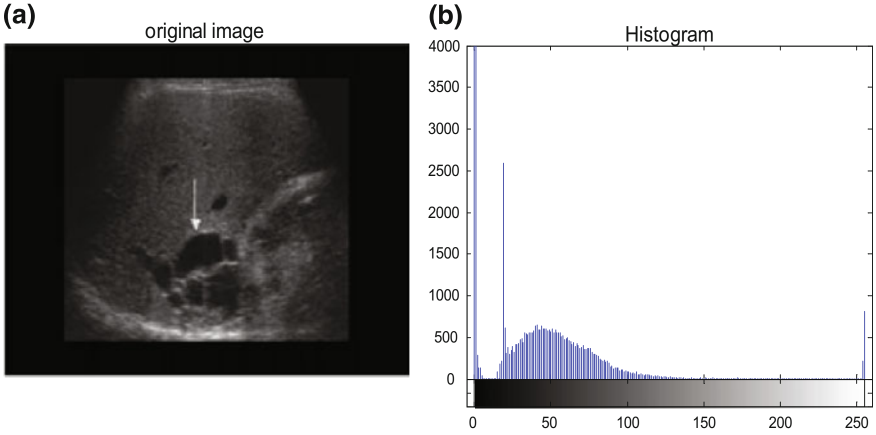


Fig. 1 **a** Original ultrasound image. **b** Histogram image

distribution function (CDF). First, HE converts the histogram of the original image to a plane histogram using an average value that is the average range of gray level [16]. Therefore, the histogram of the image is divided into two parts based on its average gray level, and the HE algorithm is separately applied on each divided section of the histogram. Secondly, histogram equalization performs the improvement action based on the overall content of the image.

HE is powerful in highlighting the boundaries and edges between different objects, but it may change the local details in these objects, particularly smooth and small areas. The other problem of HE is an abnormal increase and saturation effects of intensity and also it is not appropriate to maintain the brightness of the original image due to the changes in the brightness of the image [17].

3.2 Contrast-Limited Adaptive Histogram Equalization (CLAHE)

CLAHE is a kind of adaptive equalization of the histogram. This method divides the original image into several sub-images without overlapping [18]. The secondary histogram of the images is limited to the value of the improvement per each pixel and then equalization is performed. Details of the image are evidently revealed with respect to the background [19]. At the same time, the contrast of the image is improved equally, which results in an output contrast image with high quality [20]. In this paper, using an adaptive filtering procedure, the histogram of different parts of the partitioned image is calculated and then the histogram balancing is utilized to rearrange the brightness values of the total image. So our proposed method is different from the smoothing of the fundamental histogram, since in this method, as a traditional equation technique, only one histogram is used for the whole image [21].

Consequently, for the purpose of improving the localized image contrast and extracting more details from the image, while significant noise would be generated, the contrasting histogram is equalized.

In order to suppress these deficiencies, a generalization of Adaptive Histogram Equalization (AHE) of a contrast-limited, or concise, which is called CLAHE, is used.

This technique is designed to overcome the problem of noise exacerbation. CLAHE does not deal with the entire image, but deals with pieces that are in small areas of the image [22]. The contrast of each area is improved in such a way that the histogram of the output region corresponds to approximately the histogram expressed by the distribution parameter.

Neighbor sections are combined to eliminate abnormal induced boundaries by using bidirectional interpolations [22]. Utilizing contrast in homogeneous regions, it is possible to avoid any exacerbation of any unwanted noise that may be present in the low contrast image. Besides user friendly, simple calculation and good output in local areas are of the advantages of CLAHE. Additionally, CLAHE has less noise and can maintain the light saturation which normally occurs in the histogram equalization procedures [23, 24].

3.3 Recursive Mean-Separate Histogram Equalization (RMSHE)

One of the first suggestions to overcome the drawbacks of the HE method is to preserve the brightness of the equalized bi-histogram (BBHE). This method preserves the effective amount of image brightness while improving the contrast. Moreover, it divides the histogram into two sub-histograms based on the average amount of the brightness and equalizes each part individually if X_m denote the mean of the image X and assume that $X_m \in \{X_0, X_1, \ldots, X_{L-1}\}$. Based on the mean X_m the input image is divided into two sub level images X_L and X_U. The transform functions for the sub images are defined as:

$$F_L(X) = X_0 + (X_m - X_0)C_L(X) \tag{1}$$

$$F_u(X) = X_{m+1} + (X_{L-1} - X_{m+1})C_u(X) \tag{2}$$

According to the above equations, $C_L(X)$ and $C_U(X)$ is the respective cumulative density functions for X_L and X_U.

The output image (Y) of BBHE, is expressed as

$$Y = F_L(X_L) \cup F_u(X_u) \tag{3}$$

Now we introduce a better technique called Recursive Mean-Separate Histogram Equalization (RMSHE), which in fact performs the same BBHE algorithm as a recursive one. In aforementioned techniques the input image histograms were divided into two parts. However, in this method, instead of dividing the input image one time, the input image divides to 2^n sub-histograms using an optional criterion called n. Then, each of these sub-histograms is equalized in dependently.

When $n = 0$, it means that no sub-image is created, which is the same as the HE method [23]. Using calculations, it is claimed that with increasing n, the brightness of the output image is preserved more efficiently.

$$E(Y) = X_m + [XG - X_m/2^n] \tag{4}$$

In the above relation XG is the average of gray level and X_m is the average of efficiency. When the return level n increases E(Y) suddenly converts to an average of efficiency that is obvious from recent equality.

While RMSHE is a recursive method, it also maintains the scalability of image brightness, which is a very important parameter in image processing. The main advantage of the RMSHE method is to improve brightness with a recursive level assigned to a low contrast image.

4 Proposed Algorithm

In the optimal contrast improvement techniques mentioned in this study, histogram of input image is divided to two or more sub-histogram using different methods and then the histogram equalization (HE) method is performed on each of these sub-histograms independently. Evaluation of medical image's contrast improvement techniques, especially on mammography, shows that RMSHE and CLAHE have the best performance on contrast improvement and brightness reservation. Using these methods on MIAS database shows good developments on EME, PSNR, MSE and AMBE parameters. Also, the RMSHE technique brings the best brightness preservation to the images. Using these results leads us to utilize CLAHE in the equalization of sub histograms. Empirical results show significant improvements on contrast restorations.

However, in this paper Effective Measure of Enhancement (EME) and Peak Signal to Noise Ratio (PSNR) are used to evaluate the performance of the algorithms. PSNR is a measure of the deviation of the current image from the original image with respect to the peak value of the gray level. The EME is a quantitative measure of image enhancement. It is obtained by splitting the image into a number of blocks and using the equation:

$$\mathrm{EME} = \frac{1}{K_1 K_2} \sum_{L=1}^{K_2} \sum_{k=1}^{K_1} 20\mathrm{Log}\left(\frac{\mathrm{I}_{\max}(K,L)}{\mathrm{I}_{\min}(K,L)}\right) \tag{5}$$

In the above equation, K_1 and K_2 are the numbers of horizontal and vertical blocks of the image and I_{max} (k, L) and I_{min} (k, L) are the maximum and minimum pixel values in a given block, respectively.

Besides EME, in order to improve the confidence of the evaluation results, we use another factor named Absolute Mean Brightness Error (AMBE), which is defined to rate the performance of preserving the original brightness. Smaller values of this parameter are related to the better preservation of image brightness. AMBE is calculated as the absolute difference between original and enhanced images and is given as:

$$AMBE = \left|I(\mathrm{i,j}) - \hat{\mathrm{I}}(\mathrm{i,j})\right| \tag{6}$$

In this equation, I(i, j) and $\hat{\mathrm{I}}(\mathrm{i,j})$ are average intensity of input and enhanced images, respectively which is defined between 0 and ∞.

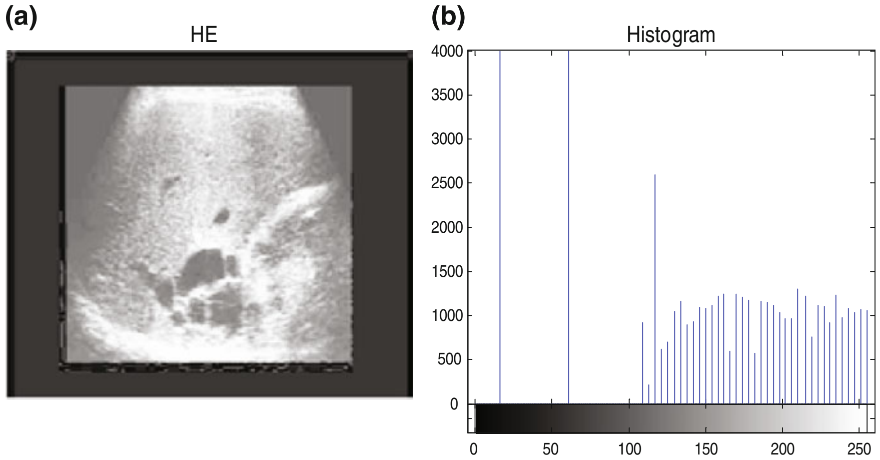


Fig. 2 **a** Contrast enhancement with histogram equalization (HE) technique. **b** Histogram image

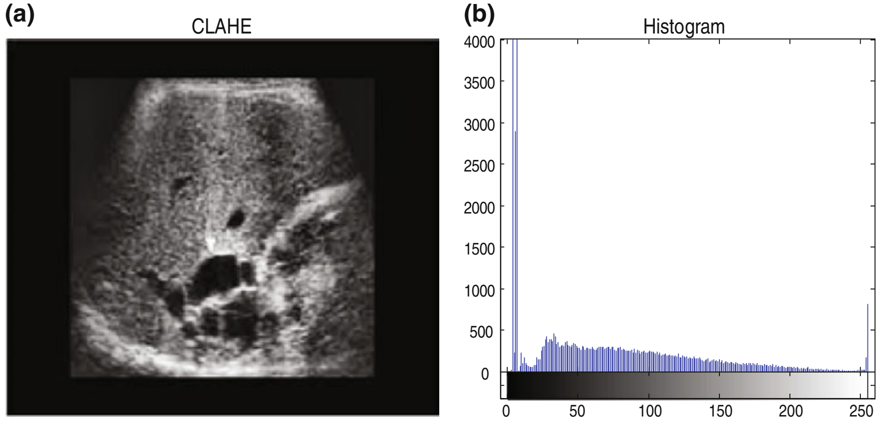


Fig. 3 **a** Contrast enhancement with contrast-limited adaptive histogram equalization (CLAHE) technique. **b** Histogram image

Besides these factors, MSE as the Mean Square Error between the original (i.e. s) and the enhanced (i.e. $\hat{s}$) images is used as illustrated in Eq. (7) (Figs. 2, 3, 4 and 5).

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{J=1}^{N} \left[\mathrm{I}(\mathrm{i},\mathrm{j}) - \hat{\mathrm{I}}(\mathrm{i},\mathrm{j}) \right] \tag{7}$$

In Tables 1, 2, 3 and 4, the results of the Effective Measure of Enhancement (EME), peak signal-to-noise ratio (PSNR), mean squared error (MSE) and absolute

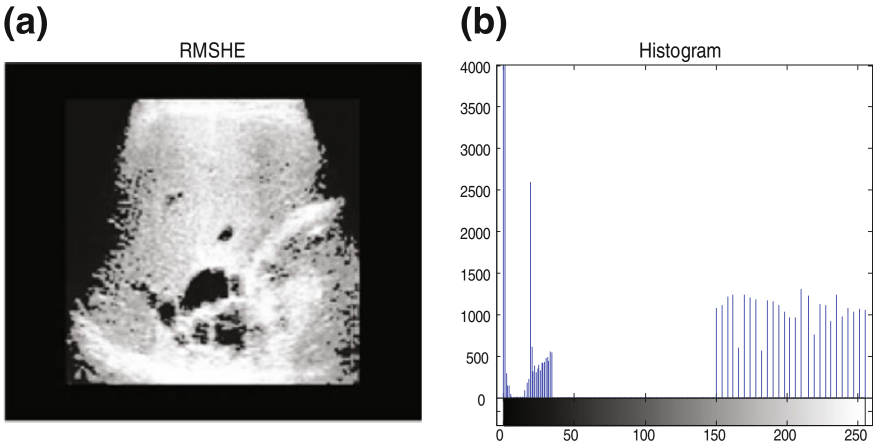


Fig. 4 **a** Contrast enhancement with recursive mean-separate histogram equalization (RMSHE) technique. **b** Histogram image

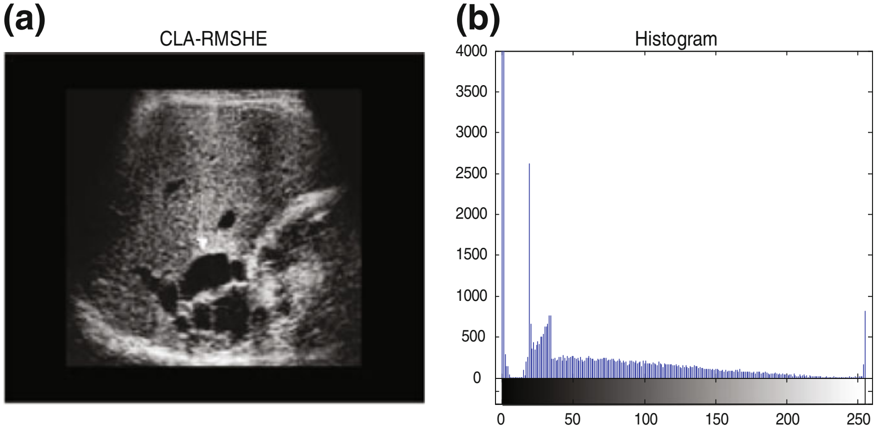


Fig. 5 **a** Contrast enhancement with suggested technique. **b** Histogram image

Table 1 EME values for different contrast enhancement techniques

Image	HE	CLAHE	RMSHE	CLA-RMSHE
Liver cysts 1	2.2486	4.3493	4.6860	6.0966
Liver cysts 2	2.6547	4.0582	5.0975	6.5524
Liver cysts 3	2.3754	3.4284	4.9945	5.4786
Liver cysts 4	1.2882	2.7928	3.4345	4.0855
Liver cysts 5	1.8542	3.4226	4.1450	4.8105
Liver cysts 6	2.2862	4.4173	4.7933	6.2311
Liver cysts 7	2.3452	3.5722	5.2153	6.3751
Liver cysts 8	1.0402	2.6869	3.9929	4.3214

Table 2 PSNR values for different contrast enhancement techniques

Image	HE	CLAHE	RMSHE	CLA-RMSHE
Liver cysts 1	9.5889	16.8176	10.6595	19.8243
Liver cysts 2	7.5803	15.8577	9.0458	18.8056
Liver cysts 3	8.1229	16.4563	9.5785	18.4956
Liver cysts 4	7.8551	17.5952	9.2693	19.4750
Liver cysts 5	8.8437	17.3986	10.1156	20.1406
Liver cysts 6	9.6820	16.7284	10.9549	19.9869
Liver cysts 7	10.5727	16.0937	12.5058	21.7694
Liver cysts 8	7.5767	17.8325	9.2233	21.0033

Table 3 MSE values for different contrast enhancement techniques

Image	HE	CLAHE	RMSHE	CLA-RMSHE
Liver cysts 1	157.8770	109.9894	149.6487	94.6371
Liver cysts 2	174.5568	115.3971	162.2234	99.5821
Liver cysts 3	169.8841	111.9944	157.9591	101.1378
Liver cysts 4	172.1744	105.7949	160.4202	96.3044
Liver cysts 5	163.8707	106.8400	153.7740	93.1521
Liver cysts 6	157.1438	110.4812	147.4545	93.8707
Liver cysts 7	150.2990	114.0433	136.4524	85.8664
Liver cysts 8	174.5882	104.5474	160.7901	89.2194

Table 4 AMBE values for different contrast enhancement techniques

Image	HE	CLAHE	RMSHE	CLA-RMSHE
Liver cysts 1	76.9183	24.6895	50.1574	11.5892
Liver cysts 2	96.1862	28.0400	56.5168	14.3583
Liver cysts 3	91.5083	24.7884	55.0843	15.1228
Liver cysts 4	94.2108	22.4383	55.7983	13.2048
Liver cysts 5	83.0150	21.8273	50.4046	11.4332
Liver cysts 6	77.4449	24.2536	48.7779	10.9278
Liver cysts 7	71.1864	25.6767	39.8442	6.2524
Liver cysts 8	97.3215	23.4840	53.7510	10.2786

mean brightness error (AMBE) are presented which have been obtained by applying the indirect contrast enhancement techniques introduced in this paper and are based on several examples of *Ultrasound Images of Focal Liver Lesions* images extracted from the Ultrasound cases database.

5 Conclusions

In this study, the well-known techniques for improving the image indirect contrast, including HE, CLAHE and RMSHE with their application in low contrast mammographic images were investigated. The traditional HE method significantly changes the image brightness; therefore the details of the image cannot be evidently verified. By comparing the obtained results of several image samples from the MIAS and Ultrasound cases database, two RMSHE and CLAHE techniques perform better in contrast of mammographic and Ultrasound images, while the RMSHE technique has the best brightness preservation. Applying the contrast-limited adaptive histogram equalization (CLAHE) to the sub-histograms derived from image decomposition with RMSHE technique, effective improvement results and a better peak signal-to-noise ratio can be achieved for improvement of the image contrast.

References

1. Hasan NMA, Alam Eldeen MH (2016) Benign versus malignant focal liver lesions: diagnostic value of qualitative and quantitative diffusion weighted MR imaging. Egypt J Radiol Nucl Med 47(4):1211–1220
2. El-Kader SMA, El-Den Ashmawy EMS (2015) Non-alcoholic fatty liver disease: the diagnosis and management. World J Hepatol 7(6):846–858
3. Winterer JT, Kotter E, Ghanem N, Langer M (2006) Detection and characterization of benign focal liver lesions with multislice CT. Eur Radiol 16(11):2427–2443
4. Chen J, Du Y-J, Song J-T (2010) Primary malignant liver mesenchymal tumor: a case report. World J Hepatol 16(41):5263–5266
5. Hapani H, Kalola J, Trivedi A, Chawla A (2014) Ultrasound evaluation of focal hepatic lesions. IOSR J Dent Med Sci 13(12):40–45 (Ver. IV)
6. Miller D, Smith N, Bailey M, Czarnota G, Hynynen K, Makin I (2013) American, overview of therapeutic ultrasound applications and safety considerations. US National Library of Medicine National Institutes of Health, available in PMC 28 Oct 2013
7. Cnossen WR, Drenth JPH (2014) Polycystic liver disease: an overview of pathogenesis, clinical manifestations and management. US National Library of Medicine National Institutes of Health, available in PMC 9: 69
8. Tuxun T, Zhang J-h, Zhao J-m, Tai Q-w, Abudurexti M, Ma H-Z, Wen H (2014) World review of laparoscopic treatment of liver cystic echinococcosis—914 patients. Int J Infect Dis 24:43–50
9. Polesel A, Ramponi G, Mathews V (2000) Image enhancement via adaptive unsharp masking. IEEE Trans Image Process 9(3):505–510
10. Cheng H-D, Xu H (2000) A novel fuzzy logic approach to contrast enhancement. Pattern Recogn 33(5):809–819
11. Agaian S, Panetta K, Grigoryan A (2001) Transform based image enhancement algorithms with performance measure. IEEE Trans Image Process 10(3):367–382
12. Tang J, Peli E, Acton S (2003) Image enhancement using a contrast measure in the compressed domain. IEEE Sig Process Lett 10(10):289–292
13. Kim YT (1997) Contrast enhancement using brightness preserving bi-histogram equalization. IEEE Trans Consum Electron 43(1):1–8

14. Wang Q, Ward RK (2007) Fast image/video contrast enhancement based on weighted thresholded histogram equalization. IEEE Trans Consum Electron 53(2):757–764
15. Hashemi S, Kiani S, Noroozi N, Moghaddam ME (2010) Image, contrast enhancement method based on genetic algorithm. Pattern Recogn Lett 31:1816–1824
16. Gonzalez RC, Woods RE (2002) Digital image processing, 2nd edn. Prentice- Hall, Englewood Cliffs
17. Pisano ED, Cole EB, Hemminger BM, Yaffe MJ, Aylward SR, Maidment ADA, Johnston RE et al (2000) Image processing algorithms for digital mammography: a pictorial essay 1. Radiographics 20(5):1479–1491
18. Al-Ameen Z, Sulong G, Rehman A, Al-Dhelaan A, Saba T (2015) An innovative technique for contrast enhancement of computed tomography images using normalized gamma-corrected contrast-limited adaptive histogram equalization. EURASIP J Adv Sig Process:32. https://doi.org/10.1186/s13634-015-0214-1
19. Pisano ED, Zong S, Hemminger BM, DeLuca M, Johnston RE, Muller K, Braeuning MP, Pizer SM (1998) Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. J Dig Imaging 11(4): 193–200
20. Sivaramakrishna R, Obuchowski NA, Chilcote WA, Cardenosa G, Powell KA (2000) Comparing the performance of mammographic enhancement algorithms: a preference study. Am J Roentgenol 175(1):45–51
21. Gupta P, Kumare JS, Singh UP, Singh RK (2017) Histogram based image enhancement techniques: a survey. Int J Comput Sci Eng 5(6) E-ISSN: 2347-2693
22. Jayaraman S, Esakkirajan S, Veerakumar T (2015) Digital image processing. Tata McGraw-Hill Education Pvt. Ltd
23. Kim JY, Kim L, Hwang S (2001) An advanced contrast enhancement using partially overlapped sub-block histogram equalization. IEEE Trans Circ Syst Video Technol 475–484
24. Chen S-D, Ramli AR (2003) Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation. IEEE Trans Consum Electron 49(4): 1301–1309

An Unequal Clustering-Based Topology Control Algorithm in Wireless Sensor Networks Using Learning Automata

Elahe Nouri

Abstract Clustering is an efficient method in saving nodes' consumption of energy within wireless sensor networks. In the majority of clustering methods, the sizes of clusters are equal to each other. This feature will lead to a rise in consumption of energy in clusters which are close to the sink node. For this purpose, a new method has been presented based on the learning automata for the unequal clustering of nodes in wireless sensor network. In this method, upon reduction of the distance between the clusters and the sink hole, the size of clusters also shrinks. This feature in addition to optimization of consumption of energy in the network, will also be efficient in a rise in the number of packages delivered toward the sink hole. The effectiveness of the proposed method has been assessed based on its implementation in a simulated environment, while the corresponding results have been compared with previous approaches. The results show that the proposed method acts better than previous methods in reduction of consumption of energy and a rise in the number of delivered packages.

Keywords Wireless sensor network · Clustering · Learning automata

1 Introduction

The sensor networks are comprised of a large number of sensors with limited energy and calculation sources. Each sensor maintains the ability to receive especial data such as temperature, light, sound, movement, and so forth, and can deliver its

E. Nouri (✉)
Department of Computer Engineering, Faculty of Engineering,
Islamic Azad University, Arak Branch, Arak, Markazi, Iran
e-mail: elahe_ni2009@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_5

received data from the environment to its neighbors. In the wireless sensor network this communication and contact is established via radio waves. In the wireless sensor networks, the sensed data are sent toward the sink node via a node in direct manner and/or via other nodes. Given the limitation of sources of energy of sensor nodes, the efficiency of consumption of energy in these networks is of paramount importance. Clustering is one of the appropriate approaches for optimization of consumption of energy in wireless sensor networks. In a sensor network, whose structure is based on clustering, the sensor nodes are in the form of categorized clusters, with each node sending its data via the cluster head toward the sink node. In data routing toward the sink node, in a sensor network which is based on clustering, the cluster heads close to the sink node, in addition to delivery of the data of the members of their cluster, are also duty-bound for routing the delivered data from cluster heads, further apart from the sink hole. This in turn results in a surge in consumption of energy by the cluster head nodes, close to the sink node. One method for prevention of this process is to reduce the size of clusters close to the sink hole and to increase the size of clusters which are further apart. This approach is referred to as unequal clustering. The application of unequal clustering technique leads to balanced distribution of load across the network and optimization of the energy consumption parameters, in addition to improvement of the rate of delivery of package. One of the important topics in regard to unequal clustering algorithms is the determination of the cluster head and the size of each cluster. Selection of optimal values for these parameters can shape an optimal topology, and lead to the improvement of the fundamental parameters of the wireless sensor network. The technique of reinforcement learning can be employed as an efficient approach for resolution of this issue. In the reinforcement learning, the learner, throughout the learning process, and repetitive interaction with the environment, attains an optimal control policy. The effectiveness of these interactions with the environment is evaluated via the maximum (minimum) numerical reward (fine) which is received from the environment. The main advantage of reinforcement learning in comparison to other learning methods, is the lack of need for any data from the environment in this approach. One of the methods of reinforcement learning, is the random learning automata. The random automata, in the absence of any data on the optimal act, intends to find an answer to the related issue. One act of automata is chosen randomly and is implemented in the environment. Thereafter, the response of the environment is received and the possible actions are updated based on the learning algorithm, with the said procedure being repeated. Based on this approach, an algorithm can be presented for controlling the topology based on unequal clustering in wireless sensor networks. In this study we intend to make use of learning automata in order to determine the cluster head and the size of sensor network clusters, and to present a new algorithm for the purpose of controlling topology based on unequal clustering in wireless sensor networks. The purpose of this study is to optimize parameters of consumption of energy, rate of delivery of package, and rear-to-rear delay in wireless sensor networks [1–5].

2 Methodology

The constant changes of topology in wireless sensor networks are the consequences of impairment and potential stimulation of host nodes. Moreover, the unsustainable presence of nodes in the said networks, alongside restriction of the band width of wireless channels, has caused many complexities and difficulties in regard to controlling topology within these networks. In order to design efficient algorithms for their implementation in wireless sensor networks; topics such as the band width of wireless connections; stimulation of nodes and low energy of the network; which are some of the main features of mobile calculations, should be attended. The stimulation and impairment of the nodes within the corresponding wireless networks lead to continuous changes in the topology of the network and surges the volatility of the network's data. The goal of this thesis is to design an efficient algorithm in a bid to solve the issue of control of topology in wireless sensor networks, so that despite the restrictions of the said networks; by reliance upon the features and capacities of learning automata and their combinations, an appropriate analytical and applied solution would be presented for this topic of importance [6, 7].

The impairment and stimulation of the nodes in wireless sensor networks is one of the main reasons behind constant topological changes across the network, which in turn escalates the volatility of data. On this basis, it is necessary to continuously update the data of the network; as a result of which a significant control is exerted on the network. On the other hand, given the restriction of the band width of wireless connections in these networks, in comparison to wired networks; in addition to restriction of sources of energy in these networks, the models and patterns which are common in wired networks are inappropriate for usage in corresponding mobile network environments that maintain a much higher level of dynamism. During the recent years, different algorithms have been designed for development and control of topology in wireless sensor networks; the majority of which are based on unrealistic common assumptions, making the theoretical analysis of the algorithms feasible. However, on the other hand, their implementation practically leads to many difficulties and challenges with due regard to the restrictions and potentials of the said networks. The development of the network's backbone is one of the most applicable and inspirational approaches for establishment and control of wireless sensor networks. Development of the spanning tree of the network, formation of a maximal independent set for the network; development of dominating set from the network's graph, and the clustering technique of the network are part of the common and most applicable methods for development of the network's backbone. In this thesis, a topology control algorithm is presented based on clustering for wireless sensor networks.

In the topology control algorithm proposed in this thesis, which is hereinafter abbreviated as CTCM (Topology control mechanism based on cluster); the network's topology is shaped based on establishment of the network's backbone. In order to build the topology of the network, in the proposed approach; initially the network is completely clustered, while later on the network's topology is shaped with the connection of cluster head nodes. On this basis, it can be said that the process of control of topology in the proposed algorithm takes place in two phases. In the first phase, the network's clustering is carried out, and in the second phase, the development of the backbone via the connection of cluster head nodes takes place [8].

Within the first phase of the proposed algorithm, upon the usage of learning automata, efforts are made for clustering of inner-network nodes in a manner that the said clusters (or cluster heads) would maintain the maximum energy, and thereby maintain the longest lifetime. For this purpose, efforts are made in each section to select a network of nodes, which maintains the highest remaining energy compared to its other adjacent nodes, as the cluster head node for that section. The selected node plays the role of cluster head for the other nodes within that section, as long as its average remaining energy is higher than the average energy of the nodes in that section. In this method, each node is equipped with a learning automata, within which each action of the automata is described as selection of that node and/or one of its adjacent neighbors as the cluster head node for that section. Throughout a learning repetitive process, each node with the assistance of its automata, selects the best and most energetic adjacent node as the cluster head node for itself. All nodes in a section will reach a joint decision on the selection of the cluster head node. On the other hand, given that the cluster head node maintains the duty to respond to all requests for transference of data, allocation of a channel and so forth from the other nodes of their cluster; it loses it energy at a faster pace compared to other nodes, as the result of which it loses its priority on playing the role of a cluster head. Hence, throughout consecutive intervals, the cluster head node should be re-selected and be once again replaced as the most energetic node in comparison to other nodes in that section. In this method, upon the application of the repetitive learning process in the automata, the permanent selection of the cluster head takes place until the end of the network's lifetime [9].

In the second phase of the algorithm, the topology of the network should be carried out via the connection of the cluster head nodes. To this end, each cluster head node, upon sending a message, identified and connects with its adjacent cluster head nodes. In the end, the most resistant network topology is established via the formation of the most energetic backbone of the network and upon the connection of the most energetic cluster head nodes. Furthermore, prior to presentation of the proposed topology control method, and with due regard to the features, specifications and strong points of learning automata; the reasons behind usage of this method in this thesis as an efficient approach for clustering wireless sensor networks, in accordance to an energy-based method, are detailed [10, 11].

As you know, in order to use learning automata; the fundamental features of learning automata, and nature of its application should be concurrently taken into consideration. The learning automata, in accordance to following features, have been applied as powerful tools for resolution of many problems.

1. Learning automata perform appropriately while there is no data at hand.
2. Learning automata perform appropriately when there is lack of assurance.
3. Learning automata search in the possibility atmosphere.
4. Learning automata need a simple feedback from the environment in order to improve their situation in each phase.
5. Learning automata are highly useful as a model for learning in distributed and multifactor environments, with restricted communications and inadequate data.
6. Learning automata maintain a simple structure and therefore can be easily placed in a software or hardware.
7. Learning automata need to use a theoretical and derived efficiency criterion for optimization applications.
8. Learning automata maintain a negligible computational load and can thus can be simply used for immediate functions.
9. There are powerful mathematical analytical methods for analysis of learning automata.

In addition to above features, usage of learning automata is beneficial for applications which maintain one or few of the following traits. Meanwhile, given the innate features of the wireless sensor networks, as a completely distributed environment with a high degree of dynamism and uncertainty, usage of learning automata would be highly effective.

(1) The application is sufficiently complicated and uncertain, such that there would be no mathematical approach available for it
(2) The application would be able to control distribution and create role models via a series of self-autonomous factors
(3) The reinforcement signal would be a random variable and would be produced based on an efficiency criterion
(4) Quantitative improvement in efficiency would be highly economically justified
(5) There would be no certain algorithm in regard to the considered application [12, 13]

Meanwhile, learning automata also maintain a number of restrictions; which makes their usage for many applications, as problematic. These restrictions are as follows:

(1) Learning automata make use of a few initial data and the additional data about the environment cannot be always used by learning automata
(2) Learning automata maintain a low rate of convergence for many functions
(3) Learning automata are non-reminiscent models

Given the above, and the features, specifications, restrictions, and capacities of wireless sensor networks, which we will mention later on; the usage of learning automata in this thesis for resolution of topology control issue based on clustering in the wireless sensor network can be justified [14].

As previously mentioned, in the wireless sensor networks, the continuous changes in the network's topology are due to emergence of malfunction of nodes;

limitation of the band width of wireless connections; restricted energy of nodes; interferences of channel; noisy connections, and a number of other factors; resulting in uncertainty, unpredictability, and changeability of the specifications of the said networks, such as channel's band width; channel's capacity; potential for delivery and receipt of nodes; and so forth; with the passage of time. This in turn causes numerous challenges in analysis and design of topology control algorithms. The majority of algorithms which have been presented for control of topology in wireless sensor networks, have been based on the following assumptions:

- In majority of algorithms, the assumption is that each node should always have complete and accurate data on the network's topology. The assumption that in wireless sensor networks, in addition to absence of a fixed and determined network infrastructure; the movement (limited) of nodes and/or their malfunction leads to emergence of continuous topologic changes across the network; imposes many communication slags on the network in order to update nodes' data. This, in turn, leads to wastage of energy in network's nodes, in addition to wastage of band width and the existing capacity of the network's channels.
- Determination of the network's topology based on momentary image of the topologic structure of the network and adoption of measures for supporting future changes; is an approach which has been adopted in majority of the proposed algorithms. Assuming a complete separation between the two phases of formation of clusters and support for change, within networks whose one innate nature is continuous topologic changes, will result in a fall in the efficiency of presented algorithm and wastage of energy within the nodes and band widths of wireless connections.
- Given the continuous changes of topology of network, due to movement (limited) of nodes and/or their impairment; the environments of wireless sensor network are completely dynamic, unpredictable, and uncertain environments. Thus, assumption of the accumulation of all topological data of network's nodes, and implementation of algorithm in a single node, especially within networks that maintain highly volatile nodes, on one hand results in transformation of the single node into a point of failure for the algorithm; and on the other hand will lead to wastage of the existing sources such as the communication channels and source of energy.
- In the majority of proposed algorithms, given the necessity to coordinate and match the performance of nodes; the presence of assuring models for publication and/or collection of control messages across the network is an obvious assumption. This comes while the wireless base of the said networks, movement (restricted) of mobile nodes; and their dynamic presence, despite the corresponding costs, put assuring conditions in place for materialization of this assumption [15].
- The wireless sensor networks maintain a completely distributed network, in which collection of data in one node and/or their publication across the network, imposes countless excessive communication demands on the network; especially in networks with large scales. Additionally, the existing limitations in the

band width of wireless connections and energy of nodes, raise inclinations toward design of algorithms that are capable of making different decisions across the network solely based on the data of nodes, and with the low consumption of sources in a distributed manner. Design of such distributed algorithms makes it possible to carry out updating and supporting operations, in a local form and shape, in case of need and upon the emergence of topologic changes [16].

- Continuous topologic changes of the network, resulting from limited stimulation of nodes and/or their impairment, restriction of the band width of wireless connections, limited energy of nodes, interferences of wireless channels; noisy connections; and a number of other factors have led to focusing upon, and assessing the specifications of the said networks, such as the channel band width; channel's capacity; capacity of node for delivery and receipt; within the framework of topology control algorithms, turning them into random, unpredictable and changeable specifications with the passage of time, causing numerous challenges in the analysis and design of the said algorithms for these random networks. Negligence of the random and variable specifications, and assumption of presence of a network with completely fixed, stable, and certain specifications leads to adoption of inappropriate decisions which are distant from the innate realities of the said networks, and are therefore inappropriate to implement, resulting in a sharp fall in output.

  Given the said cases, and upon precise focus on the features and capacities of learning automata model, which were previously mentioned, it is evident that majority of assumptions are only capable of easing and justifying the presentation and theoretical analysis of proposed algorithms, while implementation of some of them has proven to be a difficult task due to the existing capacities and limitations in wireless sensor networks, and could possibly result in a sharp fall in the efficiency of said algorithms. On this basis, learning automata, and their corresponding combined models can be considered as an appropriate model for resolution of the said issues in wireless sensor networks, due to their following features.
- Learning automata are appropriately capable of matching themselves with the environmental changes. This is a highly appropriate feature for usage in the wireless sensor networks environments that maintain a very high degree of dynamism.
- The learning automata, in addition to low computational needs, impose minor communication costs on the environment throughout their interaction with the environment. This feature makes learning automata an appropriate option for usage in environments, within which there are restrictions of energy and band width, in comparison to other models.
- Learning automata in their interactions with each other are capable of building models of distributed nature of wireless sensor networks and to simulate the changeable behavioral patterns of nodes in communication with each other and the environment, given learning automata's learning and matching potentials [17].

- Learning automata, in interaction with each other, are capable of convergence for resolution of issues related to optimization, only by reliance on local decisions. Thus, algorithms which are based on learning automata are considered as an appropriate choice for wireless sensor networks, given that they dispel the slags resulting from the collection and/or dissemination of information in centralized algorithms.
- Learning automata, in a repeatable process, and with the passage of time, complete the information that they need for decision-making from their living environment. On this basis, the endurance of algorithms which are dependent on learning automata, in the face of possible errors in receipt of data by nodes, which is common in wireless sensor networks, will not leave an impact as such on the performance of algorithm, in comparison to other algorithms.

The main goal of this thesis is presentation of a smart algorithm based on learning automata model, which upon consideration of the limitations of the said networks, and upon reliance on prudent assumptions would be able to provide an appropriate solution for the issue of topology control within wireless sensor networks.

Proposed topology control algorithm based on learning automata

Given that the proposed algorithm is an algorithm aware of energy and efficient energy; with due regard to the remaining energy of the network's nodes, and within the framework of a repeatable learning process, it tries to form the most durable topology of the network via selection of the most energetic nodes of the network and establishment of the most resistant clusters against impairment of the sensor nodes.

The phases of the proposed algorithm are as follows [18–20]:

Initially, a network of learning automata in harmony with the wireless sensor network is established. For this purpose, a learning automata (For instance Automata Ai) is allocated to each of the network's nodes (For instance node Nj). Each automata shapes its internal structure (Action set and possibility vector in selection of action) as follows:

Given that in the proposed algorithm, the action set of each automata are defined based on adjacent nodes; it is necessary for each node of the proposed algorithm to continuously update its topologic data. To this end, each node periodically airs a hello message within the boundary of its delivery. This message is sent by all nodes once in a while in order to update the topologic data of the network. The proposed algorithm in this thesis, is an energy-aware algorithm and on this basis any node, in its hello message, in addition to its ID number, sends and notifies its remaining energy. In this manner, all of the inner-network nodes are continuously informed of the amount of remaining energy in their neighboring nodes.

Each node, upon the receipt of hello message, includes its ID No. and the data on the remaining energy of the node, which has received the said message, within its list of neighbors. The number of fields for each of the nodes of the network in the list of neighbors is equal to the number of its neighboring nodes. In this manner, each of the nodes will maintain the local data of its neighbors. In fact, list of neighbors of each node such as Ni node is presented as equivalent to the length of its neighbors and each of its homes belongs to a neighboring node such as Ni node which maintain the two fields of ID and Energy within the framework of List of Neighbors.

- Assume that Ai Automata is equivalent to Node Ni.
- Assume that αi is the action set for Automata Aj.
- Assume that Ei is the remaining energy of Ni Node.
- Assume that Ei is the average remaining energy of nodes adjacent to Node Ni.
- Assume that α is the action selection of Nj Node by Ni Node.
- Assume that ρi is the vector on possibility of selectin of action of Automata Ai.
- Assume that ρj is the possibility of selection of action αj.
  Now, the action set of each automata allocated to each node is defined as follows with due regard to the location of that node in the network.
- The letter r is the number of actions of each automata which is equivalent to the number of neighboring nodes, from which the automata has received a hello message, plus one.
- Each action of automata is equivalent to its own selection and/or the selection of one of the neighboring nodes from the List of Neighbors.
- Selection of each action of α by the automata is equivalent to Node Ni, tantamount to selection of Node Nj as the cluster head for Node Ni.
- Action probability vector is formed with allocation of the initial value of 1/r to each action when r is equal to the number of actions of each automata.

Given the possibility of impairment of nodes; the network's topology is changeable with the passage of time, and since the structure of automata is shaped based on data on nodes' neighbors; the data of the topology of network should be continuously updated in nodes. Thus, the proposed algorithm continuously updates its topologic data. If, after a specified duration, a hello message is not received from a node which has been adjacent to Node Ni, it would be eliminated from the list of neighbors of the said node and its equivalent action from Automata Ai is also eliminated. Thus, the learning automata which is used in this thesis is the learning automata with a number of variable actions. One of the other reasons for usage of learning automata with a variable action set is to prevent the emergence of loop in the process of development of topology, which will be fully detailed later on.

As previously mentioned, the process of control of topology in the proposed algorithm is carried out in two phases. The first phase is the clustering of network, while in the second phase; the backbone is built via the connection of the cluster head nodes.

The first phase: Clustering of the Network [21–23]

(1) Initially, each Ni node activates its allocated Ai.
(2) Automata Ai chooses one of its possible actions in random and based on the Action Probability Vector ρi.

Each action selected by automata means that the equivalent node to the selected action is considered as the cluster head for the said node.

- Assume and select Nj as the cluster head
- Now, the learning automata equivalent to each node, assesses the optimality of the selected action as follows:

(3) Ni Node calculates the average energy of the remaining adjacent nodes Ei.

(4) If the remaining energy of the chosen cluster head (Ei) would be larger and/or equivalent to the average energy remaining from all neighboring nodes to Node Ni; in this case.

(5) The learning automata rewards its selected action via the usage of following relationship and raises the possibility of selection of that action for the following repetitions.

$$\mathrm{Pi}\,(\eta + 1) = \mathrm{Pi}\,(\eta) + \alpha(1 - \mathrm{Pi}\,(\eta))$$

The possibility of selection of other actions of the automata is reduced upon fining them via the following relationship; with a and b respectfully being the fine and reward rates [24, 25].

$$\mathrm{Pi}\ (\eta{+}1) = (1\text{-}a)\ \mathrm{Pi}\ (\eta)\text{: } \forall j \neq i$$

(6) Otherwise, if the energy of the node of selected cluster head would be less than the average remaining energy of all neighboring nodes.

(7) The learning automata reduces the possibility of selection of the said action (node) via fining it based on the following relationship:

$$\mathrm{Pi}\,(\eta + 1) = (1 - \mathrm{b})\mathrm{Pi}\,(\eta)$$

The learning automata increases the possibility of selection of other actions of automata upon rewarding them via the following relationship [26]:

$$\mathrm{Pi}\ (\eta{+}1) = b/r\text{-}1\ {+}(1\text{-}b)\ \mathrm{Pi}\ (\eta)\text{: } \forall j \neq i$$

(8) The steps one to seven in each node is repeated until the automata equivalent to that node selects one of its actions with the possibility of more than one threshold limit referred to as Pr; that is also referred to as the condition for termination of algorithm. Obviously, the higher the value of Pr, the more will be the cost of implementation of algorithm, and on the other hand the convergence of automata with the optimal cluster (most energetic cluster) takes place with further precision. This means that a more appropriate node is chosen

as the cluster head. However, on the other hand, with the reduction of the value of Pr, the possibility of convergence with the optimal cluster is reduced, in addition to the reduction of the cost of implementation of algorithm.

In this manner, in each section of the network, the node which maintains the highest level of energy in comparison to other neighboring nodes, is chosen as the cluster head node of that section. The selected node, as long as having a remaining energy higher than the energy of other nodes in that section, plays the role of cluster head for the clusters within that section. As a result, upon the usage of learning automata, the inner-network nodes are clustered such that the said clusters (or cluster heads) maintain maximum energy and thereby the maximum lifetime.

Second Phase: Development of Topology

In the second phase of algorithm; the topology of the network is established via connection of nodes of cluster head. To this end, each node of the cluster head identified itself and connects to them, via delivery of the message of nodes of its neighboring cluster head. The process of connection of cluster heads is as follows:

Each node of the cluster head CH forms a message of request for connection, known as CR, in short, sending it to its neighboring nodes.

In this case, it is possible for one of the following cases to occur for each node neighboring Nj, which receives the message CR from cluster head CH [27]:

(A) The two cluster heads are located a step from each other. The node neighboring Nj is a cluster head node. In this manner, two cluster heads Chi and CHj are directly connected to each other.
In this manner, Node CHi also receives its CR message from CHj Node.
(B) Two cluster heads are two steps apart. The Node neighboring Nj neighbors another cluster head such as Chi Cluster Head. In this case, the node neighboring Nj as the gateway node, is duty-bound to connect Chi Cluster Head and CHj Cluster Head.
(C) In the last mode, two cluster heads are three steps apart. In this case, the two cluster heads Chi and CHj are respectively neighboring two gateway nodes Ni and Nj; with the two said nodes also being next to each other. In this mode, the two cluster heads Chi and CHj are connected to each other via two gateway nodes Ni and Nj.

In the end of the second phase of proposed algorithm, the most resistant network topology is formed via the connection of the most energetic cluster head nodes, thereby shaping the most resistant backbone of the network (against the movement and impairment of sensor nodes).

One of the capacities and abilities of learning automata is compatibility with the environment and environmental changes. Given that each topology maintains a limited lifetime, and after a while, the formed topology will lose its energy and will break; under these conditions, in each of the topology control methods, the process of redevelopment of topology should be repeated. But, in the proposed algorithm, with the passage of time, whenever each node of the backbone loses its energy, the learning automata intelligently converges toward the second energetic neighboring node, in the following repetitions; preventing the failure of topology and removing the need for the repetition of algorithm. Thus, it is expected from the proposed method to significantly rise topology's lifetime and to reduce the costs of calculations and communications.

3 Conclusion

In this thesis, a fresh approach based on learning automata was presented for unequal clustering of wireless sensor networks. In unequal clustering, the sizes of formed structures are not equal and the majority of clusters close to sink node will be smaller in size, while the further away clusters will maintain larger sizes. This feature reduces the consumption of energy in the vicinity of the sink node and will increase the network's lifetime. The proposed model can maintain the following steps.

(1) Determination of the cluster head with the usage of a learning automata
(2) Determination of the size of each cluster with the application of optimization model and formation of clusters
(3) Routing data in the clustered network, with the application of routing algorithm
(4) Measures of the automata are updated in the end of each cycle

In the proposed method, the cluster heads are determined with the usage of possibility vector of the learning automata, and the number of neighbors to sensor node is determined. At the end of each cycle, the measures of learning automata are updated with the application of laws on rewards and fines. This feature results in the better performance of the proposed method in selection of the cluster head after a number of cycles.

In order to assess the performance of the proposed model, NS2 software is used. In this measure, the performance of the clustering algorithm in the network's environment was studied from different angles such as the network's lifetime, rate of consumption of energy, rate of delivery of package, and participation of nodes in the delivery process. The results of this study revealed that the proposed method maintained a better performance compared to the performance of other methods, and in addition to a longer lifetime, the proposed method will result in lower consumption of energy, and larger number of delivery of packages.

References

1. Dargie W, Poellabauer C (2010) Fundamentals of wireless sensor networks theory and practice. Wiley series on wireless communications and mobile computing, 1st ed (Wiley Publication)
2. Sohraby K, Minoli D, Znati T (2007) Wireless sensor networks: technology, protocols and applications, 1st ed. Wiley

3. Younis M, Senturk I, Akkaya K, Lee S, Senel F (2014) Topology management techniques for tolerating node failures in wireless sensor networks: a survey. Comput Netw 58:254–283
4. Mamun Q (2012) A qualitative comparison of different logical topologies for wireless sensor networks. Sensors 12(11):14887–14913
5. Alnuaimi M, Shuaib K, Nuaimi K, Hafez M (2012) Performance analysis of clustering protocols in WSN. IEEE IFIP WMNC, pp 1–6
6. Liu X (2012) A survey on clustering routing protocols in wireless sensor networks. Sensors 12:11113–11153
7. Afsar M, Tayarani M (2014) Clustering in sensor networks: A literature survey. J Netw Comput Appl 46:198–225
8. Thathachar MAL, Sastry PS (2002) Varieties of learning automata: an overview. IEEE Trans Syst Man Cybern—Part B: Cybern 32(6):711–722
9. Narendra S, Thathachar AL (2012) Learning automata: an introduction. Courier Corporation, 476 pages
10. Barto AG, Anandan P (1985) Pattern-recognizing stochastic learning automata. Syst Man Cybern 15(3):360–375
11. Forghani A, Rahmani AM (2008) Multi state fault tolerant topology control algorithm for wireless sensor networks. In: Proceedings of the IEEE 2nd international conference on future generation communication and network, vol 1, pp 433–436
12. Wang L, Jin H, Dang J, Jin Y (2007) A fault tolerant topology control algorithm for large-scale sensor networks. In: Proceedings of 8th international conference on parallel and distributed computing applications and technologies (PDCAT), vol 8, pp 407–412
13. Sergiou C, Vassiliou V, Paphitis A (2013) Hierarchical tree alternative path (HTAP) algorithm for congestion control in wireless sensor networks. Ad Hoc Netw 11(1):257–272
14. Gupta G, Younis M (2003) Load-balancing clustering of wireless sensor networks. In: Proceedings of the international conference on communications (ICC '03), vol 3, pp 1848–1852
15. Bhowmik S, Basu D, Giri Ch (2014) k-fault tolerant topology control in wireless sensor network. Adv Intell Syst Comput 235:371–377
16. Basagni S (1999) Distributed clustering for ad hoc networks. In: Proceedings of IEEE fourth international symposium on parallel architectures, algorithms, and networks (I-SPAN 99), vol 8, pp 310–315
17. Xia D, Vlajic N (2006) Near-optimal node clustering in wireless sensor networks for environment monitoring. In: Proceedings of CCECE, vol 6, pp 1825–1829
18. Jiang CJ, Shi WR, Tang XL, Wang P, Xiang M (2014) energy balanced unequal clustering routing protocol for wireless sensor networks. J Softw 23(5):1222–1232
19. Liao Y, Qi H, Li W (2013) Load-balanced clustering algorithm with distributed self organization for wireless sensor networks. IEEE Sens J 13(5):1498–1506
20. Bagci H, Yazici A (2013) An energy aware fuzzy approach to unequal clustering in wireless sensor networks. Appl Soft Comput 13:1741–1749
21. Lee S, Choe H, Park B, Song Y, Kim C (2011) LUCA: an energy-efficient unequal clustering algorithm using location information for wireless sensor networks. Wirel Pers Commun 56:715–731
22. Guha S, Khuller S (1998) Approximation algorithms for connected dominating sets. Algorithmica 20(4):374–387
23. Butenko S, Cheng X, Oliveira C, Pardalos PM (2004) A new heuristic for the minimum connected dominating set problem on ad hoc wireless networks. In: Recent developments in cooperative control and optimization, cooperative systems, vol 3, pp 61–73
24. Heinzelman WR, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In: Proceedings of the 33rd Hawaii international conference on system sciences (HICSS), vol 8, pp 1–10
25. Kour H, Sharma AK (2010) Hybrid energy efficient distributed protocol for heterogeneous wireless sensor network. Int J Comput Appl 4(6):1–5

26. Heinzelman W, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In Proceedings of the 33rd Hawaii international conference on system sciences, vol 8, Citeseer, p 8020
27. Deniz F, Bagci H, Korpeoglub I, Yazici A (2016) An adaptive, energy-aware and distributed fault-tolerant topology-control algorithm for heterogeneous wireless sensor networks. Ad Hoc Netw 44:104–117

Using an Active Learning Semi-supervision Algorithm for Classifying of ECG Signals and Diagnosing Heart Diseases

Javad Kebriaee, Hadi Chahkandi Nejad and Sadegh Seynali

Abstract Diagnosis of various heart defects and arrhythmias based on the ECG signals recorded from the patient has greatly appealed to the medical community. Biological signal processing performed by experts in the field has involved many challenges to be able to present a precise model of the recorded signals and to analyze and diagnose defects and arrhythmias based on the extracted features and to classify them into the normal and abnormal classes. It is an issue that has appealed to researcher for years to make the process of precisely diagnosing heart diseases intelligent. An efficient classification method with active and semi-supervised learning for classification of the ECG signal based on the mRMR feature selection method has been used in this research. The extracted features include the temporal features, AR, and wavelet coefficients. Finally, the indicators of validity, precision, and sensitivity for this set of selected features have also been evaluated through application of the proposed classifier. The results of simulations in the Matlab software environment suggest that the proposed system has 98.64% validity for diagnosis of 6 class types of ECG. Comparison between the obtained precision and that of the previous research demonstrates the proper performance of the proposed method.

Keywords Classification · ECG signals · Active learning · Semi-supervised learning

J. Kebriaee · S. Seynali
Computer Engineering Department, Islamic Azad University,
Birjand Branch, Birjand, Iran

H. Chahkandi Nejad (✉)
Electrical Engineering Department, Islamic Azad University,
Birjand Branch, Birjand, Iran
e-mail: Hchahkandin@iaubir.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_6

1 Introduction

Diagnosis of different types of cardiac anomaly and arrhythmia based on the ECG signals recorded from the patient has greatly appealed to the medical community. Experts in the field have been confronted with many challenges processing biological signals, so that they can present an accurate model from the recorded signals that can be analyzed making it possible to diagnose anomalies and arrhythmias and to classify them as normal or abnormal based on the extracted features. Making the process of accurate diagnosis of heart diseases intelligent is an issue that has appealed to researchers for years. In this research, an efficient classification method with active semi-supervised learning has been used for classification of the ECG signal based on the mRMR feature selection methods. The extracted features include temporal features, AR, and wavelet coefficients. Finally, the correctness, accuracy, and sensitivity indices are evaluated for this set of selected features through application of the proposed classification. The results of simulation in the MATLAB software environment suggest that the proposed system has an accuracy of 98.64% for diagnosis of 6 ECG class types, which demonstrates the desirable efficiency of the proposed method as compared to the accuracy obtained in previous research.

Diagnosis of different types of cardiac anomaly and arrhythmia based on the ECG signals recorded from the patient has greatly appealed to the medical community. Experts in the field have been confronted with many challenges processing biological signals, so that they can present an accurate model from the recorded signals that can be analyzed making it possible to diagnose anomalies and arrhythmias, to classify them as normal or abnormal, to diagnose the type of arrhythmia for presentation of the source of anomaly, and to diagnose ventricular tachyarrhythmia, atrial fibrillation, and congestive heart failure based on the extracted features. Many methods and algorithms have been presented so far for analysis of ECG biological signals in the field of frequency, which have classified the cardiac signals of a normal individual and those of one with a specific arrhythmia based on the frequency ranges of the signals. Analysis in the field of frequency is affected to a large extent by signal noises resulting from movement of the patient, the electric equipment noise of the device, etc., and it is not possible to predict and analyze beyond the time of recording and for future. Furthermore, it is impossible in the field of frequency to make a distinction between arrhythmia and noises, and primary diagnosis of CHF is impossible. For this reason, there has been concentration to some extent in the few recent years on analysis of these signals in different fields.

In this research, we will detail the different steps of implementing the proposed method, and will also investigate the effects of changing the parameters effective on implementation of the method. The method has been implemented in the MATLAB software environment. The steps of signal preprocessing will first be described, and the steps concerning extraction and selection of features will then be addressed, followed finally by data classification, for classification of the normal signal and the 5 arrhythmias [1].

2 History

It has been debated for years by researchers in all countries to make the process of diagnosing heart diseases intelligent. The process consists of steps during which the ECG signal is selected as the input to the software, and the software is expected to diagnose well-being or disease and even the type of heart disease with an acceptable accuracy. All these pieces of software extract and select the appropriate features of a signal after receiving it, and then diagnose the type of disease. We will examine different methods used in previous research below.

2.1 Classification of the ECG Signal Using a Wavelet, Morphological Properties, and Neural Networks

In this research, 15 temporal features and 15 features of transform of the selected wavelet have been used after the preprocessing and the PCA method for reduction of the feature sizes, which has resulted in selection of 8 of the best features in each class. Classification is made by a combination of the multilayer perceptron neural network and the radial basis neural network. It has been demonstrated in this research that the hybrid structure of the neural network obtains much better results than the multilayer perception (MLP) neural network [2].

2.2 Classification of Cardiac Arrhythmias Using SVM

In this research, the features of the ECG signal have been extracted through its analysis with a combination of wavelet transform and the AR model. The common methods of cardiac disease diagnosis have been optimized with such an integration. Then, a support vector machine classifier with a Gaussian kernel has been used for automatic classification of five types of cardiac arrhythmia.

2.3 Electrocardiogram Signal Processing

Biological signals are processed in 4 steps.

Measuring or recording the signal: Transformers are used for recording and collecting signals from the body.

Transforming the signal: This step is referred to as preprocessing. The purpose is to reduce signal noise and data size, so that the signal features are easier to extract in the third step. Calculating the signal parameters: The step consists of extraction of the appropriate meaningful parameters (signal features) [3]. Interpreting or classifying the signals: The physician or computer introduces the final interpretation using the extracted features.

3 Pattern Recognition

One of the important purposes of recording and processing critical signals is to interpret them and to employ the useful information in them in diagnosis and treatment. The interpretation step occurs in the recognition or classification phase. For instance, it must be specified after the ECG signal is recorded and preprocessed whether or not it concerns an individual with a particular heart disease. Signal classification actually responds to this question.

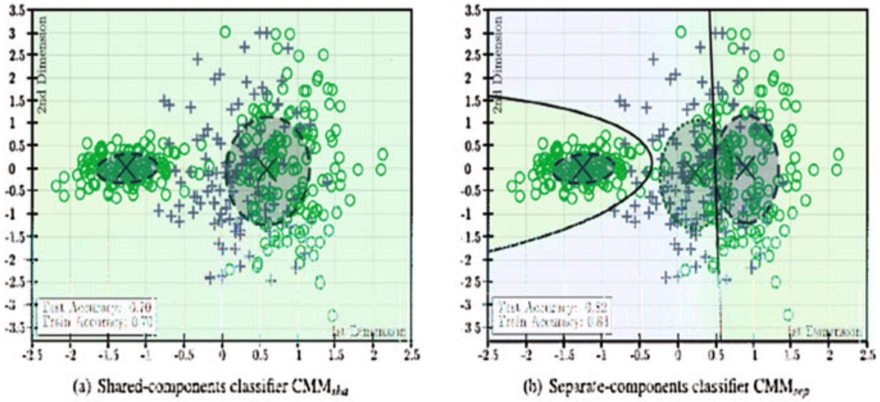
3.1 Pattern Recognition Methods

The methods for pattern recognition fall in general into three groups. Statistical methods: Statistical models are generated for patterns and classes, and classification is made using the concept of probability distribution. Structural (analytical): The pattern classes are specified by figurative structures. These methods are applied most often in cases where the patterns have specific structures [4]. Intelligent networks: Artificial neural networks are networks of units that model the brain neurons. Classification is made using these structures [5].

4 Proposed Method

Classifiers are sometimes based on generating models, since they model observed data (samples), can be trained using partially labeled datasets, and make it possible to integrate expert knowledge easily. However, it is important to identify overlapping processes belonging to different classes and incomplete data distribution matching in order to obtain great classification performance. That is why we use two different methods of training technique for the CMM probabilistic modeling approach (classifier based on probabilistic hybrid models). The first method that we use, titled shared component classifier (CMMsha), determines structural information in an unsupervised way in the first step with a density model of shared components, and is extended to a classifier using class labels [6].

The second classifier, titled separate component classifier (CMMsep), uses class information at present for training, first for construction of the density model of separate classifiers and then for allocation of the components to classes.



(a) Shared-components classifier CMM_{sha}

(b) Separate-components classifier CMM_{sep}

Fig. 1 Comparison of the two types of CMM over an artificial dataset

In Fig. 1, the differences between the two approaches of CMMsha on the left and CMMsep on the right are shown for a two-dimensional continuous (real-value) input space, where we use bivariate Gaussians (normal distributions, in other words) as model components. Each component should model one class of samples in the input space, which can be assumed to result from a process in the actual observed environment. We have 3 such processes that belong to two classes, where the circles are red, and the crosses are blue. The two processes, one blue and one red, overlap each other to a large extent. The center (average) of the Gaussian component describes its location, covariance matrix, and shape. The curves in Fig. 1 are Gaussian curvatures located at the centers represented by large crosses. The decision boundary is shown as a thick black line. In the first step of modeling, CMMsha on the left does not use label information, and is therefore incapable of differentiating between overlapping processes or classes. CMMsha recognizes only two classes, and places both in the green class, and incapability of realizing a high classification rate occurs subsequently. CMMsep on the right, which uses class information in the first step of modeling, is capable of recognizing up to 3 classes, modeling them correctly, and providing us with higher classification accuracy.

4.1 A General View of the Active Learning Approach Using Directed Learning

The standard PAL learning cycle (without our novel directed learning process, in other words) is represented by thick black arrows in Fig. 2a. PAL usually begins with a large repository U of unlabeled samples (the gray box on the left) and a small set of labeled samples L (the gray box on the right), where X = U U L and $|L| \ll |U|$. A classifier G is trained based on L. Then, a question-and-answer set of unlabeled samples is specified for labeling based on selection strategy Q, which

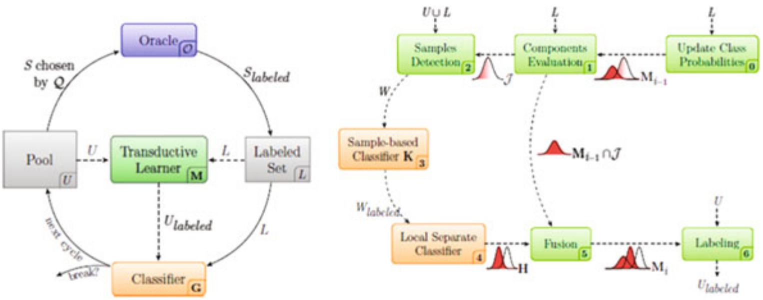


Fig. 2 Graphic description of the active learning process

considers the knowledge contained in G for sample selection, and is presented to predictor O. The labeled samples are added to L, and classifier G is updated. If a termination condition is met, PAL is stopped; otherwise, a new round of questions and answers begins (learning cycle, in other words). Our extension of the standard PAL cycle is shown using the additional dotted arrows in part a of Fig. 2 [7].

4.2 *Our Approach Differs from the Standard Approach from the Following Aspects*

The initial set L is empty.

A directed trainer has been used (part a of Fig. 2, the green box in the middle), which adopts a generating model M in each cycle, and uses M for labeling all the samples in U with a semi-supervised approach.

Classifier G is trained with the labels for all the samples in the dataset.

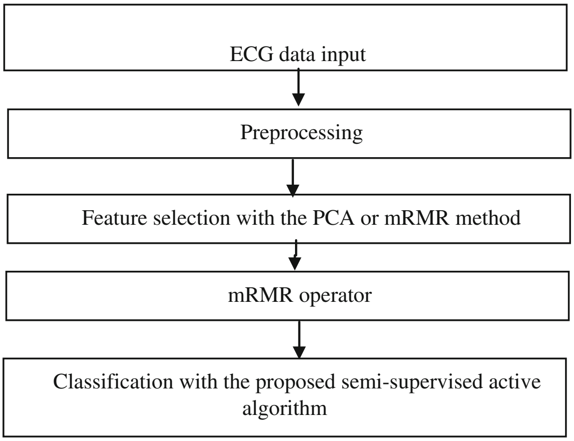
(a) PAL learning cycle (full arrows) with extensions (dotted arrows).
(b) Directed learning process, extending the standard PAL cycle, concerning the green directed learner box in part a of Fig. 2.

5 Simulation

The method has been implemented in the MATLAB software environment. The steps of signal preprocessing will first be described, and the steps concerning extraction and selection of features will then be addressed, followed finally by data classification, for classification of the normal signal and the 5 arrhythmias.

The ECG signals concerning different patients have been provided by the MIT-BIH standard database. The database has 48 two-channel ECG signals

Fig. 3 Flowchart for the order in which the algorithms and techniques applied in the proposed method are run


ECG data input
Preprocessing
Feature selection with the PCA or mRMR method
mRMR operator
Classification with the proposed semi-supervised active algorithm

obtained from 47 case studies at the BIH arrhythmia laboratory between 1975 and 1979. The signals have been stored with a frequency of 360 samples per second and an accuracy of 12 bits in a range of 10 mV, and 20 patients' signals were selected as input data (Fig. 3).

5.1 Signal Shift to the Baseline Deviation

The noise resulting from breathing when the electrocardiography signal is recorded has a low frequency of about 15 Hz. The noise causes the baseline of the electrocardiography signal to change, as a result of which extraction of the temporal properties and features of the signal is confronted with problems. The level-8 wavelet of the ECG signal has a higher amplitude than at the previous wavelet levels, and the ECG signal is corrected if it exhibits a considerable difference from the baseline at some pulses (Figs. 4 and 6) following the removal of the level, so that a uniform, regular signal is achieved.

5.2 Removal of the Noise Resulting from Mains Electricity

The ECG signal at this level includes noise resulting from mains electricity. Through application of a band-pass filter that does not pass signals in the range of 60 Hz, the noise resulting from mains electricity, which has a higher frequency than that of the main signal, can be filtered (Figs. 5 and 6) [8].

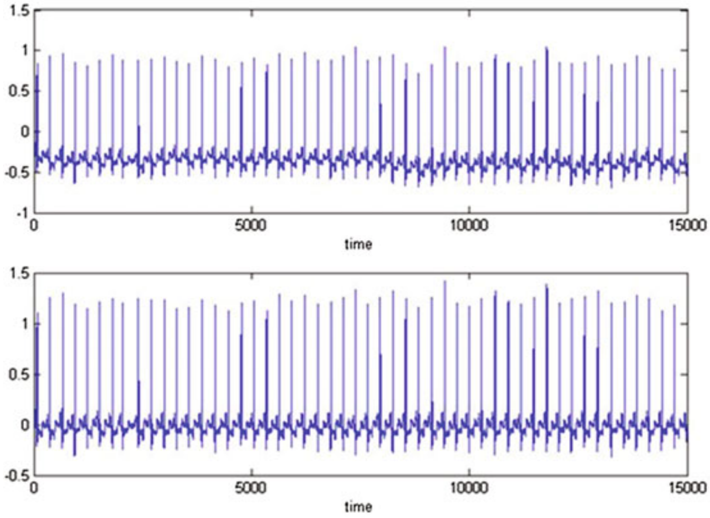
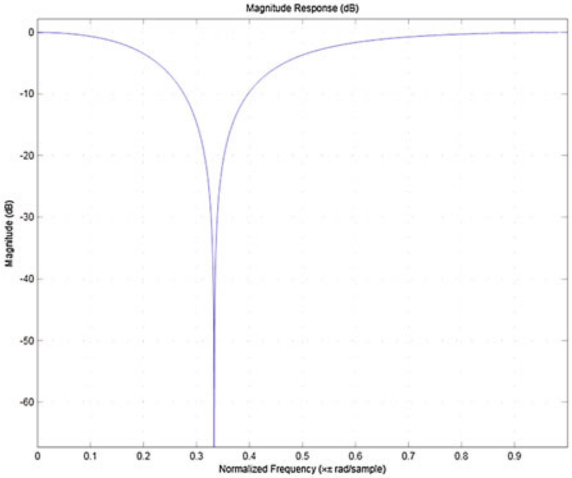


Fig. 4 50 heartbeats where the baseline deviation noise has been removed

Fig. 5 Removal of the 60-Hz frequency with a notch filter



5.3 *Signal Windowing*

Windowing is carried out using the information extracted from the software environment, such as the moment when wave R has occurred and the type of disease diagnosed by the physician for each pulse. For this purpose, 100 samples before and 200 after the moment when wave R has occurred of the smoothed signal is considered as a complete pulse (Fig. 6), and the pulse is placed in the class concerning its disease based on the physician's diagnosis on it.

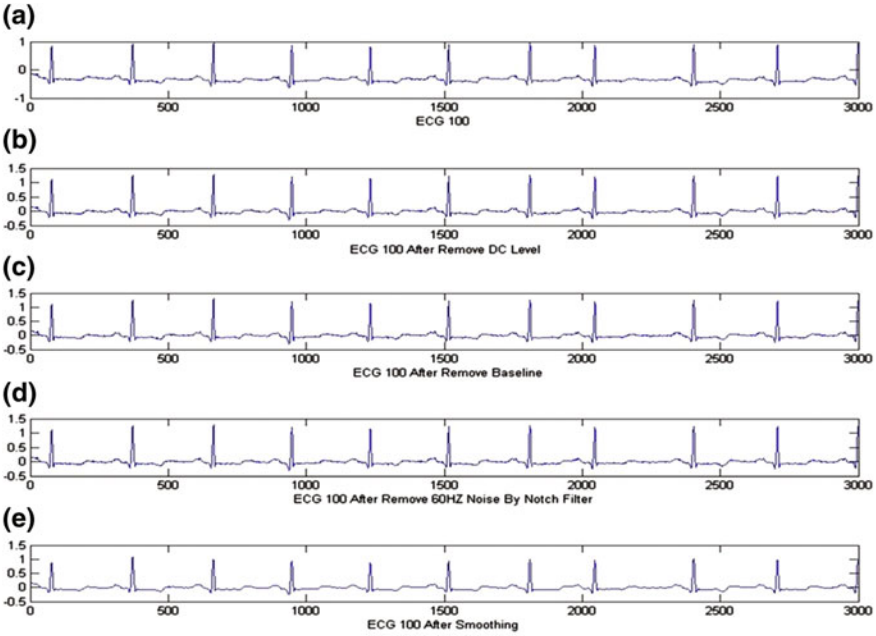


Fig. 6 **a** 100 ECG signal, **b** ECG signal after removal of the DC value, **c** ECG signal after removal of the baseline deviation noise, **d** ECG signal after removal of mains electricity, **e** ECG signal after smoothing

5.4 Selection of the Training and Test Data

At the final step of signal preprocessing, 1060 pulses are fully randomly selected from each of the six classes, out of which 750 pulses are considered as training data and 310 as test data, and are stored in a separate classification (Given the low amount of data available in the database on the disease A signal, 224 pulses have been considered as training data and 97 as test data) [9].

5.5 Feature Extraction

Temporal features, wavelet features, and AR features were selected for the feature extraction step, which were extracted for the training and test data. Each extracted feature vector consists of a total of 64 features, including 9 temporal features, 48 wavelet features, 2 AR features, and 5 PCA features.

Feature selection with PCA.

When the PCA operator is applied, the features are mapped in a new space, and sorted based on importance. The superior features are selected, and stored for

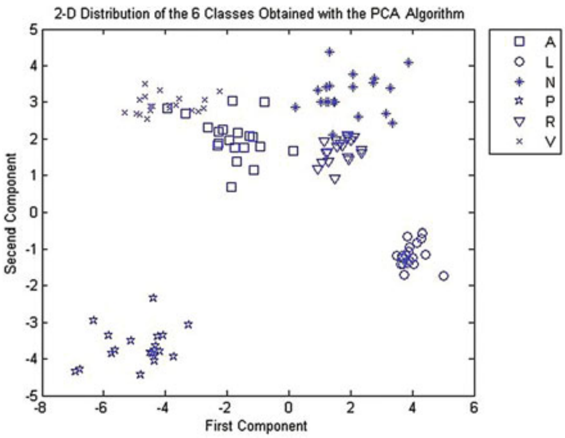


Fig. 7 Distribution of 20 feature vectors from each class after PCA mapping shown in two dimensions

Table 1 Accuracy of classification using 5 classification methods and 40 selected features

Method	A	L	N	P	R	V	Accuracy
PCA.SVM	93.67	81.44	82.58	85.66	84.02	80.30	84.61
mRMR.SVM	94.05	82.55	82.90	86.12	80.26	83.71	84.93
PCA.Porposed	94.11	99.40	98.91	99.64	99.70	99.03	98.47
mRMR.Proposed	94.11	99.05	99.54	99.33	100	99.82	98.64

classification in the following steps. The distribution of 20 feature vectors from each class after PCA mapping is shown in two dimensions in Fig. 7. Table 1 shows the output of the classifier.

Identification of the important points of the signal using PCA.

We consider all the points of the signal as the input to PCA. Given the PCA mapping, the mapped points are ranked based on importance. We select 5 superior mapped points as features to be used in the following steps [10].

6 Principal Component Analysis (PCA)

In the principal component analysis method, new coordinate axes are defined for the data, such that the first axis lies in the direction that maximizes data variance, the second axis is considered perpendicular to the first axis in the direction that maximizes data variance, and the following axes similarly lie perpendicular to all the previous axes to maximize data variance in that direction. Principal component analysis is one of the popular methods of feature extraction, used in many studies due to simplicity and high processing speed. The PCA technique is the best method for linear reduction of data dimensions; that is, less information is lost than in other methods, since the less important coefficients obtained from the transformation are eliminated. Assume that input matrix Xhas N_T samples and n features, and the N_T samples should be placed in C groups. The data mean and covariance are calculated based on the following equations.

Equation (1):

$$\begin{aligned} \mathrm{m_d} &= \frac{1}{\mathrm{N_T}} \sum_{\mathrm{i}=1}^{\mathrm{c}} \sum_{\mathrm{j}=1}^{\mathrm{N_i}} \mathrm{x_{i,j}} \\ \mathrm{COV} &= \frac{1}{\mathrm{N_T}} \sum_{\mathrm{i}=1}^{\mathrm{c}} \sum_{\mathrm{j}=1}^{\mathrm{N_i}} \left(\mathrm{x_{i,j}} - \mathrm{m_d}\right)\left(\mathrm{x_{i,j}} - \mathrm{m_d}\right)^{\mathrm{T}} \end{aligned} \tag{1}$$

In the next step, the eigenvalues and eigenvectors are calculated based on the covariance matrix. Then, k greater eigenvalues are selected of the total n. Now, input matrix X is transformed under eigenvector matrix P with k features to the principal component analysis space.

Equation (2):

$$\mathrm{Y_{ij}} = [\mathrm{P_1}, \mathrm{P_2}, \ldots, \mathrm{P_k}]^{\mathrm{T}} \mathrm{X_{ij}} \tag{2}$$

6.1 Feature Selection with mRMR

Through integration of the features and its application to mRMR, the dimensions are sorted in terms of importance. In mRMR, the feature dimensions are sorted in all the classes through selection of the best with the criteria of maximum dependency and minimum redundancy, and no mapping is carried out there.

6.2 Classification Using the Proposed Method

Learning and test: First, all the ECG segments including a particular type of heartbeat are mapped onto the feature space using the wavelet, AR, and temporal features. In the learning phase, the proposed network receives a few samples as input. These patterns are heartbeats shown by m feature parameters that can be seen as points in an m-dimensional space. The network is then capable of obtaining the labels of the new vectors by comparing the samples used in the learning phase. The network receives the training and test data, and carries out the classification. Finally, the network output is compared to the test data, and the error rate of the classifier and classification percentage are obtained by counting the pulses classified incorrectly.

In classifiers where mRMR is used as feature selection, this method of feature extraction achieves the best classification accuracy with the lowest number of selected features.

$$\begin{aligned} ERROR &= \frac{Number\ of\ False\ Classified\ Beats}{Number\ Beats} \\ Accuracy &= (1 - ERROR) \times 100 \end{aligned} \tag{3}$$

Table 1 shows the accuracy of ECG signal classification using 5 classification methods. The classification has been carried out through selection of 40 features and 750 pieces of training data. In this table, the proposed method and the support vector machine method have been compared to each other for two different instances of feature extraction.

Figure 8 shows the training procedure of the proposed network in 21 iterations.

A change in the number of selected features causes a change in classification accuracy. Figure 9 shows classification accuracy for changes in the number of selected features with 750 training data, and the proposed network selects the best solution from among the features.

In this paper, we described the steps in simulation of 5 different methods of ECG signal classification. Classification was divided into the 4 major steps of preprocessing, feature extraction, feature selection, and classification. Among the implemented methods, use of the mRMR. Proposed classification exhibits acceptable accuracy as compared to the other common methods. Table 2 shows the other comparable indicators in regard to this research, such as sensitivity.

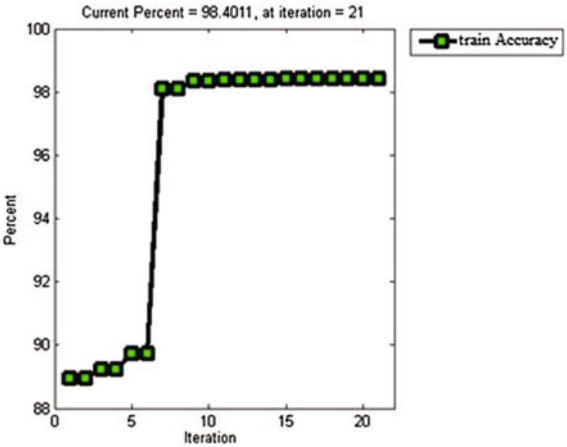


Fig. 8 Selection optimization procedure for the proposed semi-supervised network in 21 iterations of the algorithm

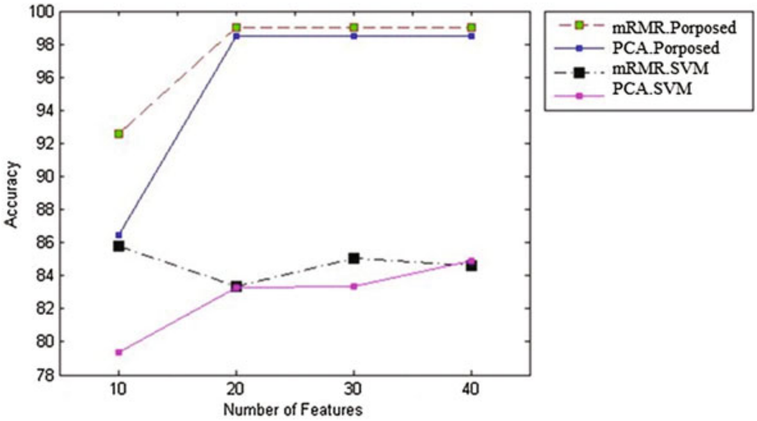


Fig. 9 Accuracy of classification using 4 classification methods for 10, 20, 30, and 40 dimensions of the feature vector

Table 2 Statistical results of implementing the proposed network and support vector machine for two feature selection methods

Feature(s)-network	Statistical parameters		
	Sensitivity (%)	Specificity (%)	Accuracy (%)
PCA-SVM	84.32	85.53	84.61
mRMR-SVM	85.32	86.63	84.93
PCA-Proposed	96.71	96.82	98.47
mRMR-Proposed	97.32	97.73	98.64

7 Conclusion

In this research, the classification of 5 cardiac arrhythmias and the normal signal selected from the MIT-BIH database were investigated. For each arrhythmia, 70% was used for training the system and the remaining 30% for testing the system. As shown, use of mRMR for optimal feature selection and of the proposed classifier network is an appropriate approach for achieving 2 purposes:

(1) increasing classification accuracy
(2) selecting fewer features (superior features).

Selection of fewer features will reduce computation, while higher accuracy was observed where fewer features had been selected than in the case where all the features had been applied to the classifier. Furthermore, selection of wavelet features and features extracted from the signal mapped by PCA has an important role in improvement of the precision of classification, in such a way that of the total of 6 superior features, 1 is a wavelet feature, and 2 are features extracted from the PCA-mapped signal. As compared to previous research with the same patient ECG signals as in this one, it is observed that application of the proposed method to these data obtains favorable results. Thus, an accuracy of 84.93 has been obtained with 750 pieces of training data and the mRMR.SVM method, while implementation of the mRMR. Proposed method with 750 pieces of training data shows an accuracy of 98.64.

8 Features Works

Finally, we make suggestions for achievement of better diagnosis of heart disease based on the results obtained from analysis of the collected dataset.

(1) Addition of further independent features can reduce algorithm error.
(2) Use of combinations of various classifiers can improve the rate of ECG signal recognition.
(3) Use of two-dimensional features using a Hankel matrix.

References

1. Ghorbanian P, Ghaffari A, Jalali A, Nataraj C (2010) Heart arrhythmia detection using continuous wavelet transform and principal component analysis with neural network classifier. In: Computing in cardiology IEEE conference, pp 669–672
2. Khazaee A (2013) Heart beat classification using particle swarm optimization. I.J. Intell Syst Appl 5(6):25–33
3. Rodrigues D, Pereira LAM, Almeida TNS, Papa JP, Souza AN, Ramos CCO, Yang X-S (2013) BCS: a binary cuckoo search algorithm for feature selection, circuits and systems (ISCAS), 2013 IEEE international symposium, pp 465–468
4. Reitmaier T, Calma A, Sick B (2015) Transductive active learning–a new semi-supervised learning approach based on iteratively refined generative models to capture structure in data. Inf Sci 293:275–298
5. Ubeyli ED (2010) Lyapunov exponents/probabilistic neural networks for analysis of EEG signals. Expert Syst Appl 37:985–992
6. Ubeyli ED (2010) Recurrent neural networks employing Lyapunov exponents for analysis of ECG signals. Expert Syst Appl 37:1192–1199
7. Guler NF, Ubeyli ED, Guler I (2005) Recurrent neural networks employing Lyapunov exponents for EEG signals classification. Expert Syst Appl 29:506–514
8. Bishop CM (2006) Pattern recognition and machine learning. Springer, New York
9. Duda RO, Hart PE, Stork DG (2001) Pattern classification. Wiley, Chichester
10. Fisch D, Sick B (2009) Training of radial basis function classifiers with resilient propagation and variational Bayesian inference. In: International joint conference on neural networks (IJCNN ’09), Atlanta, pp 838–847

Automatic Clustering Using Metaheuristic Algorithms for Content Based Image Retrieval

Javad Azarakhsh and Zobeir Raisi

Abstract Development of internet networks and mobile phone tools with image capturing capabilities and network connectivity within the recent years have led to defining new services and applications, using such tools. In this article, automatic clustering method using evolutionary and metaheuristic algorithms used in order to identify and categorize various kinds of digital images. For this purpose, a database of images prepared, and then k-means clustering method using evolutionary algorithms and optimization applied on these images. The results of retrieval indicate that automatic clustering using particle swarm optimization (PSO) algorithm has higher average retrieval accuracy in comparison with other methods.

Keywords Image retrieval · Feature extraction · Automatic clustering
Evolutionary and metaheuristic algorithms

1 Introduction

Expansion increasing of Internet and availability of imaging tools such as digital cameras and image scanners in recent decade led to generation of a huge volume of images per day [1–3]. Digital images include images of textures, natural images, images from animals and plants, digital signatures, figure print images, face images, digital maps, medical images, artistic images etc. These images, in any case, are increasing and expanding day by day. Thereby, how to find the image of concern in the increasing databases has become one of the very important research questions.

For this purpose, traditional image retrieval algorithms that are presented based on context cannot satisfy the needs of the operator any more [4]. Therefore, nowadays Content Based Image Retrieval (CBIR) algorithms are now of more

J. Azarakhsh (✉) · Z. Raisi
Faculty of Marine Engineering, Chabahar Maritime University, Chabahar, Iran
e-mail: j.azarakhsh@cmu.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_7

considered by people [5]. In a CBIR process, features of the whole images available in the database are extracted first and are effectively saved, then for a sample image (searched case) after extracting its features using a criterion, the extracted features are compared with the saved features of all images and after sorting similarity spaces, then, the results of retrieval are presented to the operator. The process of the CBPI system studied in this article is demonstrated in Fig. 1.

In CBIR, images are indexed using their own visual contents such as color [6], texture [7] and shape [8]. One of the common methods in image retrieval systems is color histogram, as it is very simple and is calculated rapidly. Furthermore, color histogram is resistant to noise and image rotation. Texture is one of the visual and internal features of images which is not dependant on color and brightness of the image. This feature reflects the homogeneity of the image. Texture contains the information of the image surface and also the information on the surrounding ambient in the image. The special information of the image is expressed quantitatively using the texture feature of image. Shape is one of the fundamental features for displaying objects in images, in a way that utilization of this feature can improve accuracy and efficiency of the CBIR systems. Generally there are two methods for displaying shape features: one is contour based and the other is area based [9, 10]. The methods that are utilized in order to extract the features of these two types of shape display are Fourier descriptor and fixed moments, respectively.

Although there are several complex algorithms designed for describing color, texture and shape, however, these models are incapable of presenting a good model for image semantics and do not return an appropriate retrieval when dealing with huge image databases. Extensive experiments on CBIT system indicate that low level contents (such as color, texture and shape) are often incapable of describing high level semantic concepts that are understood by an operator from an image [11, 12]. The performance of CBIR, therefore, is still beyond the expectations of operators.

Clustering is one of the problems that is discussed in the field of machine learning and sight and also has applications in various fields. There are several methods for solving this problem. However, the classic algorithms that are used for solving the clustering problem sometimes do not have the necessary efficiency for solving these problems. Thereby, metaheuristic algorithms or smart optimizations

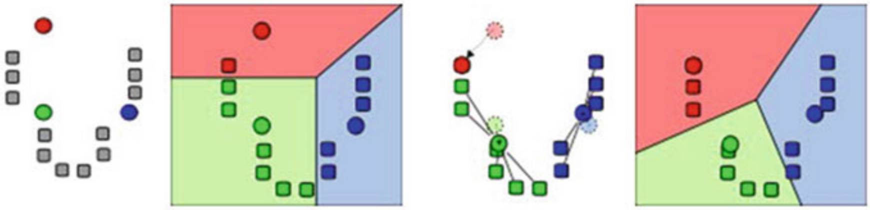


Fig. 1 Lloyd’s algorithm for solving K-means problem

can be used for solving these problems. In order to transform a clustering problem to an optimization problem, we need indices, which for two well-known index in the field of evolutionary data mining are utilized in this article.

Genetic Algorithm or in short GA is absolutely the most known method for smart optimization and evolutionary algorithm which has several applications in different scientific and engineering majors. The importance of this algorithm is so much for evolutionary calculations and computational intelligence that the first term that comes to mind after the term “Evolutionary Algorithm” is the “Genetic Algorithm”. Many people recognize other smart optimization methods as modified versions of genetic algorithm and do not believe in the authenticity of existence and nature for such other algorithms. This computational tool was emerged in the early 1970s from the heart of the outcomes derived from the efforts of the engineers and scientists of those days for simulation of the evolution process. The innovator if the idea of genetic algorithms was John Holland and after him one of his students, David Goldberg put much effort into developing genetic algorithms [13].

Particle Swarm Optimization, in short PSO, is one of the most important smart optimization algorithms that is classified under the category of Swarm Intelligence. This algorithm was introduced by James Kennedy & Russell C. Eberhart in 1995 and is designed inspired by the social behavior if animals such as fishes and birds that live together in big and small groups. In PSO algorithms, the members of the population are directly interrelated and proceed with solving the problem through exchanging information with each other and remembering the good memories of the past [14].

Differential Evolution algorithm or in short DE is a smart optimization algorithm and is based on population which was introduced by Storn and Price in 1995. The primary version of this algorithm was presented for solving continues problems, however, other version of this algorithm were presented gradually that were designed for solving discrete optimization problems [15–18].

Honeybees are among the insects that live together in relatively huge colonies. In addition to the benefits that are obtained from these useful insects in terms of agriculture, gardening and production of honey and wax, their social behavior has always been an origin of inspiration and a basis for scientific researches. Different versions of optimization algorithms are presented until now that are inspired by the group behavior of bees. The Artificial Bee Colony, in short ABC, is recognized as one of the well-known versions of the algorithms that are based on the behavior of honeybees [18].

The focus, in this article, is on applying clustering algorithm on images and finding a retrieval system based on accurate and efficient content for such images.

In the following, we applied a content-based image retrieval system on various types of images, which for, we have extracted shape, color and texture features of the images and thereby proposed a method that automatically extracts the best clusters from the image using evolutionary and metaheuristic algorithms, which this proposed method leads to increased accuracy of image retrieval.

2 Clustering

Clustering is a non-regulatory learning problem. For this problem, in general state, we categorize a set of objects in a way that nobody has provided us with no information.

Clustering has two very important feature and purpose:

(A) *Conjunction*: stating that all members in a formed group should have maximum similarity with each other or in other words, such group should have maximum conjunction.

(B) *Separation*: Stating that the groups should be defined under this condition in a way that they should have the minimum similarity with each other.

In general, therefore, the members of a group should have maximum similarity with each other and the members of groups should have maximum separation or minimum similarity with each other. These maximum and minimum terms bring this question to mind that clustering is an optimization problem. Therefore various evolutionary and metaheuristic algorithms can be used for solving such problems.

2.1 K-Means Problem

One of the most important problems that are discussed in the topic of clustering is K-means [19]. In general, K-means problem is stated as following:

Assume the dataset of x ad following:

$$X = \{x_1, x_2, \ldots, x_n\}, x_i \in \Re^D \tag{1}$$

And the cluster centers as following:

$$M = \{m_1, m_2, \ldots, m_n\}, m_j \in \Re^D \tag{2}$$

If x_i are located in a d dimension space, then m_i are in the d dimension space as well. The total number of the unknowns is $k \times d$.

2.2 Objective Function

Is the total for all clusters (j = 1, 2,…, k), therefore the total distance of all these cluster members to the center is calculated and the objective function is defined as following:

$$Obj.Fun = \sum_{i=1}^{n} \min_{1 \le j \le k} d(x_i, m_j) \tag{3}$$

In the above equation, when x is a member of cluster m_j, then it has the shortest distance to the center of such cluster. Therefore the minimum distance between such x and m_j for all data is calculated as following:

$$Obj.Fun = \sum_{i=1}^{n} \min_{1 \le j \le k} d(x_i, m_j) \tag{4}$$

In general, therefore, we have a dataset comprising of n vectors with d dimension and we want to locate k points with d dimensions in the same space as the cluster centers in a way that the objective function becomes minimum (x and k are known).

2.3 Lloyd's Heuristic Algorithm

One of the algorithms that by standard is utilized for solving k-means problems is Lloyd's algorithm. This algorithm include two phases:

Competitive Phase: in this phase, each member of the dataset is allocated to the closes cluster center or in other words each cluster center competes with other cluster centers and takes the points that are closest to it.

Update Phase: After the competitive phase, the members that are placed in a cluster proceed as holding an election and thereby replace the cluster center with their own mean. Figure 1 displays the mechanism of this algorithm:

2.4 Automatic Clustering

One of the complex types of clustering problems is addressed when the number of clusters is unknown as well and the clustering algorithm is obliged to find the number of clusters. This problem is the so called "Automatic Clustering" different than the "Automatic Classification" which represents the clustering problem itself. In the k-means problem, if the number of unknowns is not known, then this problem becomes an automatic clustering problem. Automatic clustering, in general and according to Fig. 2, is comprised of two parts: One is activation thresholds and the other is cluster centroids.

Solving a clustering problem in general and an automatic clustering problem in particular, can sometimes go beyond the power of common clustering algorithms. One of the solutions that is considered for these cases is transforming the clustering problem into an optimization problem and solving it utilizing smart optimization and evolutionary algorithms [19] which is the case of our discussion in this article

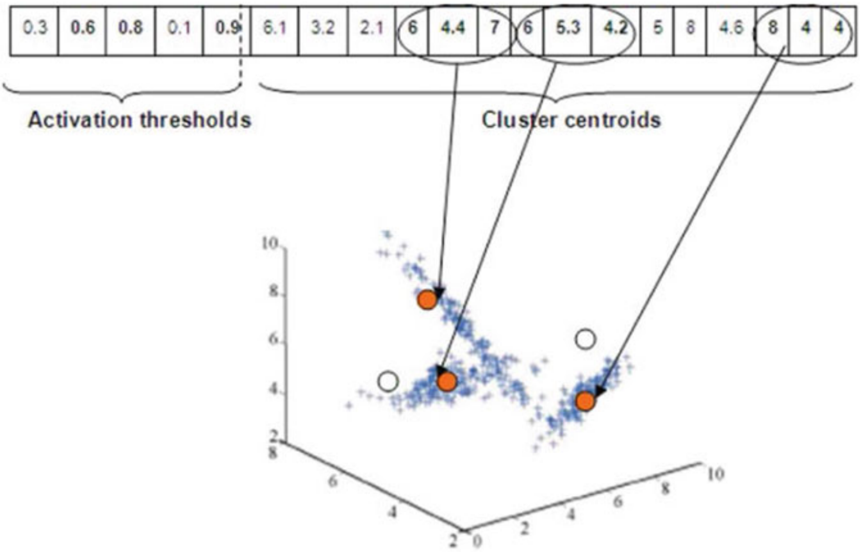


Fig. 2 Coded display of different components of automatic clustering [19]

too. In this article, five smart optimization algorithms including Genetic Algorithm, Particle Swarm Optimization, Differential Evolution and Artificial Bee Colony are used in order to solve automatic clustering problem. For this purpose we use the indexes that are provided in the following.

2.4.1 Validation Indexes

The clustering problem that was presented in the previous section did not consider separation state appropriately; especially when the number of clusters changes, then k-means problem do not provide an acceptable response. Validation index or criterion cause to consider the two important clustering factors (i.e. conjunction and separation).There are many indexes for this purpose such as Dunn's Index (DI) [20], Calinski–Harabasz Index [21], DB Index [22], Pakhira Bandyopadhyay Maulik (PBM) Index [23] and CS Measure [24]. In this article, DB and CS indexes are chosen for automatic clustering, as these indexes provide better responses comparing to others [19].

2.4.2 DB Index

DB or Davies and Bouldin Index was introduced in 1979 [21] and is used for evaluation of automatic clustering validation. In general, this index is defined as following:

$$\mathrm{DB} = \frac{\text{Distance inside Clusters}}{\text{Distance between Clusters}} \tag{5}$$

According to the above equation we are after minimizing the distance inside clusters and maximizing the distance between clusters which leads to decreased DB index value. On the other hand, as the relation between distance and similarity is reverse, therefore decreased DB value cause increased similarity.

Dispersion of the members of a cluster is defined as following:

$$S_{i,q} = \sqrt[q]{\frac{1}{N_i}\sum_{x \in c_i} d(x, m_i)^q} \tag{6}$$

where N_i is the number of the members of cluster i and $d(x,\ m_i)$ is defined as following:

$$d(x, m_i) = \|x - m_i\|_2 \tag{7}$$

The larger is the value of $S\ _{i,q}$, the space covered by cluster i is larger. The distance between two cluster centers is defined as following:

$$d_{i,j,t} = \sqrt[t]{\sum_{p=1}^{d} \left|m_{i,p} - m_{j,p}\right|^t} = \left\|m_i - m_j\right\|_t \tag{8}$$

And the distance inside clusters is defined as following:

$$R_{i,q,t} = \max_j \frac{S_{i,q} + S_{j,q}}{d_{i,j,t}}, \quad j \neq i \tag{9}$$

The above parameter indicates that cluster i has the maximum separation.

$$DB = \frac{1}{k}\sum_{i=1}^{k} R_{i,q,t} \tag{10}$$

2.4.3 CS Index

CS or Chou, Su and Lai Index was introduced in 2004 and is utilized for evaluating the validation of the automatic clustering [20]. Description of this index in general is as following:

Assume a dataset such as x_p which belongs to cluster $i\left(x_p \in c_i\right)$. The maximum distance between the members of this cluster from this data is defined as following

$$d_p^{\max} = \max_{x_q \in c_i} d(x_p, x_q) \tag{11}$$

The average of the maximum distance of the members with other members of that cluster is defined as following:

$$\bar{d}_i = \frac{1}{N_i} \sum_{x_p \in c_i} d_p^{\max} = \frac{1}{N_i} \sum_{x_p \in c_i} \max_{x_q \in c_i} d(x_p, x_q) \tag{12}$$

Which indicates the extent of the cluster and the larger is the value of this parameter, then it indicates more distance between the members. Based on this parameter, therefore, CS index is defined as following:

$$CS = \frac{\sum_i \frac{1}{N_i} \sum_{x_p \in c_i} \max_{x_q \in c_i} d(x_p, x_q)}{\sum_i \min_{j \neq i} d(m_i, m_j)} \tag{13}$$

In this index we are after minimizing CS value, which for the face should be minimized and the dominator should be maximized.

3 Main Image Retrieval Algorithms

Major image retrieval algorithms which are briefly explained in the following are utilized in this article for retrieving images.

3.1 Color Feature Extraction

Color is one of the intrinsic and obvious features of an image. This feature is resistant to changes in size, direction, noise and transparency [25]. The algorithm used in this article for extracting color feature is Color Moments. Color moments are extracted from different color spaces, however, RGB color space has a better performance in comparison with the other color spaces. The advantages of color extraction utilizing color moments include its simplicity and display of all color distributions existing in the image. Color histogram is used in this article as the extracted color feature of the image. This method is thoroughly presented in articles [26, 27].

3.2 Shape Feature Extraction

Shape is one of the fundamental features for displaying objects in images in a way that utilization of this feature can improve accuracy and efficiency of the CBIR systems. Generally there are two methods for displaying shape features: one is contour based and the other is area based [9, 10]. The methods that are utilized in order to extract the features of these two types of shape display are Fourier descriptor and fixed moments [9], respectively. In this article, Zernik moments is used for the purpose of extracting shape feature due to its simplicity of calculation and its resistance to image rotation [9, 10]. This method is thoroughly presented in articles [26, 27].

3.3 Texture Feature Extraction

Texture is one of the visual and internal features of the color that is not dependant to color and brightness [7]. This feature reflects homogeneity of the image. Texture contains the information of the image surface and also the information on the surrounding ambient in the image. The special information of the image is expressed quantitatively using the texture feature of image. Harlick et al. have defined Co-Occurance Matrix of the gray low level (Gray low-level Co-occurrence Matrix) in order to analyze the texture of image and to extract this feature which includes 14 features itself [28]. The same method is used in this article which its feature vector is comprised of 8 components and is used as texture feature [26, 29].

4 Empirical Results

4.1 Database

COREL_1K database was used for this article [18]. This database includes 1000 images that are categorized in 10 different image groups. Each group is comprised of 100 images. A sample of each image category is provided in Fig. 3. Table 1 presents the name of each of these 10 groups.

4.2 Evaluation Criterion

Precision and Recall criterion, which was first proposed by Mehtre et al. [11], is used in this study in order to evaluate each of the retrieval methods. This criterion, as addressed in most of the researches, is used as the criterion for evaluation of the performance of CBIR system [1, 5, 12, 30, 31]. Precision and Recall of the kth searched item is defined as following:

Fig. 3 Some sample images from Corel_1K database

Table 1 Group names of the image sets in Corel_1k database

Group	Group's name
1	African People
2	Beach
3	Building
4	Bus
5	Dinosaur
6	Elephant
7	Flower
8	Horse
9	Mountain
10	Food

$$precision(k) = \frac{n_k}{L} \quad \text{and} \quad recall(k) = \frac{n_k}{N_d} \tag{14}$$

where L is the number of retrieved images, n_k is the number of images related to the searched image in L retrieved images, N_d is the total number of images related to the searched image in the image database.

4.3 Results

The parameters that were used in this article concerning different algorithms are as provided in Tables 2, 3, 4 and 5.

Table 2 Parameters related to genetic algorithm

Parameter	Mutation percentage	Selection pressure	Cross over percentage	Mutation rate	Population	Maximum iteration
Value	0.3	8	0.8	0.02	100	200

Table 3 Parameters related to PSO

Parameter	Construction coefficient	Population	Maximum iteration
Value	$\varphi_1 = \varphi_2 = 2.05$	100	200

Table 4 Parameters related to ABC algorithm

Parameter	Acceleration coefficient upper band	Population number of bees	Maximum iteration
Value	1	100	200

Table 5 Parameters related to DE algorithm

Parameter	Cross over probability	Upper band of scaling factor	Lower band of scaling factor	Population	Maximum iteration
Value	0.2	0.8	0.5	50	200

This method usually requires much iteration and does not provide an appropriate response in CS Index.

This method is very fast comparing to the other methods and also provides an acceptable response for both indexes.

DB is a more appropriate index for the data of this article.

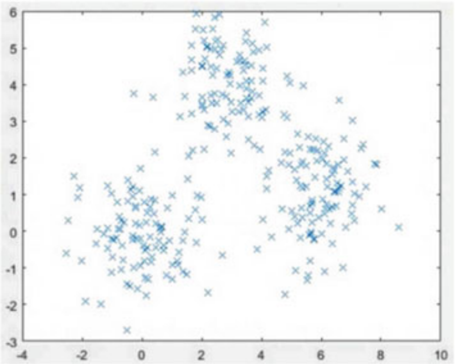
4.4 Performance Evaluation for Each of the Above Methods

For the purpose of performance evaluation for each of the above methods, we as the problem designer have produced a dataset with the following centers: $m\{1\} = [0, 0]$, $m\{2\} = [3, 4]$ and $m\{3\} = [6, 1]$, and then randomly and with distributed a number of data around each center Gaussian Normal Distribution in order to be able to better evaluate the performance of each method. The number of data around each center is 300. The set of produced data is displayed in Fig. 4.

The results obtained from applying each of the algorithms on the above dataset, for the purpose of automatic clustering with the two introduced indexes are provided hereunder (Fig. 5).

In this work we have firstly set all optimization algorithms for 10 clusters, however, the algorithms found 3 or sometimes 4 clusters considering DB and CS

Fig. 4 Produced dataset for evaluation of different algorithms



criteria. As it is evident from the above figures, DB index provided better results in comparison with CS index. In all algorithms, DB index automatically found 3 clusters, while CS index not only found the 3 clusters but also was unable to locate

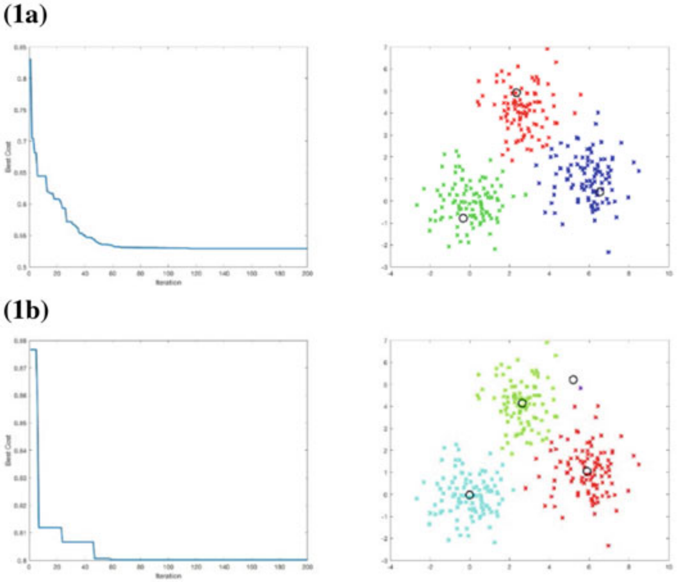


Fig. 5 Applying various evolutionary and metaheuristic algorithms on the produced dataset. **1**a Genetic Algorithm applied on the produced dataset, utilizing DB index, **1b** genetic Algorithm applied on the produced dataset, utilizing CS index. **2a** PSO Algorithm applied on the produced dataset, utilizing DB index, **2**b PSO Algorithm applied on the produced dataset, utilizing CS index. **3a** ABC Algorithm applied on the produced dataset, utilizing DB index, **3b** ABC Algorithm applied on the produced dataset, utilizing CS index. **4a** DE Algorithm applied on the produced dataset, utilizing DB index, **4b** DE Algorithm applied on the produced dataset, utilizing CS index

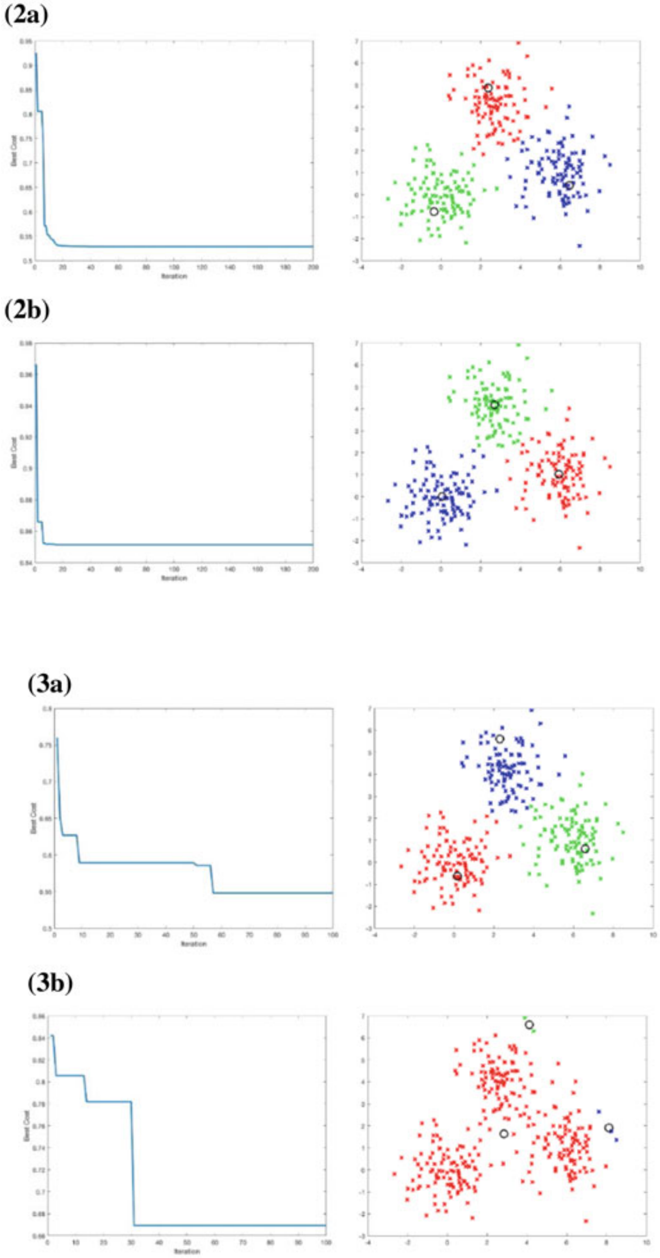


Fig. 5 (continued)

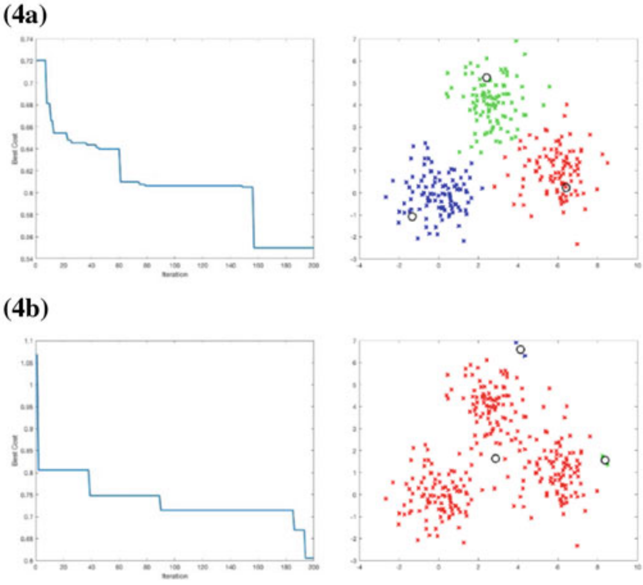


Fig. 5 (continued)

the centers of clusters correctly. Thereby it can be concluded from this stage that DB index is a more appropriate index for the dataset used in this article.

4.5 Applying Automatic Clustering for the Purpose of Image Retrieval

For this purpose, utilizing the above introduced algorithms, all images of the dataset were clustered using DB index which had a more appropriate clustering result comparing to CS index. A sample image which the above algorithms are applied on it is provided hereunder. The best image retrieval result is obtained for automatic clustering using PSO algorithm (Fig. 6).

The results obtained from this evaluation on the searched images for HSV color histogram, texture and shape features are presented in Table 6. As it is evident in the table, the results of evaluation, considering the average precisions obtained from each of the three main methods, are almost the same. Therefore, each of them can be used as the representative of the retrieval algorithm for extracting the features from the clustered images. The results obtained from retrieval using Automatic clustering with metaheuristic algorithms are demonstrated in Table 7. As you can see from this table PSO algorithm has the best result among other and the second highest retrieval is DE.

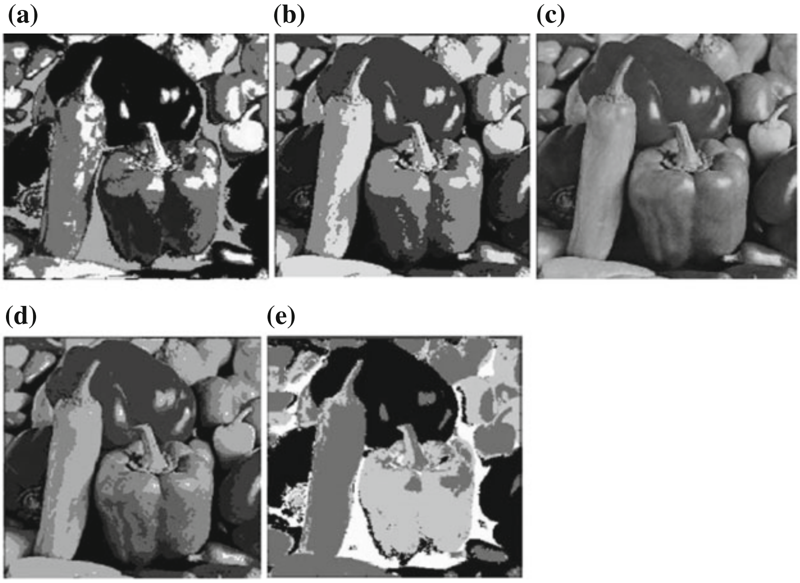


Fig. 6 **a** The original image, **b** clustering with ABC Algorithm with k = 7 clusters, **c** clustering with DE Algorithm with k = 7 clusters, **d** clustering with GA Algorithm with k = 4 clusters, **e** clustering with PSO Algorithm with k = 8 clusters

Table 6 The results of retrieval precision based on the main algorithms

Searched image	Color (HSV His.)	Texture (GLCM)	Shape (Zernike)
African People	0.754	0.798	0.415
Beach	0.598	0.415	0.598
Building	0.479	0.863	0.671
Bus	0.523	0.706	0.527
Dinosaur	0.980	0.443	0.827
Elephant	0.754	0.370	0.598
Flower	0.981	0.751	0.598
Horse	0.598	0.598	0.863
Mountain	0.598	0.598	0.618
Food	0.618	0.751	0.980
Average precision	0.688	0.629	0.669

Table 7 Results obtained from retrieval using automatic clustering with metaheuristic algorithms

Heuristic algorithms	ABC	DE	GA	PSO
Average precision	0.81	0.92	0.85	0.96

5 Conclusion

Different types of image content-based image retrieval methods which include extracting color, texture and shape features were applied on the images in this article. Furthermore, the images were automatically clustered using different metaheuristic algorithms, along with applying main algorithms on the images, in order to increase retrieval precision of all images. The results obtained from applying precision and recall criterion indicated that clustering with utilization of optimization algorithms lead to a higher average precision in comparison with solely utilizing the main image retrieval algorithms and, among them, PSO algorithm provides the highest retrieval precision.

References

1. Lin C-H, Chen H-Y, Wu Y-S (2014) Study of image retrieval and classification based on adaptive features using genetic algorithm feature selection. Expert Syst Appl 41(15):6611–6621
2. Raisi Z, Mohanna F, Rezaei M (2014) Applying content-based image retrieval techniques to provide new services for tourism industry. Int J Adv Netw
3. Furht B (ed) Encyclopedia of multimedia. Springer US, Boston
4. Hussain S, Hashmani M (2012) Image retrieval based on color and texture feature using artificial neural network. Emerg Trends …
5. Hiwale SS, Dhotre D (2015) Content-based image retrieval: Concept and current practices. In: 2015 international conference on electrical, electronics, signals, communication and optimization (EESCO), 2015, pp 1–6
6. Elshoura SM, Megherbi DB (2013) Analysis of noise sensitivity of Tchebichef and Zernike moments with application to image watermarking. J Vis Commun Image Represent 24 (5):567–578
7. Verma M, Raman B (2015) Center symmetric local binary co-occurrence pattern for texture, face and bio-medical image retrieval. J Vis Commun Image Represent 32:224–236
8. Lin C-H, Chen C-C, Lee H-L, Liao J-R (2014) Fast K-means algorithm based on a level histogram for image retrieval. Expert Syst Appl 41(7):3276–3283
9. Huang M, Shu H, Ma Y, Gong Q (2015) Content-based image retrieval technology using multi-feature fusion. Opt Int J Light Electron Opt 126(19):2144–2148
10. Singh C (2011) Improving image retrieval using combined features of Hough transform and Zernike moments. Opt Lasers Eng 49(12):1384–1396
11. Mehtre BM, Kankanhalli MS, Lee WF (1997) Shape measures for content based image retrieval: a comparison. Inf Process Manag 33(3):319–337
12. Charles YR, Ramraj R (2016) A novel local mesh color texture pattern for image retrieval system. AEU Int J Electron Commun 70(3):225–233
13. Holland JH (1975) Adaptation in natural and artificial systems. University Michigan Press, Ann Arbor
14. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of IEEE international conference neural network, pp 1942–1948
15. Storn R, Price K (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J Glob Optim 11(4):341–359
16. Bandyopadhyay S, Maulik U (2002) Genetic clustering for automatic evolution of clusters and application to image classification. Pattern Recognit 35(6):1197–1208

17. Omran M, Salman A, Engelbrecht A (2005) Dynamic clustering using particle swarm optimization with application in unsupervised image classification. In: Proceedings of 5th world Enformatika conference (ICCI), Prague, CzechRepublic
18. Karaboga Dervis (2010) Artificial bee colony algorithm. Scholarpedia 5(3):6915
19. Forgy EW (1965) Cluster analysis of multivariate data: efficiency versus interpretability of classification. Biometrics 21(3):768–769
20. Dunn JC (1974) Well separated clusters and optimal fuzzy partitions. J Cybern 4:95–104
21. Calinski RB, Harabasz J (1974) A dendrite method for cluster analysis. Commun Stat 3(1):1–27
22. Davies DL, Bouldin DW (1979) A cluster separation measure. IEEE Trans Pattern Anal Mach Intell 1(2):224–227
23. Pakhira MK, Bandyopadhyay S, Maulik U (2004) Validity index for crisp and fuzzy clusters. Pattern Recognit Lett 37(3):487–501
24. Chou CH, Su MC, Lai E (2004) A new cluster validity measure and its application to image compression. Pattern Anal Appl 7(2):205–220
25. Raisi Z, Mohanna F, Rezaei M (2011) Content-based image retrieval for tourism application using handheld devices. IJICTR J
26. Raisi Z, Azarakhsh J Content based image retrieval for marine life images using ant colony optimization feature selection
27. Raisi Z, Azarakhsh J (2016) Feature selection based-on swarm particle optimization and genetic algorithms for image retrieval. Int J Adv Biotechnol Res (IJBR) 7 Special Issue-Number 5:907–916
28. Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Trans Syst Man Cybern 3(6):610–621
29. Seryasat OR, Haddadnia J, Ghayoumi-Zadeh H (2015) A new method to classify breast cancer tumors and their fractionation, Ciência e Nat 37: 51–57. Assessment of a novel computer aided mass diagnosis system in mammograms
30. Rui Y, Huang TS, Chang S-F (1999) Image retrieval: current techniques, promising directions, and open issues. J Vis Commun Image Represent 10(1):39–62
31. Dimitrovski I, Kocev D, Loskovska S, Deroski S (2016) Improving bag-of-visual-words image retrieval with predictive clustering trees. Inf Sci (Ny) 329:851–865
32. Liu G-H, Yang J-Y (2008) Image retrieval based on the texton co-occurrence matrix. Pattern Recognit 41(12):3521–3527
33. Liu GH, Zhang L, Hou YK, Li ZY, Yang JY (2010) Image retrieval based on multi-texton histogram. Pattern Recognit 43(7):2380–2389

A Robust Blind Audio Watermarking Scheme Based on DCT-DWT-SVD

Azadeh Rezaei and Mehdi Khalili

Abstract In this paper, Watermarking hybrid sound algorithm is presented to protect the copyright of audio files, which, in addition to the clarity and consistency of the audio signal, has increased the strength and strength. To this end, a hybrid algorithm for voice signal cryptography is presented in the three domain parser transforms, discrete cosine transform, and discrete wavelet transforms. So, after discrete cosine transformation (DCT) on the host signal, by selecting the sub-band of low frequency, which contains the highest signal energy, two discrete wavelet transform (DWT) with a random wavelet filter on the low-frequency coefficients of conversion A discrete cosine applies, after selecting the approximation coefficients, the resulting one-dimensional matrix is converted to a two-dimensional matrix, and finally the resulting matrix is applied to a single value decomposition (SVD), which results in the formation of a The diameter matrix is that the watermark bits are embedded in the first layer of the dipole matrix, so that the two bits with value S (1,1), S (2,2) of the matrix The diameter S is chosen, first compares the first and second intersections of the diameter matrix S, which is multiplied by the coefficient θ multiplied by the obtained two bits and is used as a fixed value in the embedding formula and The title of the new watermark is embedded in S (1,1). The results of the implementation show that the proposed algorithm succeeded not only in achieving transparency and resistance to general audio processing attacks, such as Gaussian white noise, quantization rates, decreasing and increasing the rate of sampling, compression and low pass filtering. But has achieved better results than other similar algorithms.

Keywords Audio cryptography · Discrete wavelet transform · Discrete cosine transform · Single value decomposition

A. Rezaei (✉) · M. Khalili
Faculty of Computer Engineering, Payame Noor University, Tehran, Iran
e-mail: azadehrezaei1369@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_8

1 Introduction

Recent advances in the Internet and digital multimedia products technology have made it possible for digital signals (audio, video, and video) to be easily distributed to different regions. This ease in Transmission, allows unauthorized copies of multimedia products to be distributed and distributed. For this reason, protecting the right to digital rights has become an important topic in the world [1]. Digital Watermark has given a lot of attention to solving this problem [2]. Information storage is a way to make information. In the form of an overlapping agent with the highest degree of security precaution, it moves between the points in question so that even if the information is accessed through unauthorized persons along the route, there is no access to the hidden data [3]. In fact, the lack of a focus on art and science is the embedding of information in a carrier medium, which is becoming increasingly widespread due to the significant advancements in digital communications [4]. In mainstream preservation, security means inability to prove the existence of a message [5]. Digital media are more popular than analogue media, and unlike them, they can be easily stored, reproduced and distributed. Thus, unauthorized reproduction of digital documents, such as audio signals, has been raised as a concern in recent years [6]. Therefore, digital sounds calling is used to protect copyright and prove ownership [7]. In this regard, several methods have been proposed in different areas of frequency and time, in which the algorithms presented in the frequency domain are more resistant [8]. The most significant transformations in the transformation area used in watermarking algorithms are the discrete cosine transform, the discrete wavelet transform, and the conversion of the parsing of unique values [9]. Discrete cosine transforms are resistant to image processing operations such as low pass filter, contrast, brightness adjustment, etc. [10]. A discrete wavelet transform is presented as an alternative to the Fourier transform of short time, and its goal is to overcome resolution problems in Fourier transforms of short time [11]. Although DCT has a combination of features of the human vision system, wavelet transformation is closer to the human vision system than DCT, and in general in some applications it is better than DCT-based methods [10]. The conversion of the unique value fragmentation is an efficient and effective tool in numerical analysis, a clear feature of this transformation of stability against attacks such as rotation and noise [11]. Considering the benefits of each of the conversion areas, various studies have been done on the combination of these conversions on Watermarking, which is, of course, much of the research done in this regard, the use of watermarking of digital images in audio signals is most commonly used. Compared to Watermarking, Video and Watermarking video are more susceptible to two major causes. First, the contents of the audio signals are one-dimensional data, so it's very difficult to add hidden data without harming the quality of the audio signal. Secondly, the human hearing system (HAS) is more sensitive to the human system of vision, so the control of degradation in quality is detected by the listener. In practical applications categories, the soundtrack method must certainly meet a number of needs [12]. The three important requirements that come with the Watermarking Voice are: Inaudible, Resistance and Security [13].

Cai et al. [14] in order to protect the copyright of digital audio and video copyright in the network, a blind audio algorithm is proposed using digital discrete wavelet transform (DWT) and the conversion of unique values (SVDs). In this algorithm, an original audio signal is segmented into blocks with a length of 1024, and each block blocks the wavelet transform into two levels. Then, the approximate coefficients of the near-correct sound of the initial sound are decomposed into the SVD conversion, and the resulting matrix of transformation is obtained. Watermarked information is embedded in a diagonal matrix. The experiments show that the proposed algorithm has a transparency of 20.7000 and its robustness against common audio signal attacks, such as re-sampling, low pass filtering, quantization, Gaussian noise, MP3 compression, on average, with a correlation coefficient of more than 0.940 and an average. The error rate is less than 0.060. The non-attack extracted signal has a mean normal correlation coefficient of 1 and an average error rate of 0 that indicates that the watermark can be extracted explicitly and without blindness without any attack. In [15], a strong blind sound market with DCT-DWT-SVD is presented. In this article, the original sound is first divided into frames of length 4096. Then DCT conversion is applied to each frame at first. The low frequency coefficients derived from the DCT conversion of the five levels of the DWT conversion are applied to obtain the approximation coefficients. Eventually, the SVD conversion is applied to obtain the three matrices S, U, V, the 32-bit binary watermark image in the matrix Diagonal S is embedded. Public sounding attacks such as sampling, low pass filtering, quantization, Gaussian white noise, and MP3 compression have been applied, the results show that the average correlation coefficient is normal at about 0.980 and the average error rate is less than 0.058. In this paper, the strength of the work has been investigated, but the transparency has not been tested and the extraction is non-blind. In the proposed paper, a method is suggested that, in addition to achieving the characteristics of Watermarking, it improves the transparency component and the resistance of the audio signal, so that by combining three discrete wavelet transforms, a discrete cosine transform, and the decomposition of single values into an algorithmic representation in To prove the right to own digital audio.

2 Background

2.1 Discrete Cosine Transform (DCT)

A discrete cosine transform is a technique for returning the signal within the elemental frequency components [16]. A discrete cosine transform transforms a signal into three sub-band frequencies down, middle, and up. Most of the input signal is accumulated in low frequency components, called DCs [17].

The relation of the discrete cosine transformation is in the form of relation (1):

$$X[K] = \mathrm{w[K]} + \sum_{n=0}^{N-1} \mathrm{X[n]}\left(\cos\frac{(2n+1)K\pi}{2N}\right), \quad 0 \ll \mathrm{K} \ll \mathrm{N} \tag{1}$$

So that

$$w[k] = \begin{cases} \sqrt{\frac{1}{N}}, & k = 0 \\ \sqrt{\frac{2}{N}}, & 0 < k \leq N-1 \end{cases}$$

2.2 Discrete Wavelet Transform (DWT)

The wavelet transform method is that with two high pass and low pass filters, the audio signal is divided into two sections: high frequency, detail (CD) and low frequency, approximation (CA), the number of samples in each of these sections Half the number of samples in the main signal [18]. The transformation of a discrete wavelet acts as follows:

$$\varphi_{jk}(t) = \frac{1}{\sqrt{a_0^j}}\varphi_{jk}\left[\frac{\frac{t-ka_0^jb_0}{j}}{a_0}\right] \tag{2}$$

$$Wnf(j,k) = \langle f(j,k), \varphi_{jk}(t)\rangle \tag{3}$$

$$= \frac{1}{\mathrm{C}_\varphi}\int_{0}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{W_f}(\partial, b)\varphi_{\partial,b}(t)db\frac{d\partial}{|\partial|^2} \tag{4}$$

2.3 Singular Value Decomposition (SVD)

This conversion is an efficient and effective tool in numerical analysis. This conversion is due to the application of the matrix topic and the fact that a digital image is also a matrix can be effective in the process of watermarking [19]. Using this transformation on the desired matrix A Dimensions m * n are obtained by three matrices U, V, S [11] so that A = U * S * V and

$$S = diag(\sigma 1, \ldots \sigma i, \ldots \sigma r), \sigma i > 0(i = 1, \ldots, r)r = rank(\mathrm{A}) \tag{5}$$

3 The Proposed Algorithm

In order to protect the proposed algorithm, a binary image is first embedded as a watermark in an audio signal. First, the watermark image of the gray will be converted to a binary image before the embedding of the watermark, then the embedding and extraction process is described. The block diagram of embedding and extraction of watermarks is shown in Figs. 1 and 2 respectively.

3.1 Watermarking Process

1. Framing the original audio signal
2. Convert DCT to any frame
3. Performing two levels of DWT conversion on the low frequency coefficients derived from the DCT transform with a random wavelet filter for obtaining approximation coefficients (second level transformation on the approximation coefficient obtained from the first level transformation).
4. Conversion of the one-dimensional matrix obtained from the conversion of the coefficients of the DWT approximation to a two-dimensional matrix
5. Run SVD transform on a two-dimensional matrix according to the relationship and obtain three matrices U, S, V

$$R = U\,S\,V \tag{6}$$

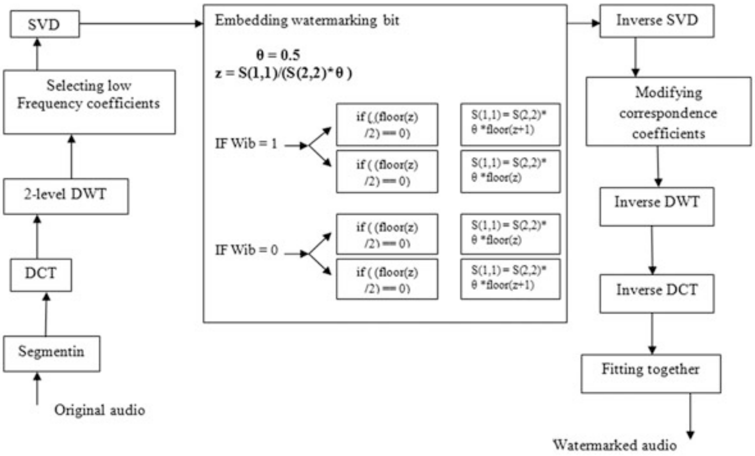
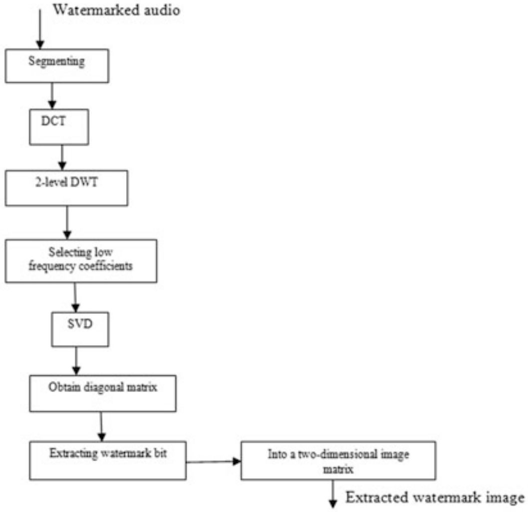


Fig. 1 Block diagram of watermark image embedding algorithm in domain DCT-DWT-SVD

Fig. 2 Block diagram for watermarked extraction of image-sound algorithm in domain DCT-DWT-SVD


Watermarked audio
Segmenting
DCT
2-level DWT
Selecting low frequency coefficients
SVD
Obtain diagonal matrix
Extracting watermark bit
Into a two-dimensional image matrix
Extracted watermark image

6. Two bits of value S (1,1), S (2,2) are selected from the diameter matrix S and embedded in them according to the conditions of the watermark. We first consider the comparison of the first and second intersections of the diagonal matrix S: if the remainder of the subdivision of the sub frame, including the integer S (1,1), on the product of S (2,2) in the defined coefficient, is 0.5 = θ, we consider the constant z, then If this value is a pair wise integer and a watermark bit, then it multiplies the individual value of z (i.e., z + 1) in the product of S (2,2) in the coefficient θ, and in the bit S (1,1) We install as new watermarked. Now, if the bit watermark is 0, then the z value unchanged is multiplied as a pair wise value in S (2,2) and θ, and we embed it as a new watermarked value in S (1,1).

$$\begin{aligned} &\theta = 0.5 \\ &\mathrm{z} = \mathrm{S}(1{,}1)/(S(2{,}2) * \theta) \end{aligned} \tag{7}$$

$$\begin{aligned} &\mathrm{if}((\mathrm{floor}(\mathrm{z})/2) == 0) \quad \text{is even} \;\&\; \text{watermark image bit} = 1 \\ &\qquad \mathrm{S}(1{,}1) = \mathrm{S}(2{,}2) * \theta * \mathrm{floor}(\mathrm{z}+1)) \end{aligned} \tag{8}$$

$$\begin{aligned} &\mathrm{if}((\mathrm{floor}(\mathrm{z})/2) == 0 \quad \text{is odd} \;\&\; \text{watermark image bit} = 1 \\ &\qquad \mathrm{S}(1,1) = \mathrm{S}(2{,}2) * \theta * \mathrm{floor}(\mathrm{z})) \end{aligned} \tag{9}$$

Now if z is a certain number and the watermark bit 1 then multiplies the pair of z, (z + 1) in the product of S (2,2) in the coefficient θ, and in S (1,1) to Insert the title of the new watermark. Now, if the bit watermark is equal to 0, then the z value unchanged as an individual value in S (2,2) and θ is multiplied and embedded as a new watermarked value in S (1,1).

$$\begin{aligned}&\mathrm{if}((\mathrm{floor}(z)/2) == 0 \quad \text{is even} \,\&\, \text{watermark image bit} = 0\\&\quad \mathrm{S}(1{,}1) = \mathrm{S}(2{,}2) * \theta * \mathrm{floor}(z))\end{aligned} \tag{10}$$

$$\begin{aligned}&\mathrm{if}((\mathrm{floor}(z)/2) == 0 \quad \text{is odd} \,\&\, \text{watermark image bit} = 0\\&\quad \mathrm{S}(1{,}1) = \mathrm{S}(2{,}2) * \theta * \mathrm{floor}(z+1))\end{aligned} \tag{11}$$

In this algorithm, the bits of diameter matrix S (1,1) with the proposed solution contain watermark bits.

7. Reversal SVD conversion on watermarked matrix.
8. Convert a two-dimensional matrix to a one-dimensional matrix.
9. Inverse the two levels of DWT conversion on a one-dimensional matrix and obtaining approximation coefficients.
10. Reverse DCT conversion on approximation coefficients.
11. Finally, the addition of all the original audio signal frames and the watermark to calculate the watermarked audio signal.

3.2 Watermark Extraction Process

1. Framing the watermarked audio signal.
2. Performing DCT transformation on the watermark signal and obtaining low frequency coefficients.
3. Implementing two levels of DWT conversion on the low frequency coefficients of the DCT conversion and obtaining approximation coefficients (second level transformation on the approximation coefficient obtained from the first level transformation).
4. Conversion of a one-dimensional matrix to a two-dimensional matrix

$$\mathrm{R_S} = \mathrm{U_S}\,\mathrm{S_S}\,\mathrm{V_S} \tag{12}$$

5. Run SVD conversion on a two-dimensional matrix.
6. Isolation of watermark bits without the need for the original audio signal by the blind method.
7. Finally, sorting the extracted watermark bit in a two-dimensional matrix to extract the Watermark image.

4 Results of Implementation of the Proposed Method

In the proposed algorithm, first, a 32-bit binary image with jpg format (Fig. 3) as a watermark in a 16-bit audio signal with a sampling rate of 44,100 kHz with a size of 2,395,137 * 1 and a wav format (Fig. 4) is embedded. The main watermark

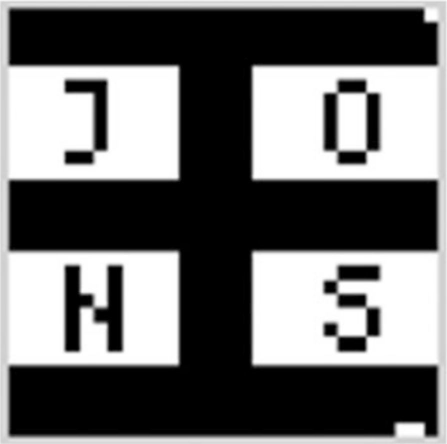


Fig. 3 Main watermark image

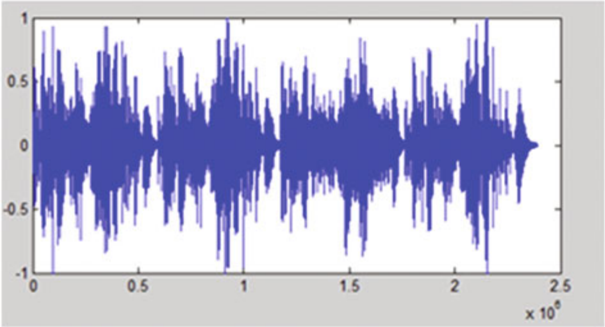
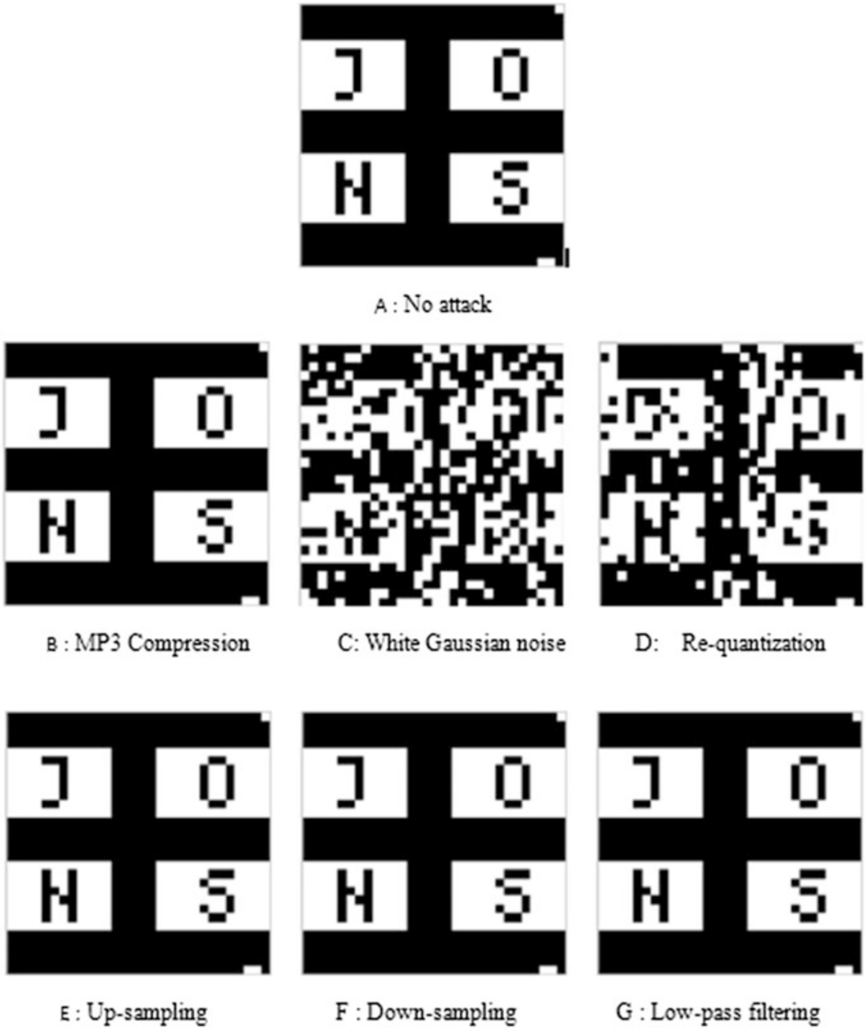


Fig. 4 Main sound signal

image and the images attacked by various attacks, including Gaussian white noise, the rate of quantization, decreasing and increasing the sampling rate, compression and low pass filter, as well as the original sound signal and the watermarked sound signal are also shown in Fig. 5. It can be seen. The test results are based on the following criteria: SNR, which represents the signal-to-noise ratio. This scale indicates the amount of noise added to the sound due to the insertion of a mark in it. The BER, which is the bit error rate between the two image images extracted with the original sign, and the NCC, which represents the normalized cross-correlation of the two signs, has been reported. The SNR value is given in Table 1, and the BER and NCC values obtained from this algorithm are shown in various tables, respectively, in Tables 2 and 3. The comparison chart for BER and NCC values is also given in Figs. 6 and 7, respectively.


A : No attack

B : MP3 Compression

C: White Gaussian noise

D: Re-quantization

E : Up-sampling

F : Down-sampling

G : Low-pass filtering

Fig. 5 Watermark image with SNR = 35.2469 with various image processing attacks

Table 1 Comparison of the proposed SNR of the proposed algorithm with similar tasks

Proposed watermarking algorithm	Method	
	Cai et al. [14]	Lei et al. [15]
35.2462	20.7000	…

Table 2 Comparison of the BER Factor of the proposed algorithm with similar tasks

Attack	Method		
	Cai et al. [14]	Lei et al. [15]	Proposed watermarking algorithm
No attack	0.0000	0.0000	0.0000
MP3 compression	0.0000	0.0000	0.0000
White Gaussian noise	0.1523	…	0.1474
Re-quantization	0.2676	0.1791	0.1033
Up-sampling	0.0000	0.0586	0.0000
Down-sampling	0.0000	…	0.0000
Low-pass filtering	0.0059	0.0586	0.0000

Table 3 Comparison of the NCC factor of the proposed algorithm with similar tasks

Attack	Method		
	Cai et al. [14]	Lei et al. [15]	Proposed watermarking algorithm
No attack	1.0000	1.0000	1.0000
MP3 compression	1.0000	1.0000	1.0000
White Gaussian noise	0.8603	…	0.8723
Re-quantization	0.7311	0.9206	0.9875
Up-sampling	1.0000	1.0000	1.0000
Down-sampling	1.0000	…	1.0000
Low-pass filtering	0.9999	0.9440	1.0000

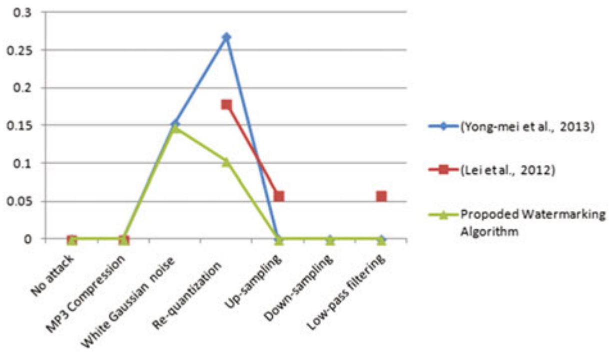


Fig. 6 Comparison of the BER Factor with the proposed method

$$SNR = 10\log_{10}\frac{\sum_{n=1}^{N} Y^2(n)}{\sum_{n=1}^{N}\left[Y(n) - Y'(n)\right]^2} \tag{13}$$

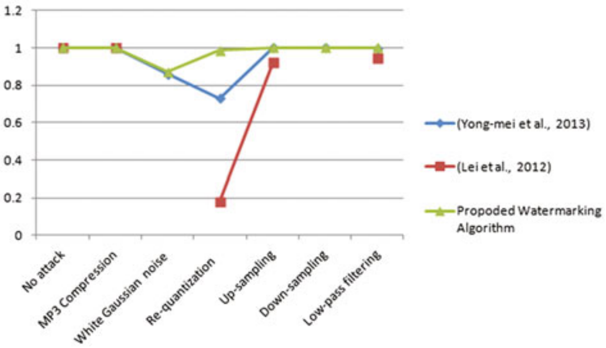


Fig. 7 Comparison of the NCC factor of the proposed method with similar tasks

$$BER = \left\{ \sum_{i} \sum_{j} \frac{\{W(i,j) \oplus EW(i,j)\}}{M \times N} \right\} \times 100\%. \tag{14}$$

$$NC(w, \hat{w}) = \frac{\sum_{i=1}^{Nw} w(i).\widehat{w}(i)}{\sum_{i=1}^{Nw} w^{2}(i)} \tag{15}$$

5 Conclusion

This paper is devoted to the presentation of a hybrid algorithm based on discrete cosine transformation, discrete wavelet transform and parsing single values for voice signal cryptography. The proposed algorithm has been able to prove the right to ownership of digital audio and video. On the other hand, the comparison of the proposed algorithm and the evaluation of their performance and resiliency and their transparency against various attacks has led to a rather comprehensive research on the methods of watermarking for digital audio signals.

In order to use the benefits of different domains simultaneously, a hybrid cognitive algorithm has been proposed in three discrete cosine transformation areas, discrete wavelet transform and single particle breakdowns. In the proposed cognitive auditing approach, a DCT operation layer is initially on the host audio signal is applied to produce high, low, and average frequencies, then on a low frequency coefficient, which contains the highest energy, a DWT conversion level is applied with a randomized wavelet filter to yield the frequency bands CA, CD. Then, on the sub-band CA, which is called the sub band of approximation and follows the initial shape of the signal, once again the DWT conversion is applied to the randomized wavelet filter, we transform the resulting one-dimensional matrix into a two-dimensional matrix, It converts a single value parsing (SVD), so that the two bits with S (1,1), S (2,2) are selected from the S-matrix matrix, first compare the first and second intersections of the diameter matrix S And embedded in the formula for the new watermarked value in S (1,1). Then the ISVD is converted and the two-dimensional matrix is converted to a one-dimensional matrix. By performing two conversion levels, the modified IDWT blocks are redone to the modified sub-band of the edited host audio, and then a conversion level of the IDCT is applied. Has been finally, to create a watermark, all the original audio frames and watermark bits are merged.

The watermarking process is also reversed by the embedding process, so that the host signal is first blocked and a DCT transformation level is applied to produce low-frequency coefficients and the DWT conversion factors are carried out with a randomized wavelet filter. And low frequency coefficients are approximated. Again, these approximation coefficients apply to a transformation DWT, and the one-dimensional matrix is derived from the approximation coefficients to a two-dimensional matrix, transforming the SVD to obtain a diagonal matrix The two-dimensional matrix is applied and the watermark bits are extracted without the need for the host's sound signal. The results indicate that the algorithm has a satisfactory performance in achieving the desired goals and has performed better than similar research.

References

1. Thanki R, Borisagar K, Borra S (2018) Speech Watermarking technique using the finite Ridgelet Transform, discrete wavelet transform, and singular value decomposition. Advance compression and watermarking technique for speech signals, pp 27–45
2. Zhang J, Han B (2017) Robust audio watermarking algorithm based on moving average and DCT. KYTZ201420 and scientific research project of department of education in Sichuan province under Grant, No. 16ZA0221
3. Ojha S, Sharma A, Chaturvedi R (2018) Centric-oriented novel image watermarking technique based on DWT and SVD. Soft computing: theories and applications. Advances in intelligent systems and computing, vol 583. Springer, Singapore
4. Mishra S, Yadav VK, Trivedi MC, Shrimali T (2018) Audio steganography techniques: a survey, advances in intelligent systems and computing, vol 554. Springer, Singapore
5. Seok J, Hong J, Kim J (2017) A novel audio watermarking algorithm for copyright protection of digital audio. ETRI J 24(3):181–189
6. Singh S, Singh R, Singh AK, Siddiqui TJ (2018) SVD-DCT based medical image watermarking in NSCT domain. Quantum computing: an environment for intelligent large scale real application. Studies in big data, vol 33. Springer, Cham
7. Kavitha KJ, Shan BP (2018) Video watermarking using DCT and DWT, a comparison. Eur J Adv Eng Technol 2(6):83–87
8. Saxena N, Mishra KK, Tripathi A (2018) DWT-SVD-based color image watermarking using dynamic-PSO. In: Advances in intelligent systems and computing, vol 554. Springer, Singapore
9. Tagesse Takore T, Rajesh Kumar P, Lavanya Devi G (2018) A robust and oblivious grayscale image watermarking scheme based on edge detection, SVD, and GA. In: Proceedings of 2nd international conference on micro-electronics, electromagnetics and telecommunications. Lecture notes in electrical engineering, vol 434. Springer, Singapore

10. Verma C, Tarar S (2018) Secure random sequence based frequency hoping spread spectrum audio watermarking. Int J Eng Sci Comput 6(5)
11. Jain R, Trivedi MC, Tiwari S (2018) Impact analysis of contributing parameters in audio watermarking using DWT and SVD. In: Advances in intelligent systems and computing, vol 554. Springer, Singapore
12. Gosavi CS, Mali SN (2017) Watermarking for Video using single channel block based schur decomposition. Global J Pure Appl Math 12(2):1575–1585. ISSN 0973-1768
13. Pohan N, Saragih R, Rahim R (2017) Invisible watermarking audio digital with discrete cosine transform. Sci Technol IJSRST 3(1). Print ISSN: 2395-6011, Online ISSN: 2395-602X
14. Cai Y-m, Guo W-q, Ding H-y (2013) An audio blind watermarking scheme based on DWT-SVD. J Softw 8(7)
15. Lei M, Yang Y, Guo Y, Mao J, Luo Q (2012) A robust blind audio watermarking scheme based on DCT-DWT-SVD. Int J Dig Content Technol Appl (JDCTA) 6. https://doi.org/10.4156/jdcta.vol6.issue21.29
16. Kakkirala KR, Chalamala SR, Rao G (2017) DWT-SVD based blind audio watermarking scheme for copyright protection. In: Audio, language and image processing (ICALIP)
17. Padungdit A (2018) Image watermarking using joined wavelet and time domain. In: ICT international conference, No. 3, pp 47–50
18. Csrvajal G (2017) Scaling factor for RGB image to steganography application, IEC comprehensive report on information security international engineering consortium. J Vec Rel 3:55–56
19. Darabkh KA (2018) Imperceptible and robust DWT-SVD-based digital audio watermarking algorithm. J Softw Eng Appl 7:859–871

A New Method to Copy-Move Forgery Detection in Digital Images Using Gabor Filter

Mostafa Mokhtari Ardakan, Masoud Yerokh
and Mostafa Akhavan Saffar

Abstract Copy-move forgery is one of the types of image manipulation which is widely used due to simplicity and effectiveness. In this method, part of the original image is copied and pasted to the desired location in the same image. The goal of detecting copy-move forgery is to find areas of the image that are identical or very similar. One of the important issues that some of the earlier algorithms suffer from is that the forged area is rotated or resized after attachment. In this research, a new approach is presented to detect copy-move forgery in digital images based on discrete wavelet decomposition along with multiple features extracted by Gabor filter to improve the function of detecting similar areas of the image. Experiments have shown that this algorithm recognizes similar areas with relatively good accuracy and is resistant to rotation and change in the scale of the forged area.

Keywords Detection of forgery · Copy-move forgery · Discrete wavelet transform
Gabor filter · Feature matrix

1 Introduction

Image forgery or manipulation has a long history. In today's digital world, it's easy to create, modify, and correct information provided by the image (without leaving any obvious traces of this operation) [1]. Image forgery can be done in different ways and for different purposes. An old sample of forged image is the following Fig. 1.

M. Mokhtari Ardakan (✉) · M. Yerokh · M. Akhavan Saffar
Department of Computer and Information Technology, Faculty of Engineering,
Payame Noor University, Tehran, Islamic Republic of Iran
e-mail: mostafamokhtari@pnu.ac.ir

M. Yerokh
e-mail: masoud_yerokh@yahoo.com

M. Akhavan Saffar
e-mail: akhavansaffar@pnu.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_9

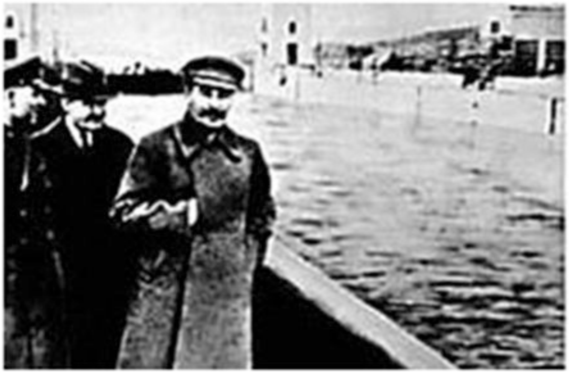


Fig. 1 Removing Nikolai Yezhov's picture

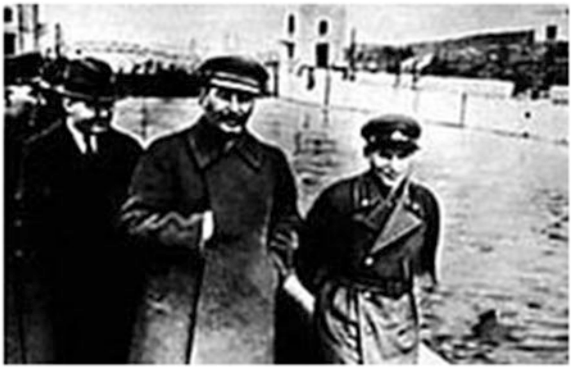


Fig. 2 The original image before forging

In this picture, the image of Nikolai Yezhov, one of the closest advisers of Joseph Stalin, the General Secretary of the Communist Party of the Soviet Union's Central Committee was removed from Stalin's photo after being jailed for corruption. The original image before the forging can be seen in Fig. 2.

Of the latest examples of image forgery, is Fig. 3. After the speech by Mr. Hassan Rouhani, President of Iran at the seventieth meeting of the UN General Assembly in New York, Foreign Minister Mohammad Javad Zarif, who was leaving the Assembly Hall, occasionally faced with President Barack Obama and Secretary of State John Kerry at the entry to the General Assembly and shook hands with them. After the publication of news, an image was published on social networks that claimed to be the photo of the moment that Javad Zarif and Barack Obama were shaking hands. A little care in watching the image shows that the image of Obama shaking hands with Zarif is manipulated in photoshop and it is fake. Studies also show that the original image is related to the visit of President Cavillion Raúl Castro and Barack Obama (Fig. 4).

The purpose of detecting forged image is the authentication of a digital image. Authentication solution is classified into two types:

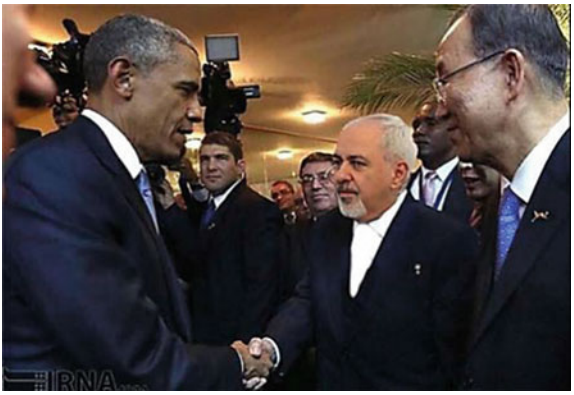


Fig. 3 A forged image published showing the moment of Zarif's meeting with Barack Obama

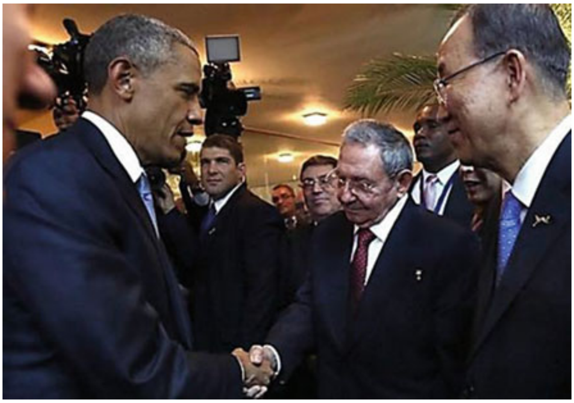


Fig. 4 Cuban President Raúl Castro's and Barack Obama

(1) active and
(2) passive or blind.

Active forgery detection techniques (such as digital watermarking or digital signatures) utilize a well-known authentication code embedded in the image content; the authentication process may be proven through the verification of the existence of such an authentication code (by comparing with the original code inserted). In addition, this method requires specific hardware or software to add an authentication code into the image (before the image is published) [2].

Blind or passive forgery detection technique uses the received images only to assess the completeness or accuracy of the images. This method is based on the assumption that while digital forgery measures may leave no visual clues of a distorted image, but most likely they distort the statistic features or image integrity compared to the normal structure of the image, resulting in new adverse effects (leading to various forms of mismatch). This mismatch can be used to identify forgery. Since this technique does not require any former information about the image, it is a commonly used technique. Existing techniques determine types of traces of manipulation and identify them (separately) by positioning the distorted areas.

2 Copy-Move Forgery (Or Area Copy Forgery)

Copy-move forgery is one of the most common techniques of image distortion, which is used due to its simplicity and effectiveness. In this method, part of the original image is copied and moved to another part in the same image and it is pasted there. This is done in order to hide particular details of the image or reproduce special effects in it. Because the uneven areas of the image have similar properties of color and noise fluctuations (which is imperceptible to the human eye in search of inconsistencies within the statistical properties of the image,) the region is used as the ideal part for cop-move forgery. Usually, fading operations (along the boundary edge of the modified area) are used to reduce the effect of disturbances between the main area and the pasted area [1]. Figure 5 presents examples of this type of forgery.

Copy-move forgery detection methods can be divided into two general categories:

1. Methods based on blocking
2. Non-block method

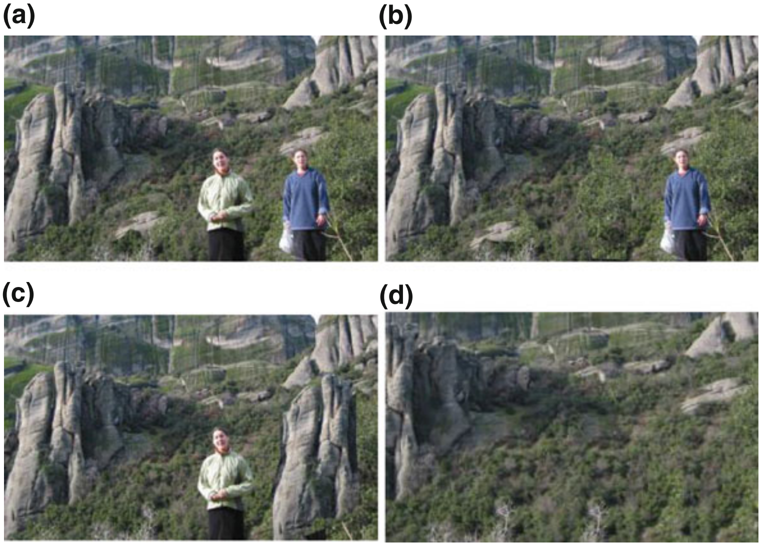


Fig. 5 **a** is the original image; **b**, **c** and **d** are forged images

Detection methods based on blocking

Most blocking methods follow a six-step process according to graph in Fig. 6.

Before the feature extraction process, a series of operations, such as image sorting, conversion of RGB images to black and white images or YCBCR conversion and the use of certain channels of the obtained images, the use of DWT or DCT conversion in order to reduce the size and improve the efficiency of classification, can be enforced on the desired images. To avoid the high computational cost of detailed search of image, comparison is made at the block level. The blocks used for comparison can be square or circle. Of course, square block use is more common [3]. If the image f(x, y) with a size of M × N pixels, and blocks with a size of b × b pixels are considered for comparison, then each block must be compared to the other blocks overlapping in the image by $(M - b + 1) \times (N - b + 1)$. Figure 7 shows the use of the two methods of blocking [4].

Accuracy, speed and complexity of forgery detection algorithm depends heavily on the ability to extract and identify similar features. Different extraction methods have been proposed for the extraction of features, most of which can be summarized in three methods: wavelet [4–7], location [8, 9] and frequency [10–12]. Some of these methods, such as methods that have been proposed in wavelet and frequency, have a good accuracy but are difficult in terms of time complexity, on the other hand, only part of these methods are resistant to factors such as Gaussian flattening and rotation. After the feature is extracted, potential copy-move pairs are identified by searching for similar feature blocks. Extracted features are initially arranged as

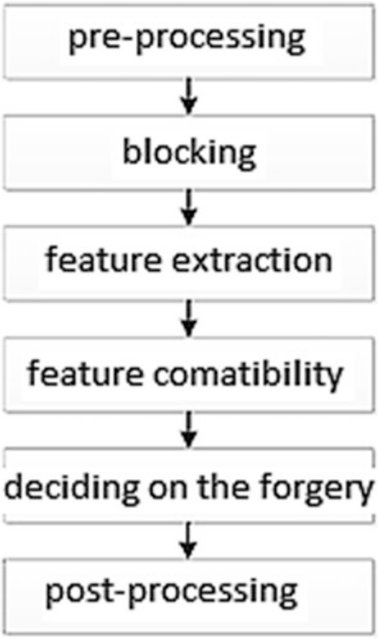


Fig. 6 Forgery detection process based on blocks

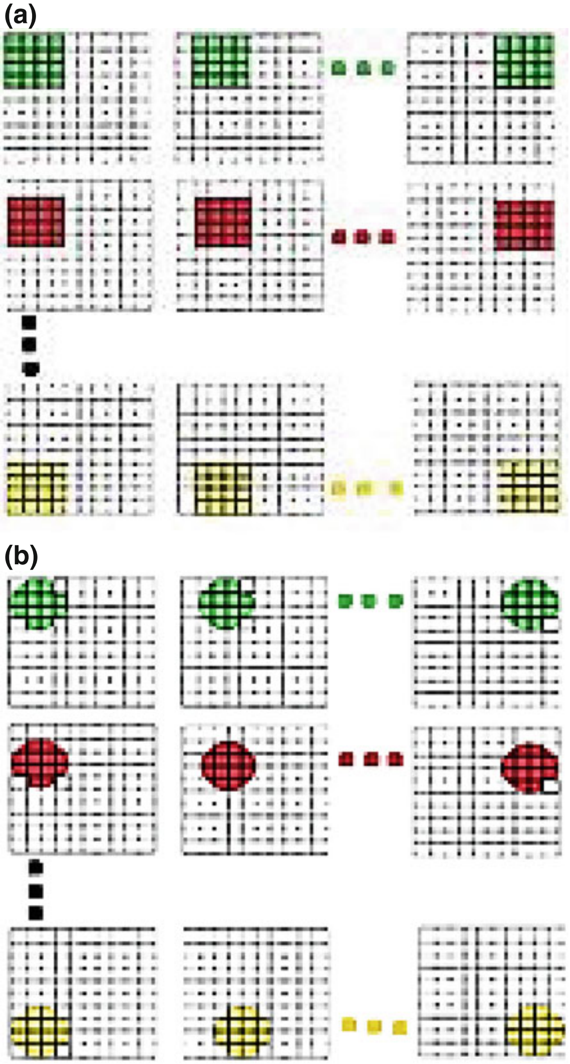


Fig. 7 Types of blocking: **a** square, **b** circular

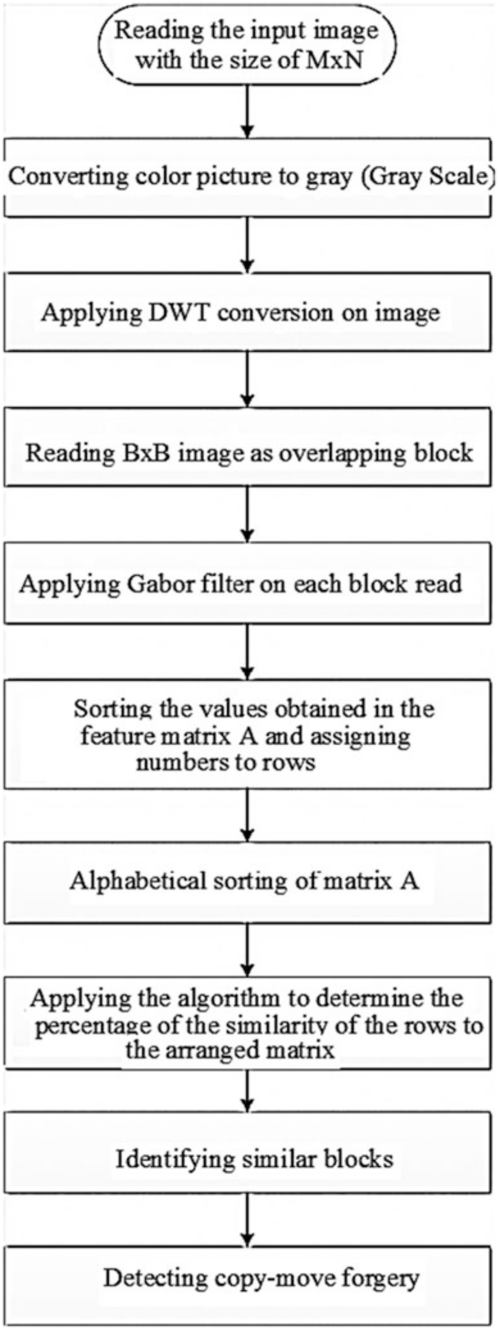
M-matrix rows, then using trivial approach, each feature is compared with all the other features, but this approach is expensive in terms of computation time. To cope with this challenge, there are many ways to set similar features close together, which prevents useless comparisons and reduces computation time. In fact, each feature will be compared only to a certain number of neighbors. Among the known methods, the most common method is "lexicographic sorting" which uses "radix sorting" to create a matrix with the same features in the neighborhood, and thus make them easier to detect.

In addition to lexicographic sorting, base sorting, sorting by the number of zeros, k-dimensional tree sorting [13], a combination of “lexicographic sorting” and “k-d tree” which is used to improve the time complexity and accuracy in the process of matching the features, Bloom filters counting, sorting based on vector components with the highest variance among all features, comparing the hash values, block linking and block clustering could also be named. As soon as the data are organized to reduce the complexity of the investigation of similarities, search for similar features using various “similarity terms” is done, some of which can be cited as follows: Euclidean distance with the size of S = 1/(1 + dis) where “dis” is the distance measured in Euclidean space; Hamming distance, Hausdorff distance, logical distance, the correlation coefficient, the phase coefficient, cross-spectrum normalized, local sensitive hashing and ratio of absolute error. In the decision-making process on forgery, one can state that, almost always, a single similarity criterion is not enough to decide on the presence or absence of duplicated space. This is due to the fact that most natural images may contain one or more pairs of very similar regions; so, wrong matches may be resulted. Therefore, it is required to identify copy-move features of the areas to distinguish them from false matches. Sometimes, the map of the duplicated areas obtained from the previous step require more processing. Along with the rest (of the methods), post-processing can be performed by methods such as morphological post-processing including opening operations, erosion, dilatation, sliding window, square kernel mean filter and random sample consensus algorithm (RANSAC) which recognize the inliers and eliminate the outliers [3].

3 Introducing the New Method

In this research, a new method is presented for identifying areas of the image that are identical or very similar. The methodology is one of the methods of blind detection based on blocking. The proposed method is shown in Fig. 8. Myna et al. [4], presented a wavelet-based approach in which the use of wavelet transform in detection of copy-move forgery was tested. In the second stage, stored blocks are repeatedly compared in each level of the wavelet transform. Finally, the last match is done on the image. This approach functions properly when the copied area is changed by scaling and rotation. In their method, to resist against the change of scale and rotation, polar logarithmic transformation is used which is a change from the Cartesian to polar coordinates. In the new method in the present paper, to resist to the change in scale and rotation of the attached area, the Gabor filter is used.

Fig. 8 Steps of the proposed method


Reading the input image with the size of MxN
Converting color picture to gray (Gray Scale)
Applying DWT conversion on image
Reading BxB image as overlapping block
Applying Gabor filter on each block read
Sorting the values obtained in the feature matrix A and assigning numbers to rows
Alphabetical sorting of matrix A
Applying the algorithm to determine the percentage of the similarity of the rows to the arranged matrix
Identifying similar blocks
Detecting copy-move forgery

3.1 Feature Extraction by Gabor Filter

Since the desired features in the image have different scales and directions, to extract information and directed features in different scales from the image is an essential step. Today, Gabor filters are widely used for this purpose due to suitable properties.

In 1946, Gabor deduced the principle of uncertainty for information on relations in quantum mechanics. According to this principle, simultaneous accuracy of a signal in two domains of time and frequency (the product of its time and frequency bandwidths) is limited by a low limit. Then he introduced a group of one-dimensional functions that achieved the low limit of uncertainty principle; in other words, the minimum simultaneous resolution in both time and frequency. These could be called fundamental (function) signals [14].

In 1980, inspired by Gabor, Dougman presented relations of uncertainty in two dimensions, and introduced a family of two-dimensional functions that reach the minimum value in the principal of uncertainty, and he called them Gabor functions. Two-dimensional Gabor function is obtained by multiplication of two-dimensional Gaussian function by a sinusoidal function in different directions of two-dimensional space. Due to very helpful properties, these functions are used in many applications as a filter in different fields of machine vision such as texture analysis, classification, image retrieval, pen detection, etc. Some of these properties to mention are simplicity, optimal simultaneous focus in location and frequency, and choice of direction and frequency for extracting image data [15, 16].

Using the two-dimensional transform of Gabor wavelet, one can extract the directional properties of the image in different scales. Physiological studies suggest that visual information processing in the visual system is done by a series of parallel mechanisms called channels; so that for each channel to use two-dimensional transform of Gabor wavelet, directional characteristics of the image at various scales could be extracted and each channel is regulated for a low frequency band width with specified direction. Mathematically, each of these channels are modeled with a pair of band-pass Gabor filters. The main advantage of Gabor filters are immutability to clearing up, rotation, scaling and image transfer. In addition, the filters can resist against photometric disorders (such as clearing changes and noise in the picture). Gabor filter in a two-dimensional spatial coordinate is a Gaussian kernel function (modulated by a complex flat sine wave), as formula (1).

$$\begin{aligned} &\mathrm{G(x,y)} = \frac{f^2}{\pi\gamma\mu}\exp\left(-\frac{x'^2+\delta^2 y'^2}{2\delta^2}\right)\exp\left(j2\pi f x'+\varphi\right) \\ &x' = x\cos\theta + y\sin\theta \\ &y' = -x\sin\theta + y\cos\theta \end{aligned} \tag{1}$$

where f is the frequency of the sinusoidal factor. θ also shows the orientation of the normal stripe of Gabor's function relative to the parallel striped of the Gabor

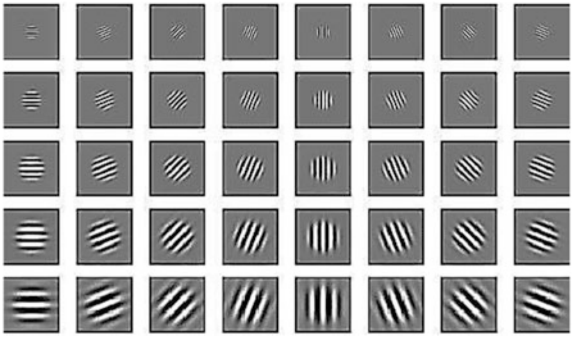


Fig. 9 Gabor filter in 5 sizes and 8 directions

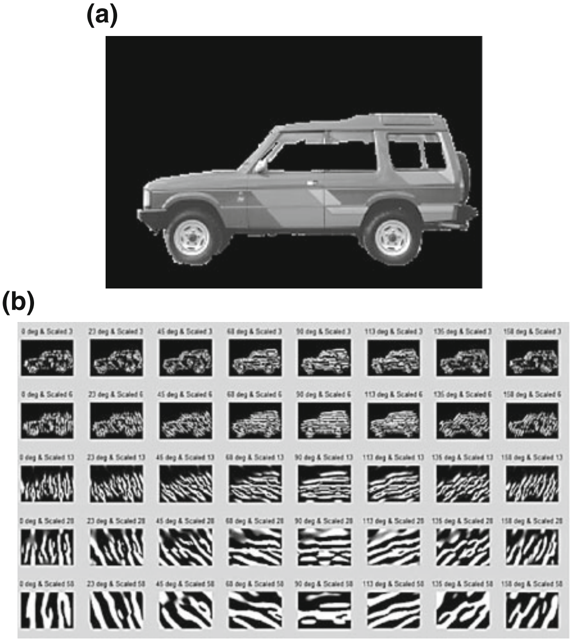


Fig. 10 **a** Vehicle image to apply to the Gabor filter. **b** The Gabor filter output on the vehicle image

function. φ is the offset of phase and σ is equal to the standard deviation of Gaussian cover. γ is the ratio of space visibility that determines the ellipticity of the Gabor function. As shown in Fig. 9, the algorithm can take advantage of forty Gabor filters (on five scales and eight directions) [17].

For example, if we use Gabor filter on Fig. 10a, the output will be the same as Fig. 10b.

Due to the fact that adjacent pixels in the image are correlated to each other, extension information could be removed through the sampling process which is less than the usual images resulting from Gabor filters [17].

3.2 Splitting the Image into Overlapped Blocks and Creating a Feature Matrix

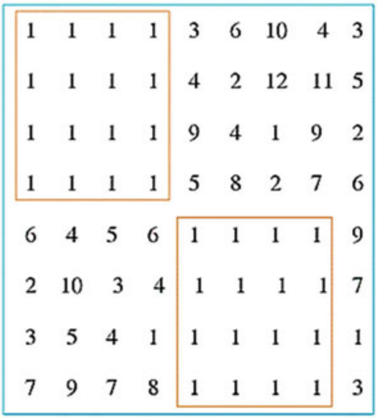
After reading the input image of the size M × N the wavelet transform is done to the "L" level, then blocks of the size b × b pixels continue from the top left corner of the image down to the lower right corner. For each position, the block is mapped to the fifth row of the Gabor filter, then the pixel values are extracted in one row of the two-dimensional A-matrix with 32 columns and $(M - b + 1) \times (N - b + 1)$ rows. Each row corresponds to a block position and to better understand the steps involved in implementing the proposed method, this algorithm is described with a small and very simple image like Fig. 11.

Because Fig. 11 is too small, a 4 × 4 window as shown in Fig. 12 is moved by applying Gabor filter on each block. According to Fig. 10, (8 × 8 block was used in the source code) overlapping blocks inserted in the feature matrix as a row vector shown Fig. 13.

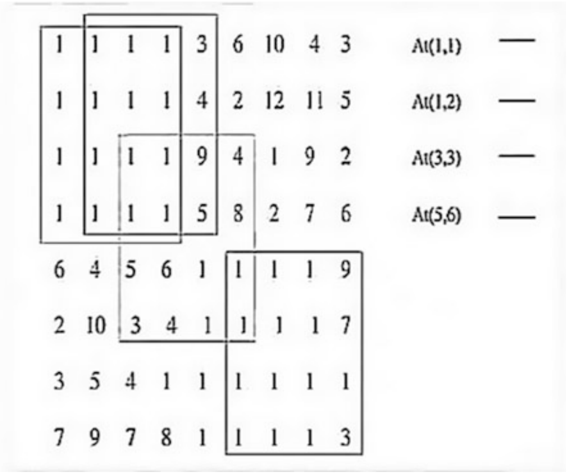
3.3 Alphabetical Sorting of Feature Matrix

To ensure the minimum number of comparisons to find the most similar blocks to each other, alphabetical sorting is applied on the feature matrix obtained from the previous step. This will locate the more similar rows next to each other and the execution time of the algorithm will reduce significantly. The result of the alphabetic sorting on the feature matrix of Fig. 13 is visible in Fig. 14.

Fig. 11 An 9 × 8 image



1	1	1	1	3	6	10	4	3
1	1	1	1	4	2	12	11	5
1	1	1	1	9	4	1	9	2
1	1	1	1	5	8	2	7	6
6	4	5	6	1	1	1	1	9
2	10	3	4	1	1	1	1	7
3	5	4	1	1	1	1	1	1
7	9	7	8	1	1	1	1	3



1	1	1	1	3	6	10	4	3
1	1	1	1	4	2	12	11	5
1	1	1	1	9	4	1	9	2
1	1	1	1	5	8	2	7	6
6	4	5	6	1	1	1	1	9
2	10	3	4	1	1	1	1	7
3	5	4	1	1	1	1	1	1
7	9	7	8	1	1	1	1	3

Fig. 12 Overlapping blocks in rows

3.4 Finding the Most Similar Blocks to Each Other Using Fourier Transform and Phase Correlation

Phase relationship is a suitable method for pattern matching. The ratio of R between the two pictures img1 and img2 is calculated according to formula (2) where 'F' is Fourier transform, and 'conj' is mixed conjunction [4, 18].

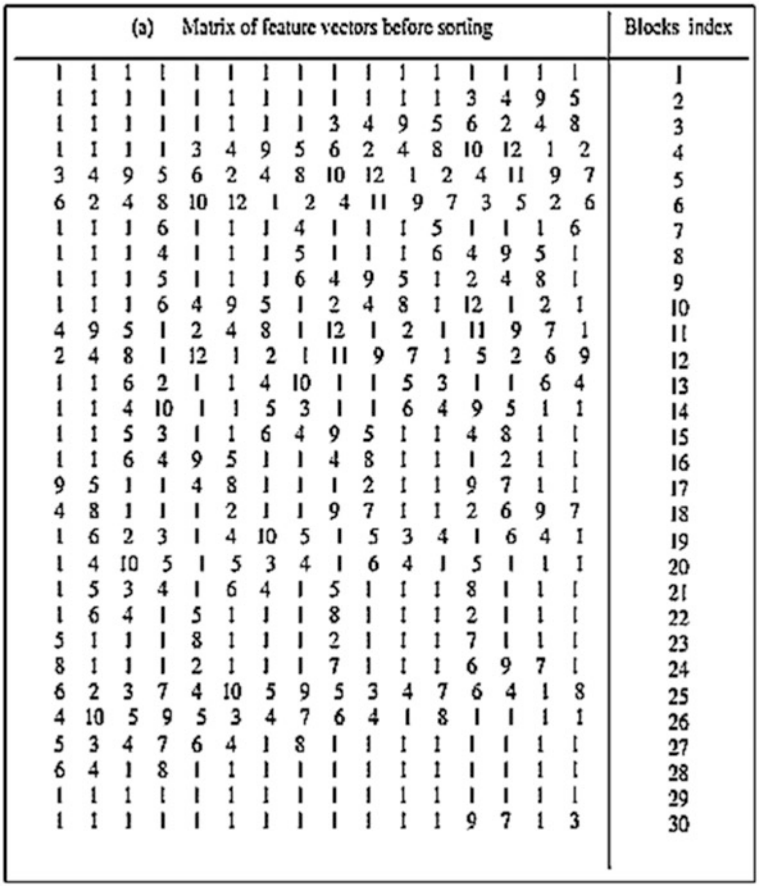
$$R = \frac{F(img1) \times conj(f(img2))}{F(img1) \times conj(f(img2))} \tag{2}$$

To find forgery in the image, a threshold proportional to the image is defined which the selection of this coefficient will be largely empirical. Surely, the more accurate this coefficient is selected, the more precise will be the locations that are detected as forgeries and also the less the extra points.

4 Investigating the results

The new program for detecting forgery by Gabor filter and the Myna [4] method was implemented in MATLAB environment version R2014a and was tested on a computer with a six gigabyte RAM and a five-core processor and Windows 8.1 operating system.

As mentioned, to resist the rotation and size change of the forged parts, the Gabor and Myna [4] filters used logarithmic-polar transformation. Results on the forged image have been investigated in different sizes and modes that shown in Figs. 15, 16, 17, 18, 19, 20, 21, 22, 23 and 24.

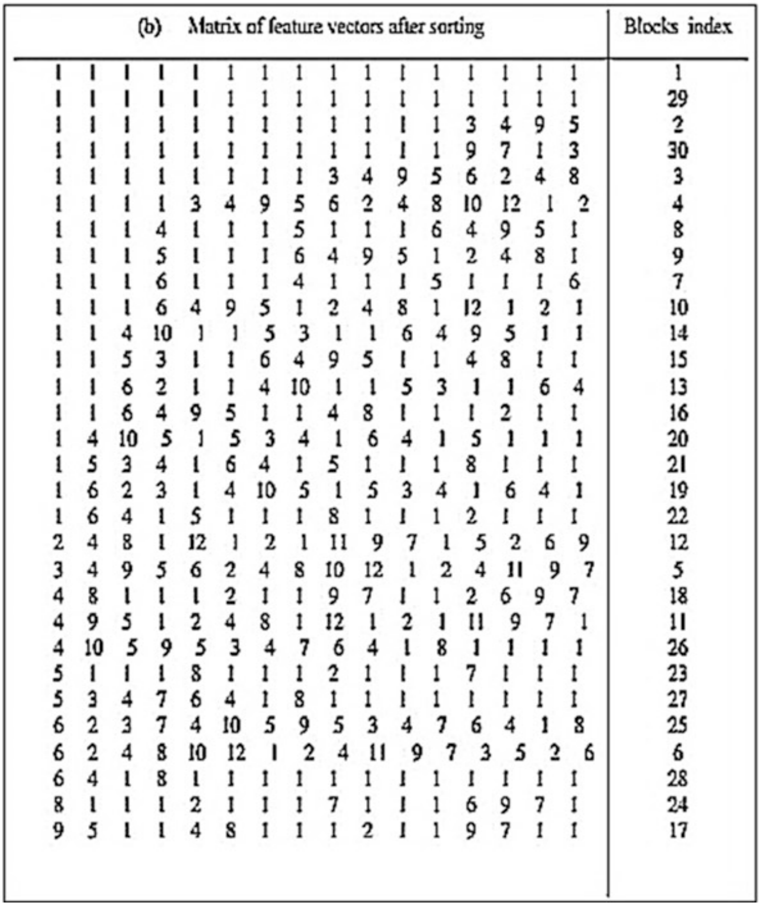


(a) Matrix of feature vectors before sorting																Blocks index
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	3	4	9	5	2
1	1	1	1	1	1	1	1	3	4	9	5	6	2	4	8	3
1	1	1	1	3	4	9	5	6	2	4	8	10	12	1	2	4
3	4	9	5	6	2	4	8	10	12	1	2	4	11	9	7	5
6	2	4	8	10	12	1	2	4	11	9	7	3	5	2	6	6
1	1	1	6	1	1	1	4	1	1	1	5	1	1	1	6	7
1	1	1	4	1	1	1	5	1	1	1	6	4	9	5	1	8
1	1	1	5	1	1	1	6	4	9	5	1	2	4	8	1	9
1	1	1	6	4	9	5	1	2	4	8	1	12	1	2	1	10
4	9	5	1	2	4	8	1	12	1	2	1	11	9	7	1	11
2	4	8	1	12	1	2	1	11	9	7	1	5	2	6	9	12
1	1	6	2	1	1	4	10	1	1	5	3	1	1	6	4	13
1	1	4	10	1	1	5	3	1	1	6	4	9	5	1	1	14
1	1	5	3	1	1	6	4	9	5	1	1	4	8	1	1	15
1	1	6	4	9	5	1	1	4	8	1	1	1	2	1	1	16
9	5	1	1	4	8	1	1	1	2	1	1	9	7	1	1	17
4	8	1	1	1	2	1	1	9	7	1	1	2	6	9	7	18
1	6	2	3	1	4	10	5	1	5	3	4	1	6	4	1	19
1	4	10	5	1	5	3	4	1	6	4	1	5	1	1	1	20
1	5	3	4	1	6	4	1	5	1	1	1	8	1	1	1	21
1	6	4	1	5	1	1	1	8	1	1	1	2	1	1	1	22
5	1	1	1	8	1	1	1	2	1	1	1	7	1	1	1	23
8	1	1	1	2	1	1	1	7	1	1	1	6	9	7	1	24
6	2	3	7	4	10	5	9	5	3	4	7	6	4	1	8	25
4	10	5	9	5	3	4	7	6	4	1	8	1	1	1	1	26
5	3	4	7	6	4	1	8	1	1	1	1	1	1	1	1	27
6	4	1	8	1	1	1	1	1	1	1	1	1	1	1	1	28
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	29
1	1	1	1	1	1	1	1	1	1	1	1	9	7	1	3	30

Fig. 13 Feature matrix before sorting

4.1 Forgery Detection Without Changing Size and Rotation and Different Rows of Gabor Filter

Result 1: The result of the forgery detection algorithm is visible using the Gabor filter in Fig. 16.
Result 2: Test on the second forged image without using discrete wavelet transform (Fig. 17).
Result 3: Test on the second forged image using discrete wavelet transform (Fig. 18).



(b) Matrix of feature vectors after sorting																Blocks index
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	29
1	1	1	1	1	1	1	1	1	1	1	1	3	4	9	5	2
1	1	1	1	1	1	1	1	1	1	1	1	9	7	1	3	30
1	1	1	1	1	1	1	1	3	4	9	5	6	2	4	8	3
1	1	1	1	3	4	9	5	6	2	4	8	10	12	1	2	4
1	1	1	4	1	1	1	5	1	1	1	6	4	9	5	1	8
1	1	1	5	1	1	1	6	4	9	5	1	2	4	8	1	9
1	1	1	6	1	1	1	4	1	1	1	5	1	1	1	6	7
1	1	1	6	4	9	5	1	2	4	8	1	12	1	2	1	10
1	1	4	10	1	1	5	3	1	1	6	4	9	5	1	1	14
1	1	5	3	1	1	6	4	9	5	1	1	4	8	1	1	15
1	1	6	2	1	1	4	10	1	1	5	3	1	1	6	4	13
1	1	6	4	9	5	1	1	4	8	1	1	1	2	1	1	16
1	4	10	5	1	5	3	4	1	6	4	1	5	1	1	1	20
1	5	3	4	1	6	4	1	5	1	1	1	8	1	1	1	21
1	6	2	3	1	4	10	5	1	5	3	4	1	6	4	1	19
1	6	4	1	5	1	1	1	8	1	1	1	2	1	1	1	22
2	4	8	1	12	1	2	1	11	9	7	1	5	2	6	9	12
3	4	9	5	6	2	4	8	10	12	1	2	4	11	9	7	5
4	8	1	1	1	2	1	1	9	7	1	1	2	6	9	7	18
4	9	5	1	2	4	8	1	12	1	2	1	11	9	7	1	11
4	10	5	9	5	3	4	7	6	4	1	8	1	1	1	1	26
5	1	1	1	8	1	1	1	2	1	1	1	7	1	1	1	23
5	3	4	7	6	4	1	8	1	1	1	1	1	1	1	1	27
6	2	3	7	4	10	5	9	5	3	4	7	6	4	1	8	25
6	2	4	8	10	12	1	2	4	11	9	7	3	5	2	6	6
6	4	1	8	1	1	1	1	1	1	1	1	1	1	1	1	28
8	1	1	1	2	1	1	1	7	1	1	1	6	9	7	1	24
9	5	1	1	4	8	1	1	1	2	1	1	9	7	1	1	17

Fig. 14 Feature matrix after sorting

Result 4: Test on the third forged image using discrete wavelet transform (Fig. 19).

Result 5: Test on a the fourth forged image without using a discrete wavelet transform (Fig. 20).

Result 6: Test on the fourth forged image using discrete wavelet transform (Fig. 21).

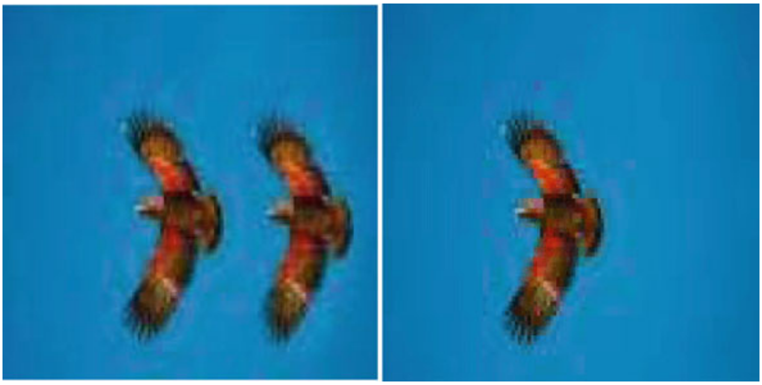


Fig. 15 The original image on the right, the forged image on the left

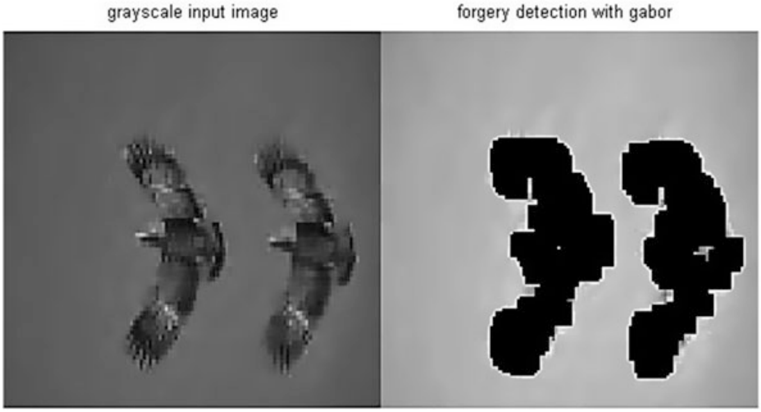


Fig. 16 Resolution: 256 × 256 pixels, block size: 88, diagnosis time: 30.737703 s, the correlation coefficient: 0.8 < R < 0.87, Gabor filter: fifth row, DWT to the first level

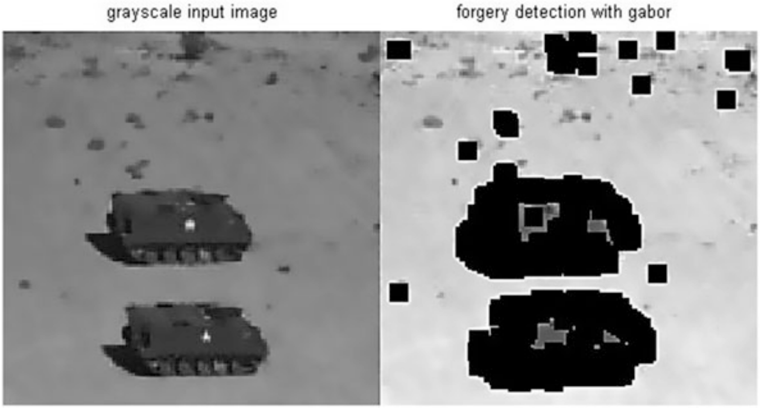


Fig. 17 Resolution: 160 × 160 pixels, block size: 88, diagnosis time: 63.503775 s, the correlation coefficient: 0.87 < R < 0.81, Gabor filter: fifth row, no DWT

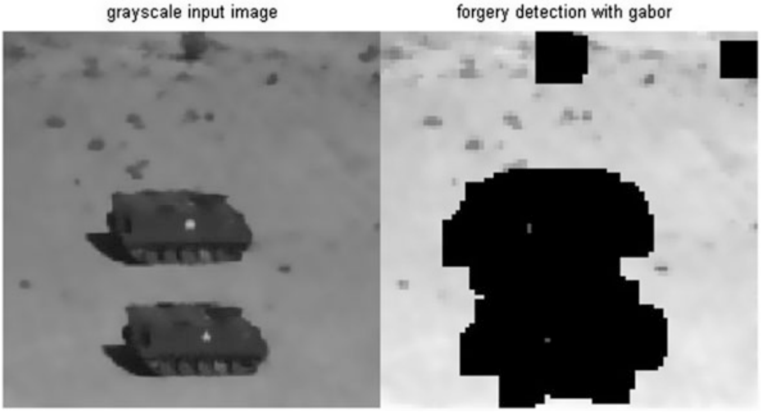


Fig. 18 Resolution: 160 × 160 pixels, block size: 88, diagnosis time: 7.636435 s, correlation coefficient: 0.87 < R < 0.81, Gabor filter: fifth row, DWT to the first level

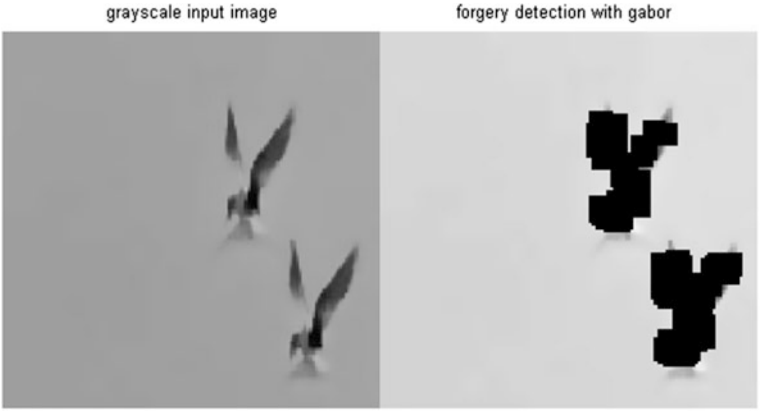


Fig. 19 Resolution: 256 × 256 pixels, block sizes: 88, diagnosis time: 31.899262 s, correlation coefficient: 0.9 < R < 0.85, Gabor Filter: fifth row, DWT to the first level

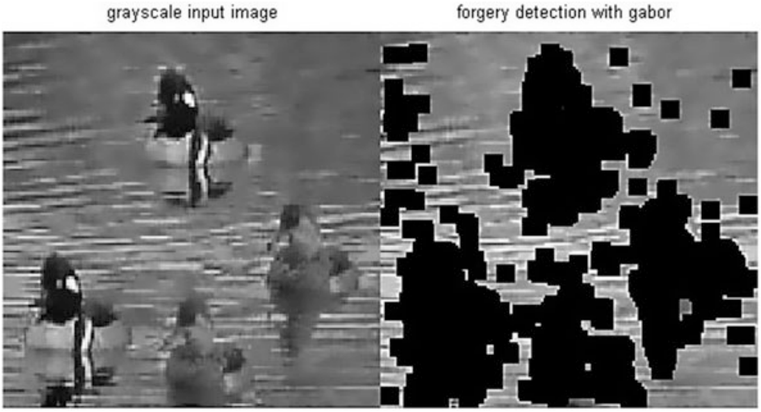


Fig. 20 Resolution: 160 × 160 pixels, block size: 88, diagnosis time: 61.1505862 s, correlation coefficient: 0.95 < R < 0.9, Gabor filter: fifth row, no DWT

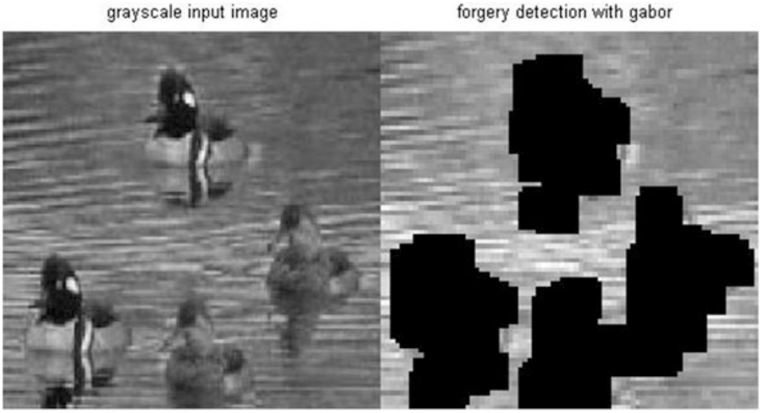


Fig. 21 Image size: 160 × 160 pixels, block size: 88, detection time: 61.1505862 s, correlation coefficient: 0.95 < R < 0.9, Gabor filter: fifth row, DWT to the first level

4.2 Resistance to Rotation

See Figs. 22 and 23.

4.3 Resistance to Resizing

See Fig. 24.

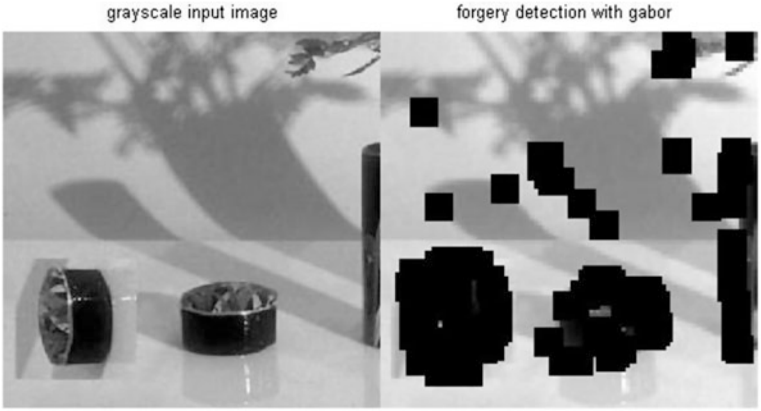


Fig. 22 The original image on the left, the forged image on the right, image size: 412 × 412 pixels, block size: 88, detection time: 15.12938 s, correlation coefficient: 0.86 < R < 0.83, Gabor filter: fifth row, DWT to second level

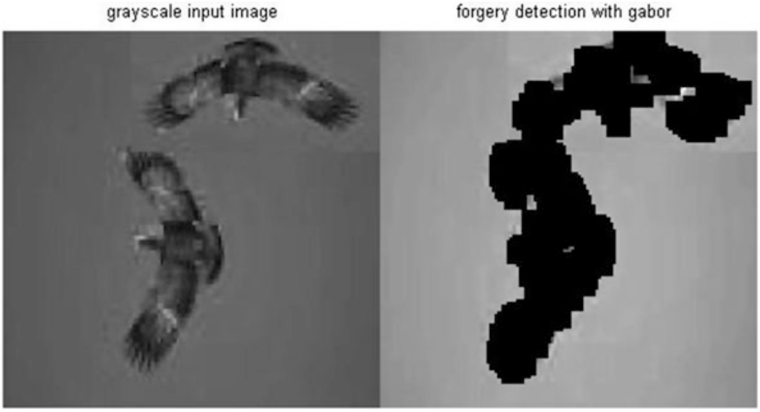


Fig. 23 The original image on the left, the forged image on the right, image size: 256 × 256 pixels, block size: 88, detection time: 30.04619 s, correlation coefficient: 0.9 < R < 0.8, Gabor filter: fifth row, DWT to the first level

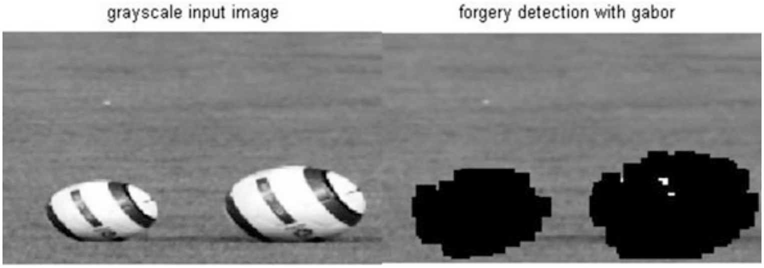


Fig. 24 The original image on the left, the forged image on the right, resolution: 300 × 600 pixels, block sizes: 88, detection time: 25.27208 s, the correlation coefficient: 0.92 < R < 0.87, Gabor filter: fifth row, DWT to second level

5 Conclusion

The obtained results and their comparison with the results indicated by Myna, it can be concluded that the new method proposed considering the time of performance is suitable, and on some images, in particular, the images in which the forged piece is resized, this method is better than Myna's method. To detect the forged area on images that forgery is not in the form of moving one part, which is not a dominant component of the image, it works well and as expected, it also works well in resize and rotation cases. However, in case of forgeries that part of the image background is used to hide part of the image or object, the performance is reduced. As already mentioned, the main advantage of Gabor filters is their immutability to clearing up, rotation, scaling and image transfer. In addition, the filters can resist against photometric disorders (such as clearing up and noise in the picture). Additional operations such as blurring may be used to eliminate the unevenness of the edge of the copied area. In such cases, the use of DCT and PCA has the advantage of being resistant to such an operation, but direct implementation lacks this advantage. It should be noted that these methods can undergo this type of operation to a certain extent. For example, if blurring is performed with high intensity, other duplicated areas cannot be identified. This occurs when blurring can be detected by eye, in which case there will be no need to search for the duplicated area. In the mentioned methods, the time complexity of the algorithm will also be reduced by reducing the length and size of the blocks.

References

1. Birajdar GK, Mankar VH (2013) Digital image forgery detection using passive techniques: a survey. Digit Invest: 226–245
2. Chauhan A (2015) Digital watermarking-revisit. J Comput Sci Inf Technol 6(1):833–838
3. Diane N, Xingming WNS, Moise FK (2014) A survey of partition-based techniques for copy-move forgery detection. Sci World J 1–13

4. Myna AN, Venkateshmurthy MG, Patil C (2007) Detection of region duplication forgery in digital images using wavelets and log-polar mapping. In: International conference on computing intelligence multimedia application, pp 371–377
5. Li G, Wu Q, Tu D, Sun S (2007) A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD. In: IEEE international conference on multimedia and expo, pp 1750–1753
6. Khan S, Kulkarni A, Khan ES, Kulkarni EA (2010) An efficient method for detection of copy-move forgery using discrete wavelet transform. Int J Comput Sci Eng: 1810
7. Gan Y, Zhong J (2015) Image copy-move forgery blind detection algorithm based on the normalized histogram multi-feature vectors. J Softw Eng: 254–264
8. Luo W, Huang J, Qiu G (2006) Robust detection of region-duplication forgery in digital image. In: 18th international conference pattern recognition, pp 18–21
9. Ryu SJ, Lee MJ, Lee HK (2010) Detection of copy-rotate-move forgery using zernike moments. Lecture notes computing science (including Subseries. Lecture notes artificial intelligence lecture notes bioinformatics), pp 51–65
10. Bayram S, Sencar HT, Memon N (2009) An efficient and robust method for detecting copy-move forgery. In: IEEE international conference on acoustics speech signal process. ICASSP 2009. IEEE, pp 1053–1056
11. AndaJan L, Fridrich J, Soukal D (2008) Detection of copy-move forgery in digital images using sift algorithm. In: Proceedings—2008 Pacific-Asia workshop on computational intelligence and industrial application, PACIIA, pp 272–276
12. Popescu AC, Farid H (2004) Exposing digital forgeries by detecting duplicated image regions. Department Computing Science, Dartmouth College. Technical Report. TR2004-515, no. 2000, pp 1–11, 2004
13. Davarzani R, Yaghmaie K, Mozaffari S, Tapak M (2013) Copy-move forgery detection using multiresolution local binary patterns. Forensic Sci Int: 61–72
14. Gabor D (1946) Theory of communication. Part 1: the analysis of information. J Inst Electr Eng III Radio Commun Eng: 429–441
15. Daugman JG (1985) Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J Opt Soc Am A:1160
16. Seryasat OR, Haddadnia J, Ghayoumi-Zadeh H (2015) A new method to classify breast cancer tumors and their fractionation. Ciência e Nat 37:51–57
17. Haghighat M, Zonouz S, Abdel-Mottaleb M (2013) Identification using encrypted biometrics. In: Computer analysis of images and patterns, pp 440–448
18. Kang X, Li Y, Qu Z, Huang J (2012) Enhancing source camera identification performance with a camera reference phase sensor pattern noise. IEEE Trans Inf Forensics Secur: 393–402

Temporal and Spatial Features for Visual Speech Recognition

Ali Jafari Sheshpoli and Ali Nadian-Ghomsheh

Abstract Speech recognition from visual data is in important step towards communication when audio is not available. This paper considers several hand crafted features including HOG, MBH, DCT, LBP, MTC, and their combinations for recognizing speech from a sequence of images. Several classifiers including SVM, decision trees, K-nearest neighbor algorithm and the sub-space K-nearest algorithm were tested feature evaluation. Further, the application of PCA for dimensionality reduction was considered in this study. Two sets of tests were carried out in this study: lip pose recognition and recognition of isolated words. For evaluation, the MIRACL-VC1 data set was considered. Self-dependent tests reached an accuracy of over 95% while in the self-independent tests, the maximum accuracy of recognition was about 52%.

Keywords Speech recognition · Temporal features · Spatial features
Dimensionality reduction · Classification

1 Introduction

Talking is a relationship between people that includes voice transmissions, facial expressions, hand movements, and body language. The features of the human face, such as the eyes, ears, nose and lip, are widely used for diagnostic tasks in the field of vision. Lip reading or visual speech recognition is a method for speech recognition using lip, face, and tongue movements, when no audio is available.

Terissi et al. [1] proposes a visual speech classification scheme based on wavelets and Random Forests (RF). To model the sequence of visual parameters, wavelet multiresolution analysis is used and the coefficients associated with these

A. J. Sheshpoli · A. Nadian-Ghomsheh (✉)
Cyber Space Research Inst., Shahid Beheshti University, Tehran, Iran
e-mail: a_nadian@pmail.sbu.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_10

representations are considered as features to model the visual information. In most of previous work, visual features of lip shapes are represented by Snake's contour. They represent the features of lip shape by six points on the lip contour extracted from the Snake model. Six point representation of lip's Snake contour, is expected to reduce the burden on recognition stage. For example, Faridah et al. [2], uses the Snake contour to find the visual feature of lip shapes. Six points from the outer edge of the lip were taken as a feature. These are the leftmost, rightmost, uppermost and lower points of the lobe. These points are taken from the snake points that result from the contour extraction process. For a lip that is not in line with the horizon, it is necessary to rotate the lip image according to its deviations.

Chung et al. [3] try to identify phrases and sentences pronounced by a talking person with or without sound. Instead of focusing on identifying a limited number of words or phrases, their work was addressed as an open world issue for unlimited natural language sentences and in wild videos. The key words in this article are, firstly, "look, listen, pay attention and spell", which the network learns to transmit a video of lip movements to the character. Second is a learning strategy for learning to accelerate education and reduce costs and third is a collection of visual speech samples consisting of more than 100,000 natural sentences from British television. Paleček [4] in his article, examined the effect of depth information using Kinect for visual speech recognition. The feature based on depth information with active appearance model (AAM) was used. This system consists of detecting the desired area, computing visual features, enhancing features, and integrating features to extract visual features. Zhang et al. [5] focuses on the impact of Kinect 3D data on the ability of the lip reading system. To find the complete lips, the left lobe and the right lobe are reconstructed using three-dimensional coordinates obtained from. Their main contribution is that the two sides of the lobe profile are reconstructed according to the three-dimensional coordinates taken by Kinect to complete the lip data. Wand et al. [6], show that the neural network based learning on raw images of the oral area has a better accuracy than systems based on the processing and extraction of features and classification. Feedforward and recurrent neural network layers, namely, Long Short-Term Memory (LSTM) were stacked to form a single structure which is trained by back-propagating error gradients through all the layers.

Although many features and classifiers are tested for visual speech recognition, previous studies have not evaluated the importance of each feature separately. Further, no evolution on the classification strategy chose for visual speech recognition is available.

The objective of this study is to explore various feature extraction methods and classifiers for visual speech recognition and find the best combined technique among extensive choices possible. In order to achieve this, we designed a system based on conventional visual descriptors and applied several well established classifiers in the classification stage. The MIRACL-VC1 dataset which contains 10 word classes collected from 15 individuals with ten repetitions was used for evaluations. Results showed that with pre-labeled data from individual accurate speech recognition is possible.

2 Method

To evaluate the performance of spatial and temporal characteristics of the lip during the act of speech, two states were considered: first, the impact of features for lip pose detection (Fig. 1) and the second, the impact of features for visual speech recognition (Fig. 2).

Detecting the pose of the lips is regarded as assigning a class label to individual frames while a certain word is spoken. For example, the classes that show if the lips are open or closed. To implement this stage, first, the face region was detected using the well-known Viola and Jones algorithm [7]. Consequently, the region consisting of lips and mouth where extracted from the detected face region. Then, by using different methods of feature extraction, the lip features were extracted and categorized into 5 classes.

For isolated words recognition using the visual data, first the mouth region was extracted as above. Then, the feature vector was extracted for each frame. In order to handle variable length videos, the features of each frame were arranged in an F × N array, where, F is the number of frames, and N is the size of the feature vector. Then, using bi-cubic interpolation, the feature vector for an input video was converted to a 20 × N array. Several classifiers where tested to obtain the accuracy speech recognition. Features and classifiers used in this study are briefly explained in Sects. 2.1 and 2.2.

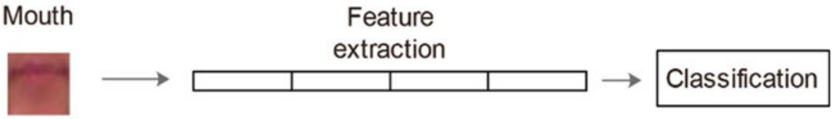


Fig. 1 Lip pose estimation

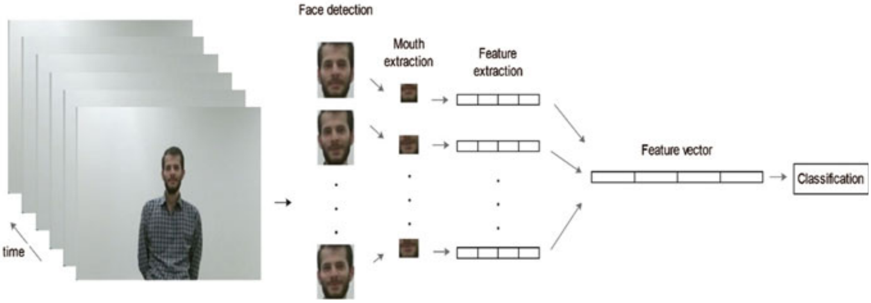


Fig. 2 Lip-reading process for speech recognition

2.1 Hand-Crafted Features

Features applied to describe the lip image include: Histogram of Oriented Gradients (HOG) [8, 9], Local binary pattern (LBP) [9], Modified Census Transform (MCT), Motion Boundary Histogram (MBH) [8], and Discrete Cosine Transform (DCT) [4].

The local appearance and shape of the object in an image can be described by the distribution of intensity gradients or edge directions provided by the HOG feature. In this study, the input lip image was divided into small connected regions called cell, and the pixels of each cell were used to create a histogram of gradient directions. The final descriptor was produced by integration of these histograms. To increase accuracy, the local histogram was normalized by the intensity obtained from a larger area of the image called blocks. The magnitude of the edges where calculated via (1) and edge angles were obtained by (2). Since the information between frames are not considered, the feature is regarded as a spatial feature.

$$g = \sqrt{g_x^2 + g_y^2} \tag{1}$$

$$\theta = \frac{g_y}{g_x} \tag{2}$$

In contrast to the HOG descriptor that is closely related to the derivative in a single frame, MBH describes a region of interest based on horizontal and vertical derivatives the optical flow image separately. The descriptor encodes the relative motion between pixels. Since MBH represents the gradient of the optical flow, locally constant camera motion is removed and information about changes in the flow field (i.e., motion boundaries) is preserved. Optical flow describes the pattern of apparent motion of the lip regions using consecutive frames of the input image sequence. MBH can be regarded as a temporal feature as it considers could describe the information between consequent frames.

LBP, another spatial feature, can be used to further include the texture information of the input region of interest. LBP is a texture operator which labels the pixels of the region of interest by thresholding the neighborhood of each pixel and considers the result as a binary number. In this study, 8-connected neighbours were considered for describing the texture of the input region of interest. The region was split into cells using, then each pixel $(x_c,\ x_y)$ was compared in a cell with its eight neighbors, and an 8-bit binary number was created using (3). The binary numbers were then converted to decimal format using (4). In each cell, the histogram was normalized and the histogram of all cells were merged to form the final descriptor. The process of producing the LBP code for a pixel (x_c, x_y) is shown in Fig. 3.

$$S(x) = \begin{cases} 1 & x \geq 0 \\ 0 & otherwise \end{cases} \tag{3}$$

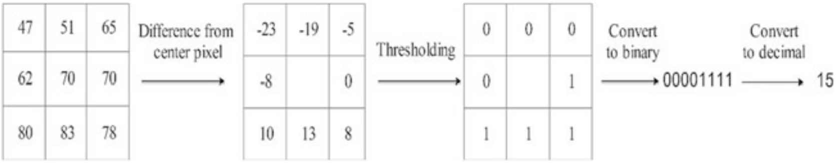


Fig. 3 An example of LBP feature for one pixel

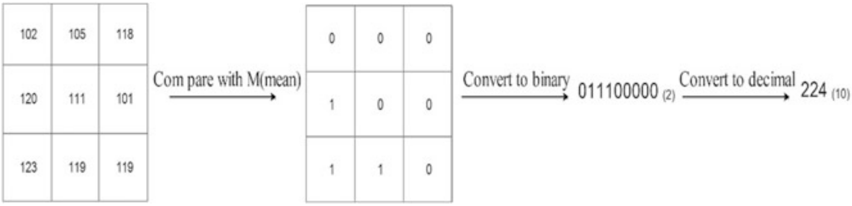


Fig. 4 Implementation of MTC feature for an example pixel

$$LBP_{P,R} = \sum_{p=0}^{p-1} s(g_p, g_c)^{2p} \tag{4}$$

MTC provides an illuminant-robust constancy assumption for solving correspondence problems appropriate for computation of displacement field in image sequences. To implement the MTC transformation, a 3 × 3 neighbor of a pixel was considered. First the average value of the neighborhood matrix was calculated (5) and the average values was compared with the corresponding pixel neighbors. If the neighbor was larger than the average value, it was replaced with one and otherwise it was filled by zero (6). Finally, put together the zero and the one that came from the comparison with the neighbors, and converts the binary string into decimal (7). Figure 4 shows the implementation steps for an example pixel.

$$M = \frac{1}{9}\sum_{x=1}^{9} I_x \tag{5}$$

$$C = \sum_{x=1}^{9} B_x \times 2^P \tag{6}$$

$$B_x = \begin{cases} 1 & I_x > M \\ 0 & otherwise \end{cases} \tag{7}$$

DCT expresses a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. DCT is a powerful feature for describing an image in the frequency domain. In this study, eight point DCT was applied to the region of interest and DCT coefficients were considered as features of the region. This study considers the two dimension frequency information of the image sequences. Thus, the feature acts as a spatial features in this study. It should be noted that DCT actually represent the image features in the frequency domain, however, for the means of this study, it was categorized as a spatial feature.

All the above mentioned features were used to describe the lip region, i.e. region of interest in an input image. In the results section, it will be shown how each feature affects the problem visual speech recognition.

2.2 Classifiers

Several classifier were considered in the classification stage: Support Vector Machine (SVM), k-Nearest Neighbors (KNN), Decision Tree (DT), and Ensemble Learning [10].

SVM is a supervised machine learning algorithm which can be used for both classification and regression challenges. In this algorithm, each data point is plotted as point in n-dimensional space, where n is defined by the number of features, where the value of each feature is the value of a particular coordinate. Classification is performed by finding the hyper-plane that differentiate the two classes the best. If features can't be mapped into two classes with a linear hyper-plane, a kernel can be used to map the data into a high dimensional vector space where linear relations exist among the data. Then, a linear algorithm can be applied in this space. In this study, the cubic kernel function as formulated in (8) was used for classification.

$$k(x_1, x_2) = (x_1^T x_2 + 1)^3 \tag{8}$$

KNN is another simple algorithm that stores all available cases as training data and classifies new cases based on a similarity measure (e.g., distance functions). A case is classified by a majority vote of its neighbors, with the case being assigned to the class most common amongst its K nearest neighbors measured by a distance function. In this study we chose K = 1, thus, each feature vector is simply assigned to the class of its nearest neighbor.

Decision tree learning uses a decision tree as a predictive model to go from observations about an item represented in the branches of the tree to draw conclusions about the item's target value which are represented in the leaves. An important notice about decision trees is that they are robust against noise. In the case of speech recognition, where the region of interest in the image comprising the mouth is normally a small region, noise could play in important role in the classification results. Thus, the effectiveness of this classifier for visual speech recognition was evaluated in this study.

Ensemble learning is a machine learning paradigm in which several generic learners have been trained to solve a similar problem. Compared to commonly used machine learning methods that try to learn a hypothesis from train data, ensemble methods try to build a set of hypotheses and combine them for use. Since a large number of features were considered in this study, ensemble learning was also considered to explore how effective it could be for the task at hand.

3 Results

This section reports the results of evaluations. MIRACL-VC1 dataset was used for this evaluation, which is a lip reading dataset consisting on 1500 samples (15 persons × 10 phrases × 10 instances) and 1500 phrases (15 persons × 10 phrases × 10 instances) [10]. The dataset covers phrase such as *navigation*, *connection*, etc., and everyday phrases like *Nice to meet you*, *I love this game*, etc. The Kinect sensor was used to acquire 2D images and depth maps with a resolution of 640 × 480 pixels and at an acquisition rate of 15 fps. The distance between the speaker and the Kinect is about 1 m. In this study, only color images acquired by the Kinect were considered. The depth images in the dataset were used for extracting the lip region in the image. However, our observations showed that only using color images and the Viola-Jones algorithm suffices for the purpose of this study [7]. Several tests were conducted for evaluations:

- Evaluation of lip pose classification.
- Evaluation of speech recognition.
- Speaker independent (SI) test.
- Speaker dependent (SD) test.

For SI configuration, leave-one-speaker-out strategy was utilized where data from a single speaker were used as the validation data, and the records of the remaining speakers were used for the training stage. The same procedure was repeated for each speaker in the dataset.

In the SD configuration, the training and the testing data were obtained from the same speaker. For each of speakers in the dataset, the leave one video out cross validation was carried out, that is, two videos were used for testing and the rest were used for training. The results of this article are based on the accuracy measure:

$$Accuracy = \frac{t_p + t_n}{N} \tag{9}$$

where N is the number of total test data for each experiment. First, we evaluated the accuracy of lips pose recognition. To achieve this, the mentioned dataset was used to extract the lips and then, based on the shape of the lips, the extracted lobe area was classified into five classes. Classes were recognized manually. The lips were categorized into five classes based on the apparent visual movements of the mouth as illustrated in Fig. 5.

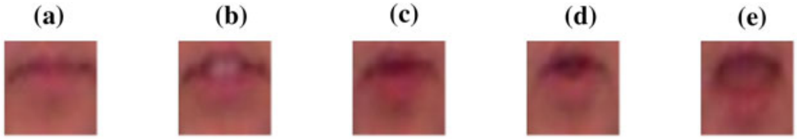


Fig. 5 Five classes of lip poses. The classes were obtained considering 5 levels of movement in a lip pose

Table 1 Results of the lips poses recognition. **a** Without principal component analysis (PCA). **b** With principal component analysis (PCA)

	DT	Cubic SVM	KNN	Subspace KNN
a				
DCT	73.7	74.40	72.0	86.1
HOG	71.0	**86.5**	83.9	84.5
LBP	51.5	68.8	63.1	70.3
HOG + LBP	69.6	86.8	83.2	84.9
MTC	64.2	84.4	82.3	82.4
b				
DCT	74.0	**86.2**	84.0	85.0
HOG	71.4	**85.8**	82.5	82.8
LBP	53.5	65.9	62.9	65.5
HOG + LBP	69.0	84.5	81.7	82.8
MTC	68.1	76.8	77.0	82.3

Table 1a, b represent the results of lip pose classification. Since no motion data is available in this stage, only spatial feature were considered. In both tables, applying PCA did not make a significant difference on the results. DCT coefficients in this context was an exception. When PCA was applied to the feature vector obtained by applying DCT, the accuracy were increased to more than 10%. Further, SVM with cubic kernel showed to have a better accuracy compared with other methods. In these tests, DT had the lowest accuracy no better than 73.7%. Overall, the lip pose recognition stage showed an accuracy of 86.5% when SVM and DCT features were used.

Table 2a, b show the results for evaluating visual speech recognition using the pre-mentioned features and classifiers. When image sequence were used for recognition, except for the DT classifier, the accuracy rate was significantly increased compared to the case of lip pose recognition. Further, the accuracy of DCT compared with other features was also reduced. The best results were obtained when HOG and MBH were combined. This is an expected result, because, when image sequences are used for classification, MBH can describe the relation between different frames and thus better results could be obtained. When DCT, HOG, LBP and MTC, i.e. only spatial features were used for speech recognition, the accuracy of recognition was significantly reduced. This is due to the same fact that these features do not consider the motion information for classification. In these tests,

Table 2 Results of speech recognition with SD strategy. **a** SD training and testing without principal component analysis. **b** SD training and testing with principal component analysis

	DT	Cubic SVM	KNN	Sub-space KNN
a				
DCT	41.4	58.5	75.7	92.7
HOG	36.5	91.2	87.1	94.1
LBP	20.7	64.1	68.6	84.6
HOG + LBP	32.7	91.1	87.6	94.4
MBH	31.2	88.4	88.2	95.1
MBH + HOG	33.4	91.6	88.8	**95.6**
MBH + LBP	28.3	89.0	86.5	**95.7**
MBH + DCT	40.0	82.3	87.3	94.9
MTC	27.2	85.4	83.9	88.5
b				
DCT	52.9	84.2	87.4	91.9
HOG	48.4	87.1	90.4	93.3
LBP	26.6	52.6	64.5	83.7
HOG + LBP	43.0	85.2	90.4	93.5
MBH	41.4	87.1	88.2	94.6
MBH + HOG	45.4	90.9	91.7	**95.0**
MBH + LBP	42.2	89.0	91.4	**95.5**
MBH + DCT	36.3	86.1	91.1	94.3
MTC	35.1	83.4	83.5	83.4

applying PCA did make a significant change in the recognition accuracy. A notable difference between lip pose recognition and speech recognition was improvement of accuracy using the Sub-space KNN classifier. For the case of speech recognition, Sub-space KNN classifier provided accuracy significantly higher than the SVM classifier. In fact, the sub-space classifier shows a better performance when the size of the feature vector is large. It should be notices for an N frame video, the feature size is 20 times larger than when a single frame is used for describing the lip region in the image. Please refer to the Method section for information on the size of the feature vector.

Table 3 provides the results of speech recognition using the SI evaluation strategy. As the results show, in the case of SI tests, the accuracies were significantly reduced. This is probably due to the fact the different people show have different visual features for saying each phrase. In SI tests, SVM showed to better classifier compared with other classifiers tested in this paper. Temporal features had a negative impact on the results. This is probably because transitions from one frame to other frames does not correlate among different samples of the data set.

In summary, when temporal features had no effect on the results, DCT, LBP, and HOG features showed to be significant features for lip pose and speech recognition, in addition, in such cases the cubic SVM showed to be the superior classifier. In the SD tests, were temporal information were significantly appropriate for speech

Table 3 Results of speech recognition with SI strategy

	DT	Cubic SVM	KNN	Sub-space KNN
DCT	23.1	**51.7**	35.5	09.8
HOG	16.2	**48.1**	30.2	10.6
LBP	15.4	28.1	19.9	9.20
HOG + LBP	16.1	45.2	30.1	10.8
MBH	17.7	37.0	22.8	11.7
MBH + HOG	18.2	45.9	30.2	10.2
MBH + LBP	17.6	39.8	27.4	11.6
MBH + DCT	24.1	41.9	25.5	10.8
MTC	19.9	45.1	29.7	10.1

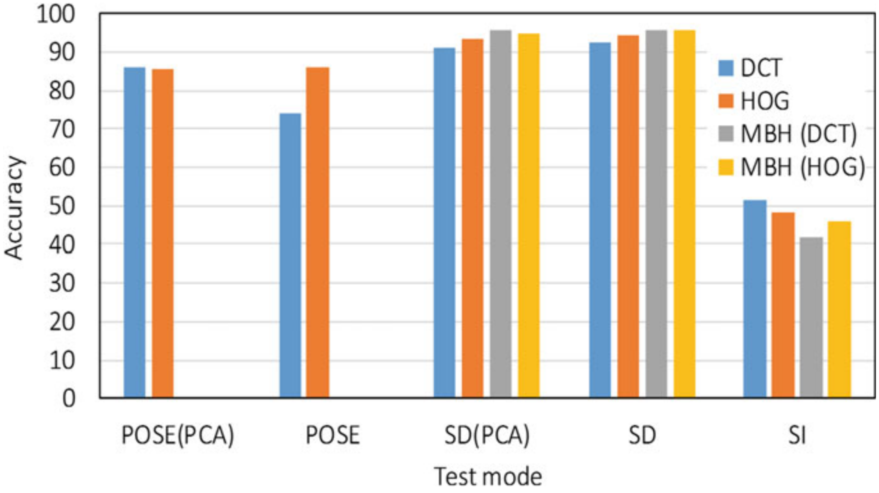


Fig. 6 Comparison among significant features in all test strategies

recognition, KNN and sub-space KNN showed to be the superior classifiers. Figure 6 shows summarizes the significant results among all tests. Based on results shown in this figure, SI testing does not show significant accuracy, thus, however, if a recordings from the same person is available, it is possible to recognize a conversation with significant accuracy.

4 Conclusion

In this paper, several spatial and temporal features were tested for visual speech recognition. Also, 4 classifiers accompanied with PCA for feature size reduction were used for evaluation of the selected features. The results showed that DCT and

HOG were significant features for visual speech recognition. Further, in SD test strategy, the added temporal features showed to have a significant impact on increasing the accuracy of visual speech recognition. In the SI test strategy, the results were mostly insignificant. That is, in all evaluation cases, the accuracy did not increase more than 52%. In general, the results showed that providing improved solution for anonymous visual speech recognition requires further research and more sophisticated tools. However, if pre-labeled information of each individual is available, accurate visual speech recognition can be achieved.

References

1. Terissi LD, Parodi M, Gómez JC (2014) Lip reading using wavelet-based features and random forests classification. In: 22nd international conference on pattern recognition, Sweden, 24–28 Aug 2014
2. Faridah F, Achmad B (2015) Lip image feature extraction utilizing snake's control points for lip reading applications. Int J Electr Comput Eng 5(4):720–728
3. Chung JS et al (2016) Lip reading sentences in the wild. In: Asian conference on computer vision. Taiwan, 20–24 Nov 2016
4. Paleček K (2015) Comparison of depth-based features for lipreading. In: 38th International conference on telecommunications and signal processing (TSP), Prague, 9–11 Jul 2015
5. Wang J et al (2015) Lipreading using profile lips rebuilt by 3D data from the Kinect. J Comput Inf Syst 11(7):2429–2438
6. Wand M, Koutník J, Schmidhuber J (2016) Lipreading with long short-term memory. In: International conference on acoustics, speech and signal processing, Shanghi, 20–25 Mar 2016
7. Viola P, Jones MJ (2004) Robust real-time face detection. Int J Comput Vis 57(2):137–154
8. Rekik A, Ben-Hamadou A, Mahdi W (2016) An adaptive approach for lip-reading using image and depth data. Multimedia Tools Appl 75(14):8609–8636
9. Pei Y, Kim T-K, Zha H (2013) Unsupervised random forest manifold alignment for lipreading. In: Proceedings of the IEEE international conference on computer vision, USA, 1–8 Dec 2013
10. Ho TK (1998) Nearest neighbors in random subspaces. In: 1998 proceedings joint IAPR international workshops advances in pattern recognition, Australia, 11–13 Aug

The Application of Wavelet Transform in Diagnosing and Grading of Varicocele in Thermal Images

Hossein Ghayoumi Zadeh, Hamidreza Jamshidi, Farshad Namdari and Bijan Rezakhaniha

Abstract Varicocele is the abnormal dilation and tortuosity of venous plexus (venous) above the testicles. The pattern of abnormal heat distribution in the scrotum can be diagnosed by the help of thermal imaging. Thermal Imaging is a distant, non-contact, and non-invasive method. Thermal imaging was conducted on 50 patients who referred to the hospital 501 (AJA). It was implemented by non-contact infrared camera VISIR 640. Capabilities of thermography method was then evaluated. In order to evaluate and diagnose the varicocele, thermal asymmetry and Haar wavelet techniques were used based on thermal imaging. In two methods, with the help of heat distribution, varicocele can be detected using a thermal camera; one of these two methods includes thermal asymmetry and increased temperature in venous plexus (pampiniform) with no thermal increase in the testicle of the same side (It is grade II of varicocele with a temperature difference of about 1 °C), and the other includes the increased temperature of venous plexus (pampiniform) with increased temperature of the testicle in the same side(It is grade III of varicocele with a temperature difference of about 1.5 °C). The accuracy of the recognition of thermography is 76% in different grades of varicocele. According to the results of the study, thermography is a useful method for the initial screening process. In addition, it can be applied as a supplement to other diagnostic techniques due to lack of exposure, low cost and its exact diagnostic capacity in varicocele.

Keywords Thermal imaging · Varicocele · Heat distribution

H. G. Zadeh (✉)
Department of Electrical Engineering, Vali-e-Asr University of Rafsanjan,
Rafsanjan, Iran
e-mail: h.ghayoumizadeh@gmail.com

H. Jamshidi
Department of Electrical Engineering, Islamic Azad University
Kermanshah Branch, Kermanshah, Iran

F. Namdari · B. Rezakhaniha
Department of Urology, AJA University of Medical Sciences, Tehran, Iran

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_11

1 Introduction

Varicocele is the abnormal dilation and tortuosity of venous plexus (venous) above the testicles. This disease rarely occurs at ages under 10 but its prevalence among young adults and in fertile men is 15 and 20–40%, respectively [1]. Among those with secondary infertility (already have child), its prevalence may reach 70% [2]. 90% of cases occur in the left side and only 10% of them are bilateral [3]. This is because of higher length of the vein discharging the left testicle blood into the renal vein and its more vertical angle compared to the right testicle. One of the main theories explaining the pathophysiology of varicocele is the theory of testicular temperature increase [4]. What is certain is that the etiology of temperature rise in the standing and Valsalva maneuver positions seems questionable [5]. It should be noted that increased intra-abdominal pressure is considered a weak risk factor for varicocele [6]. However, there are still many discussions on the treatment of varicocele by surgery. Most of men with varicocele are able to have child, so spermatic vein ligation surgery is not recommended for all common cases varicocele. In some men, varicocele is a progressive condition which causes the loss of former fertility [7]. Current guidelines recommend surgery for infertile men with diagnosed conditions and semen disorders [8]. However, recent research has led to the revision of these suggestions and it has been shown that surgery option in the treatment of varicocele does not increase the chance of fertility in cases that varicocele is considered the only proof of infertility [9]. This problem needs a solution which is timely selection of men with deteriorating varicoceles and reduced semen quality. The negative toxic effect of varicocele on the testes is proven by evidence; which also indicates that untreated cases varicocele can have unpleasant results such as male infertility [10]. In cases where the quality of semen is deteriorated as a result of progressive varicocele, surgical operation should be performed. The lack of diagnostic criteria on performing surgery is an issue in this type of varicocele. Semen quality analysis is not regarded as a screening method. Currently, varicocele diagnosis depends on physical examination and ultrasonography/Doppler scrotum [10, 11]. Abnormal thermal patterns can be easily detected by thermal imaging. On the other hand, clinical thermal imaging was performed occasionally in the past as a short objective diagnostic method [2]. Although it is a non-specific method and highly dependent on background and environmental factors in some cases, there are several reasons that have caused thermal imaging to be generally accepted and welcomed in the medical community. First of all, thermal Imaging is a distant, non-contact, and non-invasive method [12]. Duration of imaging is very short, so it is possible to simultaneously monitor a large part of population. The colors of thermograms can be easily and quickly interpreted. In addition, this method records only the natural radiations from the surface of the skin and there would be no trace of harmful rays. Therefore, this method can be repeatedly used for a long term. Finally, thermal imaging is an immediate method which can monitor the dynamic changes of temperature. Digital infrared thermography of scrotum is a non-invasive, sensitive diagnosis tool for the detection of primary varicocele using the scrotal skin surface temperature measurement. In spite of objective truth and short duration of this diagnostic method, it used to be sporadically used during clinical procedures in the past. A recent research has proposed the diagnostic criteria of scrotal thermography for varicocele detection [13]. It has been also proven that scrotal thermography is a useful diagnostic method for mild varicocele and the postoperative period [14] Thermography has been also applied as a successful for follow-up [12, 15, 16]. Asymmetry technique is one of the basic principles and methods for analysis of thermal images. Some researchers [17] have conducted the analysis of asymmetry based on temperature change and skewness and kurtosis of the image and the study area. According to studies conducted in other countries, it has been revealed that the thermography imaging systems, considering the proper and correct diagnostic results and a few numbers of false positive or negative responses, can have an appropriate performance not as a precise and absolute method but as a supplement to other techniques [18–20].

The present study also aims to identify the strengths and weaknesses of thermography systems in scrotal thermography for diagnosis of varicocele. The main assumption is that scrotal digital infrared thermography is the main tool in the diagnosis of varicocele.

2 Materials and Methods

In the present study, a thermography device with non-contact infrared camera (VisIR 640) was used. This system has a resolution of 110,592 pixels per image and a minimum thermal resolution of 0.01 °C. SatIr Wizard software was used for analysis and presentation of images. For conducting thermography, the patient was placed in a certain distance of about 30 cm from the infrared thermal camera and the body radiations that are in the range of 0.7–0.9 μm were sent to the computer processing system of images after passing through a focusing lens that acts as a filter. Points which must be observed when thermal imaging include the following: (1) The patient must relaxed and comfortable psychologically before the performance of thermal imaging. (2) Room temperature is set in approximately 25 °C (not too cold and not warm). (3) It is best that patient take off their shirts for 10 min before shooting and somewhere sits calmly. (4) If fluorescent lamps are used in place of imaging, it must be turned off.

After determining the location and based on the received wavelength, temperature of each point of the body appears on the display in a certain color. The accuracy of detection of radiations from the surface of the skin depends on the skill of the specialist in reading the thermography images, the operator's skill in setting the visibility window, the selected color scale, and patient's distance to the device. In this study, 50 patients suspected of varicocele were clinically examined by a specialist and the results were recorded. It is worth mentioning that all patients have been randomly selected, and all of them had referred to urologist. The average age

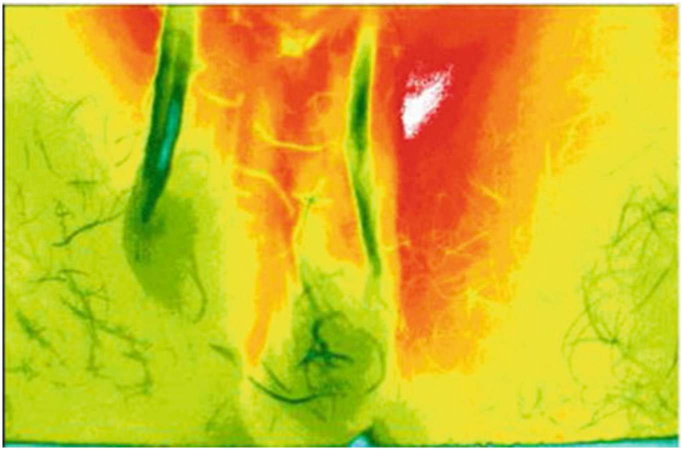


Fig. 1 Thermographic image of grade 2 varicocele

of patients is 35 years. Before conducting the thermal imaging, some items such as imaging conditions, room temperature, patient's comfort, and so on, all are effective in providing false or true responses of thermography, were investigated. Then, the patients were placed in front of the camera in a standing position with naked lower part of the body. The patients were asked to hold the tip of their penis upward with their finger in a way that the patients' legs were stretched, testicles hung freely, and the head of penis was kept against the abdominal wall. Finally, thermography images were taken of the testicles by the operator. Increased scrotal temperature is considered the main reason for semen disorder in patients with varicocele. Monitoring the temperature of the scrotum is part of the diagnostic evaluation of varicocele. In digital infrared thermography of the scrotum, highly sensitive infrared cameras are used for tracking and measuring the temperature of the scrotum area [22]. Thermal images of a patient with grade II, varicocele is noticeable in Fig. 1.

To evaluate the temperature of the thermal images which are as thermal pixels, two strategies are proposed. One of them is thermal asymmetry evaluating, and the other is a model based on Haar wavelet transform. In the first method, the subject of thermal asymmetry is examined in venous plexus pampiniform. Secondly, a model is proposed based on Haar wavelet transform. Both methods are useful for the diagnostic purposes and they also complement each other.

Noise elimination is one of the complicated and ambiguous subjects, and their importance cannot be ignored in signal processing and analysis system.

Specialists in the field of signal analysis and systems, examined different Filters and methods and conversion and mapping in order to remove noise. They have done several studies to improve the functioning of the filters and methods. Many experts have assessed Strengths and weaknesses of each of these methods, by using real samples in different fields to remove the potential noise from the data, signals and systems. And they have tried to obtain efficiently and accurately results. One of the new methods in recent decades, which has become important from different aspects is wavelet method. An exceptional feature of this method is that the wavelet analysis combines both time and frequency domains together. A field of wavelet transform is uses it as a preprocessing tool for smoothing [21]. The main form of a wavelet transform is defined as follows, that ψ (x) known as the mother wavelet.

$$\psi_{jk}(\mathrm{x}) = 2^{\frac{j}{2}}\psi\left(2^{j}x - k\right) \tag{1}$$

j and k are used for changing the scale and shift (transmit) of the wavelet. In 1910, Alfred Haar introduced the first recognized wavelet, indicating, every continuous function f (x) in the range [0, 1] can approximate it by using a series of step functions. These functions are shown below:

$$\psi(\mathrm{t}) = \begin{cases} 1 & (t \in [0, 0.5)) \\ -1 & (t \in [0.5, 1)) \\ 0 & otherwise \end{cases} \tag{2}$$

According to the following features, ψ_{jk} function is an orthonormal function.

$$\int \left(\psi_{jk}.\psi_{lk}\right) = 0, ((j \neq 1)\forall(k \neq m)) \tag{3}$$

$$\int \left(2^{j/2}\psi\left(2^{j}\mathrm{x} - \mathrm{k}\right)\right)^{2} dx = 1 \tag{4}$$

The transfer function of f(x) is expressed as follows in which ψ (t) is known as the parent wavelet or Comparator The function.

$$f(x) = c.\psi(\mathrm{t}) + \sum_{j=0}^{n-1}\sum_{k=0}^{2^{j}-1} c_{jk}\psi_{jk} \tag{5}$$

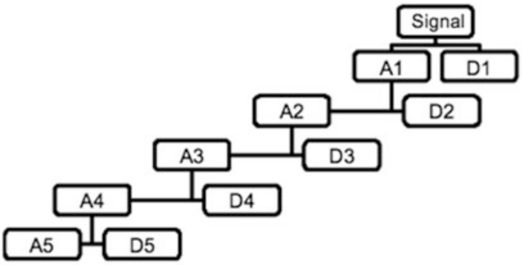
$$\psi(\mathrm{t}) = \begin{cases} 1 & 0 \leq t < 1 \\ 0 & otherwise \end{cases} \tag{6}$$

$$\int \psi(\mathrm{t})dt = 1, \int \psi(\mathrm{t})dt = 0 \tag{7}$$

Using the Haar wavelet, the signal is decomposed into an original signal known as the approximation shown as $a_{i,}$ and more minor signals called details and indicated as d_i . At each stage of decomposition, a_i can be decomposed again using Haar wavelet so that the sum of a_n, d_i create the same original signal (Fig. 2).

The original signal = a_n + d_1 + d_2 +⋯+ d_n.

Fig. 2 Decomposition of wavelet tree based on the original signal



The model proposed for using Haar wavelet transform for investigating the thermal images includes these aspects. First, thermal line from temperature pixels is drawn in vertical direction of the left and right testicles. Then it is seen as a continuous signal. Now, Haar wavelet transform is conducted on it. The image of the obtained signal is shown in Fig. 3.

Now, Haar wavelet transformation is applied on the signal in three stages. The composed figure can be seen in Figs. 4 and 5 concerning, healthy and patient samples.

If we notice the level of difference of peaks in the d1, d2, d3, in the person with varicocele, we will find that these differences are more than healthy samples. The next stage includes thermal asymmetry which is examined In addition to the

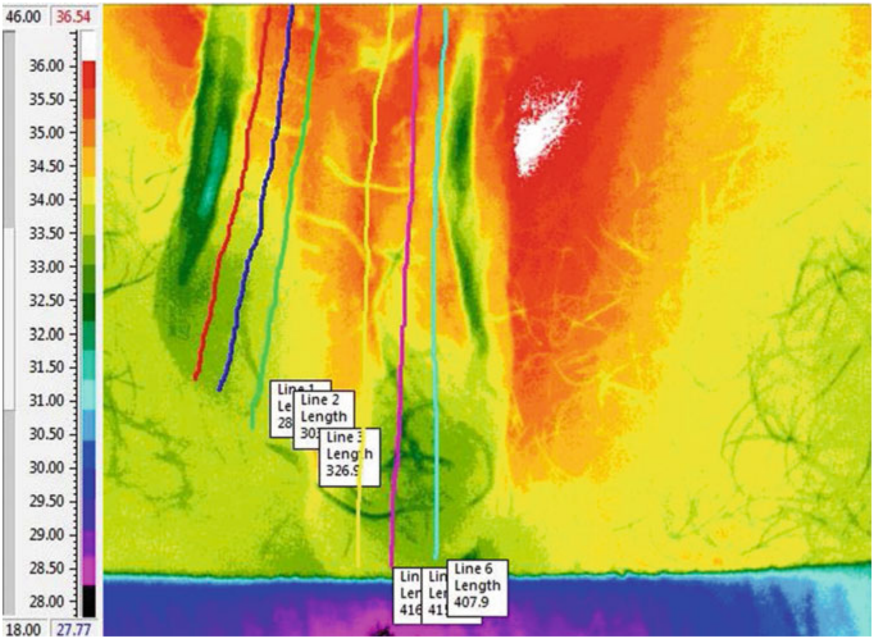


Fig. 3 Creating thermal lines in parallel with the penis on both sides of the testicles

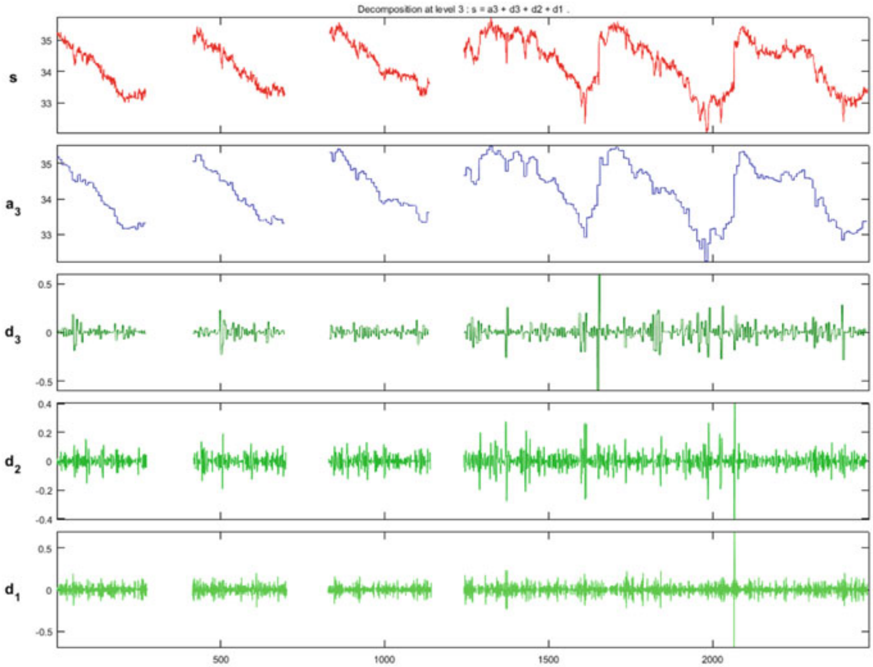


Fig. 4 Decomposed Tipaks index by Haar wavelet transform in the 3-step process of decompose for the sample with grade 3 varicocele

Features of the wavelet transform method. As it was mentioned before, at the first stage, the patient's history was recorded. Then clinical examination was performed by a specialist in the field of stereotypes urology. The following thermal imaging was conducted on the patient. Finally, the results were compared with each other to identify whether the samples suffer varicocele or not. According to the previous studies, the temperature of the right and left testicular and venous plexus are the focused points. If the temperature of the both sides of the scrotum is uniform, the result of the related sample is normal [14]. A sample of Thermal images taken from the patient is shown in Fig. 6. It is evident from the image that the thermal pattern is uniform and asymmetric on both sides of the scrotum. Then, the obtained results were compared with the related doctor. Finally, it was shown that the analyzed sample is not suffering varicocele.

The first pattern with asymmetric temperature distribution is described in the pampiniform venous plexus of scrotum which exclusively engages the upper part of the testicles. An example of this pattern obtained from the results has been shown in Fig. 7. As it can be observed, an asymmetric thermal pattern can be seen in the pampiniform venous plexus area at the top of the right and left testicles.

The second pattern with asymmetric temperature distribution is in the pampiniform venous plexus that is expanded towards the testis of the same side or

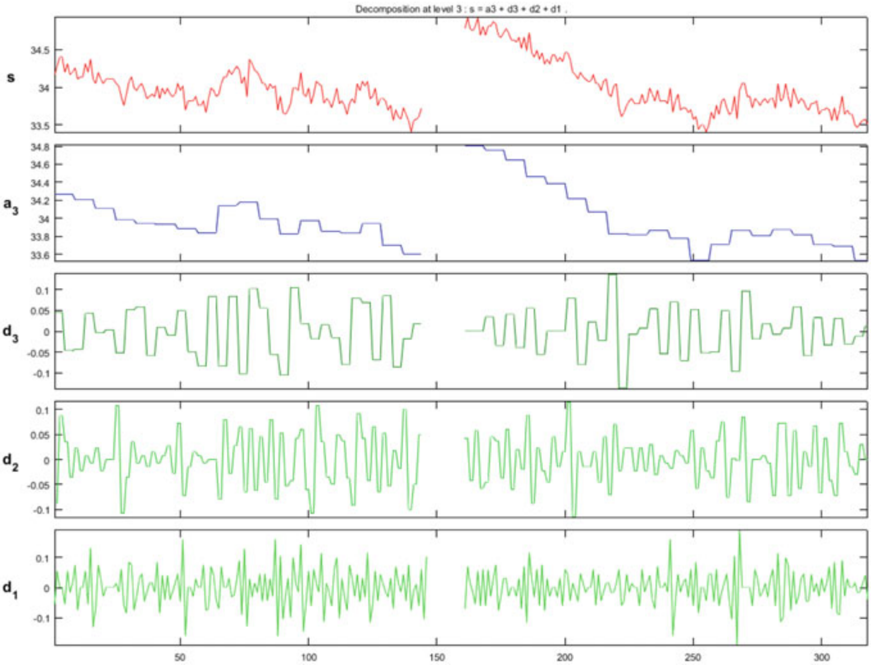


Fig. 5 Decomposed Tipaks index by Haar wavelet transform in the 3-step process of decompose for the healthy sample

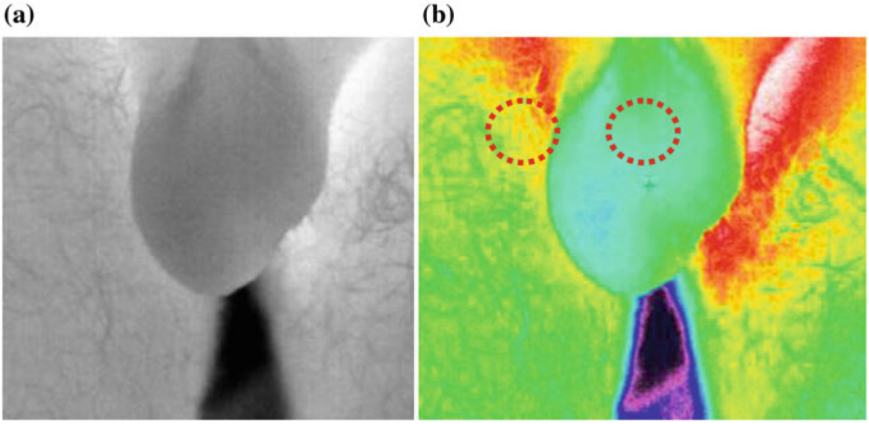


Fig. 6 Thermographic image of testicles of a healthy person in two states; **a** gray, **b** colorful

involves the hyperthermia of the whole testicles. Although the full bilateral hyperthermia of the scrotum is rare, it is a definite symptom of varicocele. As previously shown, localization and extension of hyperthermia area are very

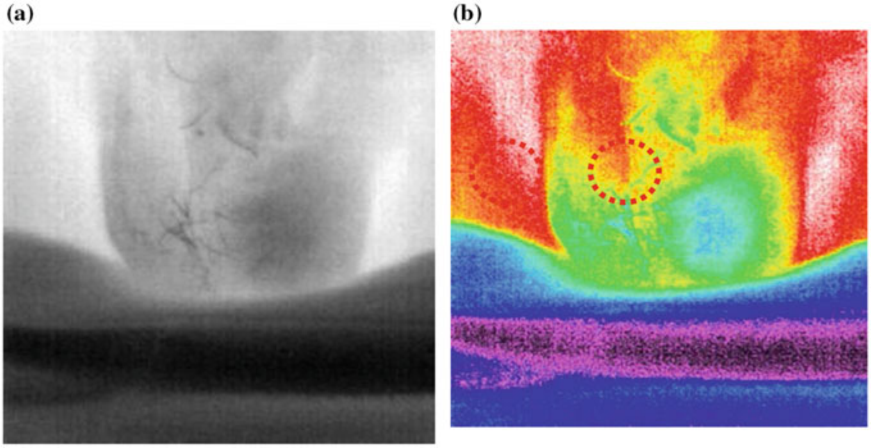


Fig. 7 Asymmetric thermal pattern in the upper part of the testicles in thermography image **a** gray, **b** colorful taken from a patient with grade 2 of varicocele

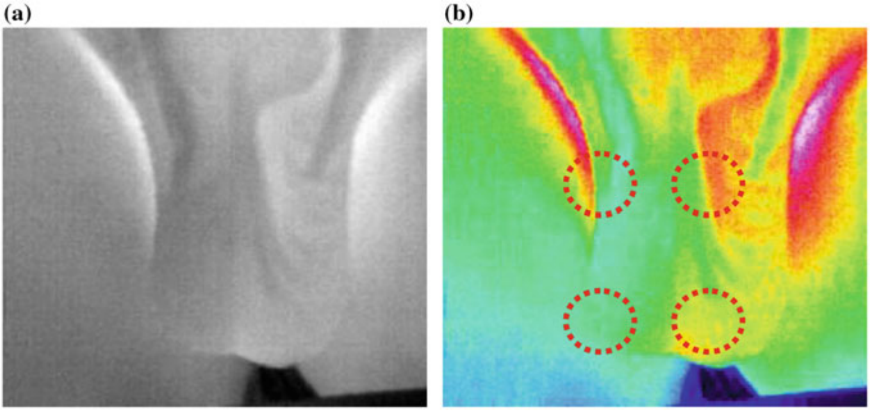


Fig. 8 Asymmetric thermal pattern in the pampiniform venous plexus and the scrotum in thermography image **a** gray, **b** colorful taken from a patient with grade 3 of varicocele

important in the interpretation of thermography. An example of this pattern has been shown in Fig. 8. According to this figure, increased temperature involves both the pampiniform venous plexus and the scrotum.

3 Results

The obtained features using Haar wavelet transform are presented in Table 1.

Table 1 The obtained features of Haar wavelet transform in healthy and varicocele samples

Features of Haar wavelet transform	Varicocele	Health
Average of level peak D1	0.1 ± 0.4.	0.1 ± 0.2
Average of level peak D2	0.1 ± 0.5	0.05 ± 0.2
Average of level peak D3	0.1 ± 0.4	0.01 ± 0.2

Table 2 Measured temperature (Celsius) by thermal cameras in different areas

Temperature	LP	RP	LT	RT	LTH	RTH	ΔLPRP	ΔLPLTH	ΔRPLTH
Mean	34.15	32.32	33.4	32.22	33.1	33.11	1.92	1.18	0.01
Median	34.15	32.32	33.65	32.22	33.1	33.21	1.83	1.43	0.11
Standard deviation	0.74	0.65	1.21	0.51	0.89	0.52	0.09	0.7	0.37
Minimum	33.1	31.1	31.2	31.3	33.3	32.2	2	0.1	1.1
Maximum	36	33.5	35.55	33.2	34.88	34.4	3.5	2.35	0.48

Temperature areas related to pampiniform plexus, scrotum, and thigh in a sample of patients have been presented in Table 2. In this table, L, R, P, T, TH, and Δ denote left, right, pampiniform plexus, testicle, thigh temperature, and the difference between the measured temperatures, respectively.

In general, the approximate temperature difference between patients with varicocele with their grades is shown in Table 3.

50 persons were tested and evaluated. Initially, all were tested with Doppler ultrasound. From surveyed cases, 35 patients had diseases related to testicular. The results of Ultrasound are shown in Table 4.

Then, the obtained results of ultrasound methods were compared with thermal imaging method. The Results of Testicular ultrasound of patients are shown in

Table 3 The approximate difference of temperatures among the patients with varicocle

Samples	Temperature differences in venous plexus Pampiniform
Healthy case	$\Delta T < .5$
Grade I	$0.5 < \Delta T < 0.75$
Grade II	$0.75 < \Delta T < 1$
Grade III	$1 < \Delta T$

Table 4 Scrotal abnormalities detected by scrotal ultrasonography in 35 infertile men

Properties	Number of cases (%)
Left varicocele	30(66.7)
Epididymal cyst	2(4.45)
Right varicocele	2(4.45)
Unilateral testicular cyst	1(2.22)

Table 5 Comparison of the results between diagnosis of thermography and ultrasound in grade varicocele

	Grade no (%)				
	I	II	III	No(healthy)	Total
Ultrasonography	8	12	10	15	45
Thermography detection	5	10	8	22	45

Table 6 Venous diameter of left spermatic vein

	Diameter of vein in left pampiniform plexus, mm*
Left varicocele none present (No)	3.0 ± 0.9
Grade 1	3.7 ± 1.2
Grade 2	4.1 ± 1.3
Grade 3	5.1 ± 1.5

Table 7 Identification of individuals only with the help of the proposed thermal patterns regardless of the temperature difference between the proposals

Proposed pattern	Number
Healthy pattern	22
Grade I of thermal pattern	5
Grade II of thermal pattern	7
Grade III of thermal pattern	5

Table 4 with considering their varicocele grade. Also diagnostic functions of thermography were evaluated according to presented patterns in this article. It is worth mentioning that from 50 cases, 5 patients were excluded from Examining due to cystic masses and etc. (Table 5).

The values related to venous diameter of left spermatic vein are presented in Table 6.

It is noteworthy that Thermography is only able to offer temperature changes of the skin surface. According to proposed thermal models in the paper, the results of studding these patterns can be observed in Table 7 on the collected cases.

4 Discussion

Currently, varicocele diagnosis depends on physical examination and ultra-sonography /Doppler scrotum. Physical examination is subjective and also it cannot be helpful alone in the diagnosis of subclinical varicocele [17]. Clinical experience of the examiner and the interpreter is one of the disadvantages of ultra-sonography/Doppler. In addition, the use of ultrasound in the postoperative stage as a follow-up is restricted [22]. Thermography allows imaging of the surface temperature distribution. Skin temperature depends on complex relationships of heat exchange between the skin tissue, internal tissue, and local vascular and metabolic activity. Thermography was applied in medicine for the first time in 1957 [23] and its application for diagnosis of varicocele goes back to 1970 [24]. However, at the time, thermal measurement equipment was expensive, large, and of low-resolution and also did not support the relevant software for interpretation of images. Recent developments in the field of focal thermal cameras as well as mobile software have made digital thermography to be presented as an affordable and easy method. Testes temperature is about 3 °C lower than the body temperature (37 °C) [25]. If varicocelectomy is done in the early stage of the disease and at a young age, it will produce better results in fertility [26]. Recent studies have shown that patients with the early abnormalities of semen quality are more at risk of PDSQ (Progressive Deterioration of Semen) than patients with normal primary quality of semen. In addition, varicocele patients with normal initial semen quality and higher scrotum temperature are more likely to suffer from PDSQ [27]. Although varicocele is a common condition, the infertility caused by is not much prevalent. Effects of varicocele are progressive and varicocele over time can affect sperm production and fertility, in a way that can cause azoospermia [28]. It has been also reported that the size of varicocele cannot make for the prediction of the final status of fertility and even a subclinical varicocele can be similar to a large and clinically significant varicocele in terms of damages [29]. It has been revealed that bilateral varicocelectomy significantly leads to improved sperm production compared to one-way varicocelectomy, even if there is a little varicocele on the right side [30]. The effect of varicocelectomy on male fertility has been often challenged [31]. It is really difficult and impossible to predict that who will benefit from this surgery. It would be valuable to study that whether a degree of varicocele on physical examination is always in relationship with the similar temperature increase that can be detected by thermography. In other words, is it possible to relate temperature to visible changes in semen analysis? Some authors have reported their experiences on digital thermography and briefly expressed their own diagnostic criteria for varicocele. However, we still do not have standardized diagnostic criteria and specifications for varicocele. The present study aimed to analyze the thermography images of patients with typical varicocele and try to propose elements for these criteria. The thermography defined in this study correctly confirmed the left varicocele diagnosis in all patients. Merla et al. [32] stated that temperatures above 34 °C in the pampiniform network vein and/or scrotal temperatures above 32 °C are indicative of varicocele. In the studied sample in this research, 83% of patients had a temperature above 34 °C and 92% of them showed temperatures above 32 °C. Tucker reported that retention of breath can help diagnose of varicocele, and in a normal mode (i.e. absence of venous reflux), this leads to a decrease in temperature by 0.5 °C [33]. However, no similar effect was observed in this study and further studies should be conducted to determine the usefulness of this parameter. Nogueira et al. (2009) and Yamamoto et al. stated that a temperature difference of 0.3 and 0.8 in the right and left sides of upper part of the pampiniform network are indicative of one-sided varicocele [16, 34]. It should be emphasized that this clinical sign alone cannot be taken into account in one-sided varicocele. Temperature of the scrotum skin is lower than that the upper part of the thigh [35]. In this study, the temperatures of the central part of the upper thigh was measured as the calibration temperature. In all patients, temperature of the left pampiniform network was higher than that of the upper part of the thigh. In the study conducted by Merla et al., the researcher raised the possibility for measuring the difference of temperature return speed after cooling of the scrotum [32]. The obtained Mean difference of temperature in the venous plexus (pampiniform) of patients with grade III varicoceles is 1.5 °C and in patients with varicocele grade II is almost 1 °C.

However, this potential diagnostic method was not studied in this research that was mainly due to the complexity of its practical approach. Finding at least three potential symptoms leads to the diagnosis of varicocele. During the present research, a common described pattern of thermal distribution was found in men with varicocele. In some cases, high scrotal temperature can be observed only in the venous plexus (pampiniform) and there is no increase in temperature in the testicles of the same side. In other cases, increased temperature of the venous plexus (pampiniform) expands to the testicles of the same side.

5 Conclusion

The importance of studying and research in this field is due to the fact that few studies and comparison have been conducted on varicocele diagnosis in relation to thermography. The findings from the present study show that thermography has some advantages and disadvantages in the detection of varicocele. With the advent of new generations of infrared detectors, infrared thermal imaging has become a thorough medical diagnostic tool for measuring the abnormal areas in the thermal pattern. In addition, sensitivity to temperature, spatial resolution, and its non-contact nature, and safety are some of the features of thermal imaging. Thermal images can be digitally stored and then processed using different software packages. Thermography does not provide information on morphology of testicular structures, but presents information on temperature function and vascular conditions of testicular tissue. The results of this study suggest that thermography can be used for primary diagnosis or quick screening. In addition, it can be applied as a supplement to ultrasound. In other words, although thermography can be helpful in the initial screening for determining the positive or negative result of affliction with varicocele, determining the grade of diseases requires higher accuracy and more studies. However, a pattern for grading can be reached in a large number of images. Another point is that asymmetry has a key role in early diagnosis, which can be achieved through primary settings of the camera. To achieve perfect certainty, thermal characteristics of Haar wavelet transform can be used in this research. The results can be used to identify two things in relation to varicocele: a temperature difference in venous plexus (pampiniform) and the other thermal pattern that was discussed in the paper. If the temperature difference is noticed in venous plexus (pampiniform), Varicocele can be suspected. Due to the temperature difference the thermal pattern can be related to varicocele grade. The potential defect of scrotum thermography is inseparability of varicocele from other pathological states of scrotum (such as testicular tumors and inflammation of the epididymis). Given the advances in this technology and increased demands for a low price and radiation-free method of screening, thermography has great potential for being selected as proper varicocele imaging technique. It is recommended that more studies to be conducted on larger number of patients and healthy subjects to investigate the sensitivity of thermography method and its features and also to study the diagnostic parameters for investigation of thermal measurement in varicocele detection.

Acknowledgements We are thankful for this research has been supported by AJA University of medical sciences.

References

1. Romeo C, Santoro G (2009) Varicocele and infertility: why a prevention? J Endocrinol Invest 32(6):559–561
2. Kulis T, Knezevic M, Karlovic K, Kolaric D, Antonini S, Kastelan Z (2013) Infrared digital thermography of scrotum in early selection of progressive varicocele. Med Hypotheses 81 (4):544–546
3. Namdari F, Dadpay M, Hamidi M, GHayoumi-zadeh H (2017) Evaluation of thermal imaging in the diagnosis and classification of varicocele. Iran J Med Phys 14(2):114–121
4. Goldstin M (2002) Surgical management of male infertility and other scrotal disorder, Vol. I. Campbell's urology Patrick C Walsh, Alan B Retik, Vaughan (eds) 8:313–316
5. Namdari F, Dadpay M, Hamidi M, Zadeh HG (2017) Providing a model for the diagnosis of varicocele in the scrotum thermal images. Biomed Res 28(9)
6. Said S, Aribarg A, Virutamsen P, Chutivongse S, Koetsawang S, Meherjee P et al (1992) The influence of varicocele on parameters of fertility in a large group of men presenting to infertility clinics. Fertil Steril 57(6):1289–1293
7. Scaramuzza A, Tavana R, Marchi A (1996) Varicoceles in young soccer players. Lancet 348 (9035):1180–1181
8. Witt MA, Lipshultz LI (1993) Varicocele: a progressive or static lesion? Urology. 42(5):541–543
9. Sharlip ID, Jarow JP, Belker AM, Lipshultz LI, Sigman M, Thomas AJ et al (2002) Best practice policies for male infertility. Fertil Steril 77(5):873–882
10. Cozzolino DJ, Lipshultz LI (2001) Varicocele as a progressive lesion: positive effect of varicocele repair. Hum Reprod Update 7(1):55–58
11. Evers J, Collins J, Clarke J (2008) Surgery or embolisation for varicocele in subfertile men. Cochrane Database Syst Rev 3
12. Watanabe Y (2002) Scrotal imaging. Curr Opin Urol 12(2):149–153
13. Ng E-K (2009) A review of thermography as promising non-invasive detection modality for breast tumor. Int J Therm Sci 48(5):849–859
14. Kulis T, Kolaric D, Karlovic K, Knezevic M, Antonini S, Kastelan Z (2012) Scrotal infrared digital thermography in assessment of varicocele–pilot study to assess diagnostic criteria. Andrologia 44(s1):780–785
15. Gat Y, Gornish M, Chakraborty J, Perlow A, Levinger U, Pasqualotto F (2010) Azoospermia and maturation arrest: malfunction of valves in erect poster of humans leads to hypoxia in sperm production site. Andrologia 42(6):389–394

16. Yamamoto M, Hibi H, Hirata Y, Miyake K, Ishigaki T (1996) Effect of varicocelectomy on sperm parameters and pregnancy rate in patients with subclinical varicocele: a randomized prospective controlled study. J Urol 155(5):1636–1638
17. Gat Y, Bachar GN, Zukerman Z, Belenky A, Gorenish M (2004) Physical examination may miss the diagnosis of bilateral varicocele: a comparative study of 4 diagnostic modalities. J Urol 172(4):1414–1417
18. Zadeh HG, Haddadnia J, Seryasat OR, Isfahani SM (2016) Segmenting breast cancerous regions in thermal images using fuzzy active contours. EXCLI J 15:532
19. Collett AE, Guilfoyle C, Gracely EJ, Frazier TG, Barrio AV (2014) Infrared imaging does not predict the presence of malignancy in patients with suspicious radiologic breast abnormalities. Breast J 20(4):375–380
20. Nicandro C-R, Efrén M-M, María Yaneli A-A, Enrique M-D-C-M, Héctor Gabriel A-M, Nancy P-C, et al (2013) Evaluation of the diagnostic power of thermography in breast cancer using bayesian network classifiers. Comput Math Methods Med 2013
21. Agarwal P, Prakash N (2013) An efficient back propagation neural network based face recognition system using haar wavelet transform and PCA. Int J Comput Sci Mobile Comput (IJCSMC) 2(5):386–395
22. Cvitanic O, Cronan J, Sigman M, Landau S (1993) Varicoceles: postoperative prevalence–a prospective study with color Doppler US. Radiology 187(3):711–714
23. Lawson R (1957) Thermography; a new tool in the investigation of breast lesions. Can Serv Med J 8(8):517
24. Kormano M, Kahanpää K, Svinhufvud U, Tähti E (1970) Thermography of varicocele. Fert Steril 21(7):558
25. Mieusset R, Bujan L (1995) Testicular heating and its possible contributions to male infertility: a review. Int J Androl 18(4):169–184
26. Shin JW, Kim SW, Paick JS (2005) Effects of varicocele treatments in adolescents: changes of semen parameters after early varicocelectomy. Korean J Urol 46(5):481–486
27. Chen S-S, Chen L-K (2012) Risk factors for progressive deterioration of semen quality in patients with varicocele. Urology 79(1):128–132
28. Poulakis V, Ferakis N, De Vries R, Witzsch U, Becht E (2006) Induction of spermatogenesis in men with azoospermia or severe oligoteratoasthenospermia after antegrade internal spermatic vein sclerotherapy for the treatment of varicocele. Asian J Androl 8(5):613–619
29. Dhabuwala C, Hamid S, Moghissi K (1992) Clinical versus subclinical varicocele: improvement in fertility after varicocelectomy. Fertil Steril 57(4):854–857
30. Elbendary MA, Elbadry AM (2009) Right subclinical varicocele: how to manage in infertile patients with clinical left varicocele? Fertil Steril 92(6):2050–2053
31. Ficarra V, Cerruto MA, Liguori G, Mazzoni G, Minucci S, Tracia A et al (2006) Treatment of varicocele in subfertile men: the Cochrane review–a contrary opinion. Eur Urol 49(2):258–263
32. Merla A, Ledda A, Di Donato L, Romani GL (2004) Assessment of the effects of varicocelectomy on the thermoregulatory control of the scrotum. Fertil Steril 81(2):471–472
33. Tucker AT (2000) Infrared thermographic assessment of the human scrotum. Fertil Steril 74 (4):802–803
34. Nogueira FE, das Chagas Medeiros F, de Souza Barroso LV, de Paula Miranda E, de Castro JD, Mota Filho FHA (2009) Infrared digital telethermography: a new method for early detection of varicocele. Fert Ster 92(1):361–362
35. Pochaczevsky R, Lee W, Mallett E (1986) Management of male infertility: roles of contact thermography, spermatic venography, and embolization. Am J Roentgenol 147(1):97–102

A Review of Feature Selection Methods with the Applications in Pattern Recognition in the Last Decade

Najme Ghanbari

Abstract The present study is a review of recently-done research (in the past 10 years) on the feature selection methods and a set of the applications in pattern recognition. The study aimed to introduce the latest research on the feature selection methods and applications. The study findings can be the basis for further and more practical research in this field. Significant advances have been made in the last decade. Particularly in recent years, the evolutionary algorithms related to random methods were widely used to solve feature selection problems.

Keywords Feature selection · Feature selection methods · Applications of feature selection

1 Introduction

Feature selection is an important topic in modeling, classifying, discovering knowledge and data mining in a variety of fields such as signal processing, computer vision, statistics, neural networks, pattern recognition and machine learning. Due to the speed of data collection, feature selection problem is one of the most important factors in reducing or increasing the recognition rate. The feature selection process consists of four basic steps as follows.

Step 1: Finding a starting point to create a subset of features. Usually, researchers use three starting points in their research; (a) Begin with an empty set, (b) Begin with a full set, and (c) Begin with a subset of features selected randomly.

Step 2: Applying one of the feature selection methods. The feature selection methods are divided into three main categories: complete methods, heuristic methods, and random methods.

N. Ghanbari (✉)
Department of Electrical Engineering, Faculty of Engineerging,
University of Zabol, Zabol, Iran
e-mail: najme.ghanbari@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_12

Complete methods are categorized into exhaustive or non-exhaustive methods. Exhaustive methods can find the optimal subset by creating and controlling all candidate subsets. Such methods are used in cases where time is not a problem and the total size of the good set of features is small. In non-exhaustive methods, the size is larger, i.e. more features.

In heuristic techniques, the number of optimal features that can be selected is predefined. Adding more features will increase the classification error. In addition, the more information we have about the problem, the less error we will face. Sequential search methods and principal components analysis (PCA) are among heuristic techniques. Although such methods may not be able to find minimum size subsets, they can find subsets with a size close to the minimum in less time, in cases where the total number of features and the number of good features are large.

In heuristic techniques, a candidate subset is created randomly, and a supervised instruction is applied during the search for logical leaps in searching other regions of the feature space. There is no guarantee for obtaining the subset of the optimal features through such techniques. Evolutionary algorithms are an example of random methods.

Although the complete methods guarantee that the optimal candidate feature subset is obtained, high costs are required to implement such methods, which itself requires high computational complexity. On the other hand, the size of databases today is large in bioinformatics. Therefore, the complete methods are used less commonly. In contrast, the random and exploratory methods have been widely used, despite the lack of any guarantee for optimality.

Step 3: Applying an evaluation strategy for feature selection. There are two methods for the evaluation strategy.

1. A filter approach
2. Dubbed the wrapper

Step 4: Determining stopping criterion or condition. There are different stopping criteria. For example, a criterion could be that the number of predetermined features to be selected has already been obtained or the maximum number of iterations as another criterion.

Feature selection can result in several concurrent improvements. For example, removing inappropriate features that cause a deterioration of the recognition rate will increase computational efficiency. Another example, feature selection increases the generalization power and classification accuracy by reducing the dimension of the input vector. The selection of appropriate features also reduces the computational and convergence time during the training procedure. Section 2 is an overview of some of the feature selection methods proposed for pattern recognition over the last few years. There is an overview of some of the applications of feature selection in pattern recognition in the last 10 years in Sect. 3. And, Sect. 4 includes the conclusion.

2 Some of the Feature Selection Methods Proposed for Pattern Recognition Over the Last Few Years

In order to solve the feature selection problem, a modified feature selection criterion was proposed using fuzzy logic [1]. In this method, the number of features was defined as a fuzzy number and the fuzzy criterion was obtained by applying the principle of fuzzy development. This method was implemented and evaluated on 6 datasets of UCI machine learning. The results confirmed its effectiveness. In this method, fewer features are obtained with higher classification accuracy than the previous methods or with almost the same.

An improved version of the binary ant optimization algorithm was introduced to solve the feature selection problem [2]. This algorithm presents the characteristics of both the Ant Algorithms for Discrete Optimization and Binary Ant Algorithm. It is worth noting that the feature selection problem had already been solved separately by both algorithms, and the combined method proposed solved the previous weaknesses. This method was implemented on 12 standard classification (UCI) databases, which were varied in number of samples, features, and classes. And, the evaluation results confirmed the proper performance of the algorithm. K-Nearest Neighbors Classifier was used for classification (k = 1).

A classification-based selection feature method was proposed using the Genetic Algorithm (GA) and Decision Tree (DT) [3]. This method was implemented on the polarimetric features of the land cover classification (RADARSAT-2 images, San Francisco). In addition to being highly precise in distinguishing between urban classes and vegetation, it produced high speed performance. The parameters extracted from the distribution matrix, covariance, and coherence, as well as the parameters extracted from the target analysis methods were considered features. Furthermore, the feature selection was done using the GA-SVM algorithm for comparison. The results indicated that the GA-DT algorithm was more efficient in distinguishing "urban areas and vegetation" and the GA-SVM algorithm in distinguishing "urban and water areas".

Several filter and wrapper based feature selection methods, including were reviewed [4]. It was mentioned that evolutionary algorithms have been widely used to resolve the feature selection problem in recent years. In the reference, seven examples of these algorithms were studied, i.e. ant colony algorithm, genetic algorithm, firefly algorithm, bee colony algorithm, harmonic search algorithm, cuckoo algorithm, and colonial competition algorithm.

The filter and wrapper feature selection methods were described [4]. Moreover, the feature selection problem was solved and evaluated using the aforementioned seven methods. However, it cannot be claimed that a particular evolutionary algorithm is more efficient for all problems. In other words, any algorithm can be used according to the intended project.

For the first time, a hyper-heuristic approach was proposed to select optimal features among all the features for classification [5]. According to this approach, the search space is efficiently searched by applying a few local searches, each of which is neighborhood structures and/or simple local searchers. Each part of the search space had its own features. In the search path, a local search should be selected and applied in the current solution. This selection is made by an observer based on the history of local search performance. This method represents a good compromise between search and productivity, unlike the existing heuristic algorithms. In this method, the genetic algorithm was used as an observer and the 16 other heuristic algorithms were used for local search. The method was evaluated on the UCI databases with very good results.

3 The Applications of Feature Selection in Pattern Recognition in the Last Decade

The goal was hyperspectral image classification using Support Vector Machine (SVM) [6, 7]. SVM gives good stability in high-dimensional spaces. For a better SVM performance in this article, (a) the parameters of the support vector machine are optimally determined, and (b) an optimal feature subset is selected from all the input features. One of the most powerful methods to accurately and rapidly classify the hyperspectral images using SVM is feature selection by which noise and irrelevant bands are eliminated. In 1 and 2, hyper-heuristic algorithms were used. In feature selection using Genetic Algorithm, Data Envelopment Analysis (DEA) was used. In order, two criteria were considered to evaluate the quality of the selected features, i.e. classification accuracy and the number of selected features. In other word, the aim was higher accuracy and fewer features. By applying 1 and 2 simultaneously, the classification accuracy increased by 5 and 15% for the data obtained by AVIRIS hyperspectral sensor for Gaussian and Polynomial kernel. In addition to the genetic algorithm, a simulated annealing algorithm for gradual optimization was used. And, the results showed that genetic algorithms, especially as the search space grows, produces a better performance.

An intelligent feature selection method was proposed to solve the recognition problem of Persian handwritten digits [8]. In this method, a binary gravitational search algorithm was used. This algorithm minimally optimizes the fitness function, which is the recognition system error. Accordingly, the features that affect the increase of the recognition rate are selected and applied, and other features are eliminated. In addition to reducing the number of features and computational rate, the method increases the recognition rate dramatically. The classifier used was a simple fuzzy method and the features used for digits were the zooning features.

An application of feature selection from DNA microarray data that play a key role in diagnosis of cancerous tissues was presented [9]. The microarray data are a matrix with thousands of columns and a few hundred rows, i.e. each column represents a gene and each row a sample. The high-dimensional feature vector and the small number of samples caused problems in analyzing the data such as increased classifier complexity, reduced ability to generalize classifiers, and their reduced credibility for predicting new samples. Therefore, reducing the number of genes (the selection of distinctive genes) is an important step in data analysis. An effective gene selection method can greatly improve the classification and diagnosis of cancer. The microarray data are pre-processed and their information-less genes are identified and discarded. The method proposed in this paper was a combination of the Binary Particle Swarm Optimization (BPSO) algorithm and the Bayesian Linear Discriminate Analysis (BLDA). Effective genes were selected by BPSO and BLDA evaluated the quality of the subset of genes selected by the PSO. The algorithm was implemented on four datasets of cancer database and was able to find a small subset of genes with information so that classification accuracy increased significantly. The results were evaluated based on the 10-fold cross-validation method.

In order to increase the recognition rate in the power quality disturbance classification, the feature selection was used through Gram-Schmitt method [10]. A combination of Hyperbolic S-transform and Wavelet Transform was used for feature extraction. SVM Multi-Class Classification was also used. The SVM parameters were optimized using PSO. For the classification, there were 6 single disturbances, 2 combined disturbances, and normal mode. Evaluations were carried in different noise conditions with different levels of signal with noise.

The problem of land cover classification, especially in urban areas, has been improved with feature selection [11]. Land cover classification is one of the applications of radar polarimetric imaging. Feature extraction is organized into three groups of main data features, target analysis features, and SARs. As a search tool, Non-dominated Sorting Genetic Algorithm II (NSGA-II, a multi-objective optimization algorithm), two SVM classifiers, and Adaptive Network-based Fuzzy Inference System (ANFIS) were used for classification. The proposed algorithm was implemented on the images of RADARSAT-2 (located in San Francisco) and its effectiveness was confirmed.

A method was presented for solving the problem of representing the behavior of the brain in the state of visual selective attention, in which optimal feature selection was also used [12]. This is based on Event-Related Potentials (ERP). Feature extraction was done using wavelet coefficients (12 features) and scaling coefficients (18 characteristics). Optimal features were selected by P-value and Dispersion Index, and classification by using SVM with Gaussian and Polynomial kernels. To evaluate the proposed algorithm, 5-fold cross-validation was used. The maximum resolution between responses ranged from 100 to 400 ms created using stepwise discriminant analysis. With Dispersion Index, and SVM with Gaussian kernel, both target and non-target classes were distinguished with high accuracy (86.7%). The highest accuracy was related to parietal and temporal regions. The database used included 250 channels of brain signals recorded by ActiveTwo (BioSemi). The data were sampled at 256 Hz and recorded using a 24-bit analog-to-digital converter.

The genetic algorithm was used to select the optimal features in implementing an effective validation model for bank customer to provide credit facilities appropriate to each class [13]. Decision trees were created to validate bank customers. Genetic algorithms was used to select optimal feature and create the decision trees.

The development process is also used in pattern recognition and CRISP. In addition to decision trees, feature selection, and genetic algorithm, clustering was also used. In this paper, the classification accuracy was higher than of the similar studies, and also the number of leaves and size of the decision tree were smaller with less complexity.

An application of feature selection in bioinformatics proposed was the selection of single-nucleotide polymorphisms (SNPs) [14]. The SNPs provide useful information for the detection of genes associated with diseases such as cancer, cardiovascular disease, diabetes, and mental diseases. Two important steps were proposed to select important SNPs in the databases. In the first step, Correlation based Feature Selection was used to reduce the number of features. The selected features were used as inputs in the second step, which is the neural network. A genetic algorithm was used to search the problem space in both steps. The final features were the most important SNPs in the database.

An application of feature selection was proposed in the Intrusion Detection System (IDS) [15]. Feature selection in intrusion detection systems improves accuracy and speed. Using scattered search and some criteria, including linear correlation and intraclass correlation, the interclass correlation, the feature selection problem was solved step by step. The method implemented on NSL-KDD dataset, increased the recognition rate and eliminated many extra and non-efficient features.

A brain-computer interface is a system that can directly communicate between the brain and the outside world without the help of muscles. The basis of the work is the recognition of the individual's movements that can be used to help patients who have lost their physical and muscular abilities. A method for feature selection in brain-computer interfaces is presented using a new evolutionary algorithm called ProbPSO [16]. This algorithm is a combination of particle group accumulation and distribution estimation algorithms for feature selection. It can also be called class-based feature selection. In this paper, combinatorial classifiers were used instead of a separate classifier. This method was implemented on data from the 2005 Brain-Computer Tournament, and the classification accuracy increased by 20% compared to when the feature selection was not considered.

To predict the direction of changes in the index of the most active companies listed in the Tehran Stock Exchange, a combined feature selection method was used [17]. The classifier used was K-NN. The proposed feature selection algorithm was a combination of principal components analysis and genetic algorithm. The advantage of this combined algorithm is the use of both filter and wrapper methods in selecting the optimal subset of the features. This method produced a higher accuracy than similar methods. Different feature selection methods were described and presented in the form of a diagram in terms of three types of producer functions [18] (Fig. 1).

An efficient feature selection algorithm was developed for the recognition of breast cancer in diagnostic systems, which has used fewer features than similar systems in addition to a recognition accuracy of 100% [19]. In this method, the WDBC database features were used, which included 569 FNA samples. BPSO was

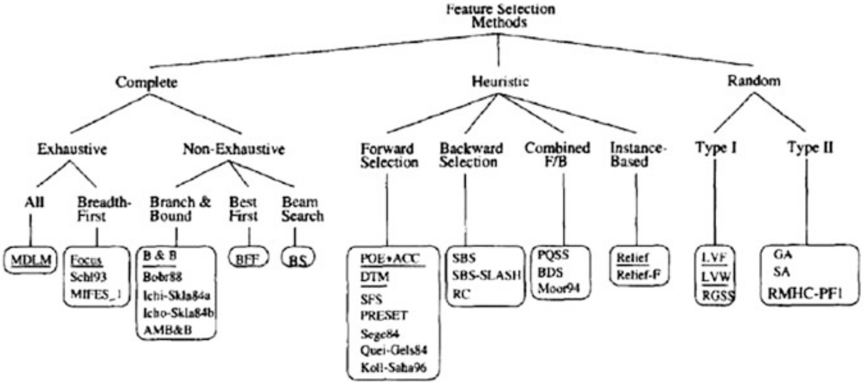


Fig. 1 Feature selection methods in a category

used for feature selection and SVM for classifier. The recognition rate of 100% was obtained using 28 features in the form of 5 SVM models.

Two new class-dependent feature selection methods were presented for Persian handwritten digits recognition [20]. These methods were implemented on a Persian Optical Character Recognition system. Four different feature classes (with 65, 129, 193, and 257 features) and three different classifiers (1NN, 7NN, and SVM) were used and the results indicated that the proposed class- dependent methods increased classification accuracy by up to 7.9% compared to similar non-dependent methods.

Optimal spectral and texture features were selected for analyzing multi-time remote sensing images using the genetic algorithm and Bayesian classifier [21]. The evaluation of the changes in Sahand City (in the northwest of Iran) were carried by IRS-P6 and Geo-Eye1 imaging screenshots collected on 14th July, 2006 and 1st September 2013, were reviewed. The implementation was done by MATLAB R2013a. The study revealed that texture features as a complementary source of information improved the recognition of changes in urban areas. The study results also showed that the feature selection produces an acceptable performance in recognition of changes based on spectral and texture features. The proposed method in this paper exhibited better results than the two commonly used techniques, i.e. PCA and SAA. Kappa coefficient and overall accuracy of the map of changes also increased.

4 Conclusion

The research reviewed several articles on new feature selection methods (5 items), as well as several articles on the applications of feature selection in classification problems (15 items). In most of the papers that have been published in recent years, evolutionary algorithms were used that are considered random feature selection techniques. Most of the algorithms yielded good results. That which evolutionary algorithm is better option to use in order to achieve the desired result depends on the experience of the researchers and a particular one may not be an always good algorithm for all problems. The evolutionary algorithms in feature selection proved to be very effective in all disciplines, such as electronic, computer, medicine, geography, management sciences etc. Even in one case, a hyper-heuristic algorithm, which was proposed by using several evolutionary algorithms (17), was used for feature selection and good results were obtained.

References

1. Nosrati Nahouk H, Eftekhari M (2013) A new method for fuzzy logic based feature selection. Intell Syst Electr Eng 4(1):71–83
2. Kashef Sh, Nezamabadipour H (2015) A new version of binary ant algorithm to solve feature selection problem. Iran J Electr Eng Comput Eng 12(2):127–134
3. Khosravi A, Mousavi MM, Amini J (2015) A feature selection method based on genetic algorithm and decision tree for the classification of radar polarimetric images. Eng J Geospatial Inf Technol 3(2):75–88, K. N. Toosi University of Technology
4. Shabani R (2015) An overview of evolutionary algorithms to solve the feature selection problem. In: Third national conference on new ideas in electrical engineering, Isfahan Islamic Azad University, Khorasgan Branch
5. Montazeri M, Soleimani Baghshah M, Niknafas AA (2011) Finding effective features using a hyper-heuristic approach, Master's thesis, Shahid Bahonar University of Kerman, Faculty of Computer Engineering
6. Samadzadegan F, Hassani HA (2012) Optimal support vector machines in the classification of hyperspectral imaging based on genetic algorithm. J Inf Commun Technol 4(13–14):9–24
7. Seryasat OR, Aliyari-shoorehdeli M, Honarvar F (2010) Multi-fault diagnosis of ball bearing based on features extracted from time-domain and multi-class support vector machine (MSVM). In: IEEE international conference on systems man and cybernetics (SMC), pp 4300–4303
8. Ghanbari N, Razavi SM, Nabavi Karizi SH (2011) An intelligent feature selection method based on binary gravitational search algorithm in the Persian handwritten recognition system. Iran J Electr Eng Comput Eng 9(1):29–36
9. Joroughi M, Shamsi M, Saberkari HR, Sedaghi MH, Momen Nejad A (2014) Gene selection and classification of cancer cells based on microarray data using BPSO and BLDA combined algorithm. J Smart Syst Electr Eng 5(2):29–45
10. Hajian M, Akbari Foroud A (2013) A new design for automatic classification of power quality disturbances based on signal processing and machine learning tools. Iran J Electr Eng Comput Eng 12(1):1–13
11. Salehi M, Maghsoudi Y, Sahebi MR (2014) Improvement of Urban area classification with radar polarimetric data and multi-objective optimization methods. Radar Mag 1(2):45–56
12. Akbarzadeh Totounchi MR, Hosseini SA, Naghibi Sistani MB (2016) Evaluation of visual selective attention by analyzing event-related brain potentials. J Electr Eng, Univ Tabriz 46(75):13–24
13. Alborzi M, Mohammad Pourzarandi MA, Khan Babaei M (2010) The application of genetic algorithm in optimizing decision trees for validating bank clients. J Inf Technol Manage 2(4): 23–38

14. Sanat Nama H, Esmaeilizadeh Kashkuiyeh A, Holaku F, Eftekhari M (2011) Using a combination of correlation coefficients, neural networks, and genetic algorithms for selection of single-nucleotide polymorphisms (SNPs). In: 7th national conference on biotechnology in the Islamic Republic of Iran, Tehran, Power Research Institute
15. Ostovar S, Nasser Sharif B (2014) Selection of feature subset using distance and correlation based criteria for intrusion detection system. In: 1st national conference on industrial mathematics, Tabriz
16. Sarabian Moghaddam A (2014) Feature selection and reducing data dimensions in brain-computer interfaces using an evolutionary algorithm. Master's Degree in Computer Science, Tabriz University of Technology
17. Pouyan Far A, Fallahpour S, Noruzi Jan Lekwan A, Farhadi Shooli OH (2016) Using hybrid feature selection method and nearest neighboring algorithm to predict the direction of the daily movement of the index of 50 most companies listed in Tehran stock exchange. J Fin Eng Manage Securities 25
18. Hatami Khah N (2013) Exploring methods based on feature selection. Malek Ashtar University of Technology, ICT Complex
19. Alipour M, Haddadnia J (2009) Introduction of an intelligent system for the precise diagnosis of breast cancer. Iran J Breast Dis 2:2
20. Nematollahi MM, Mahmoudi H, Kazimi Pour B, Salehi B (2011) "Two new class dependent feature selection methods and their application in Farsi handwritten digits recognition", Conference
21. Sadeghi V, Enayati H, Ebadi H (2017) Improving the detection of changes in Urban areas by selecting optimal spectral and spatial features based on genetic algorithm. J Geogr Inf 24:96

A Review of Research Studies on the Recognition of Farsi Alphabetic and Numeric Characters in the Last Decade

Najme Ghanbari

Abstract The present paper is a review of research on the recognition of Persian alphabetic and numeric characters in the last ten years. The goal is to provide researchers with the latest research findings for further research after discovering the gaps that still exist in this field, i.e. Optical Character Recognition (OCR) in the Farsi language. However, there has been a dramatic improvement in this field in the last decade as it will be discussed. Although the researchers have achieved very good results, there is still a gap due to the lack of commercial and functional software.

Keywords Farsi alphabetic characters recognition · Farsi numeric characters recognition · Handwritten alphabetic and numeric characters · Printed alphabetic and numeric characters

1 Introduction

We humans need to learn how to work with computers as a tool, and provide facilities for working with computer software programs. One of the most important type of such programs includes Optical Character Recognition (OCR) for Farsi characters, literally meaning converting a printed or handwritten text within an image into a machine-readable text image. The image will be obtained by optical scanning or digital imaging.

Over the past few decades, the recognition of written patterns that mostly include printed or handwritten alphabetic and numeric characters attracted many researchers. The first efforts date from the early 1970s. Although extensive research led to the development of many efficient methods for importing information in documents, books, etc. into computers, most of the methods are still not fully functional. Filling this gap requires greater effort.

N. Ghanbari (✉)
Department of Electrical Engineering, Faculty of Engineering,
University of Zabol, Zabol, Iran
e-mail: najme.ghanbari@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_13

Various fields such as digital signal processing, image processing, machine vision, machine learning, fuzzy logic, statistics, neural networks et cetera are used in character recognition. There are many applications of alpha-numeric character recognition, such as postal address recognition, license plate recognition, automatic bank cheque processing, and security applications like passport authentication, etc.

An overall categorization for OCR systems in terms of the type of input pattern is as follows: 1. Recognition systems for printed texts. 2. Recognition systems for handwritten texts.

Today, traditional writing tools like paper and pen are used more commonly than keyboards and computers. In other words, information is usually handwritten. Therefore, the digitization of such information requires handwriting recognition (HWR).

In a different categorization, OCR systems are divided into two categories: 1. online systems, and 2. offline systems.

The application of online recognition is in handwriting writing, and offline recognition in printed and handwritten texts. The input to the offline systems is scanned images of texts, and to the online system is the coordinates of the pen movement points by a digitizer pen and tablet.

In the offline method, there are location information and highlighted image points. In the online method, the sequence of the line parts written by the user is accessible (the two-dimensional coordinates of a point sequence in writing are recorded as a continuous function of time). That is, the online method includes space-time representation.

Today, the demand for online systems has increased due to; (1) commercial tools used to communicate with users through pressure-sensitive screens instead of keyboards (such as PDA and Tablet PC, (2) writing being simpler than typing, (3) typing being impossible in some cases, (4) being difficult to type characters, and (5) lack of a full keyboard on small computers.

The main stages of a numeric character recognition system are presented in Fig. 1.

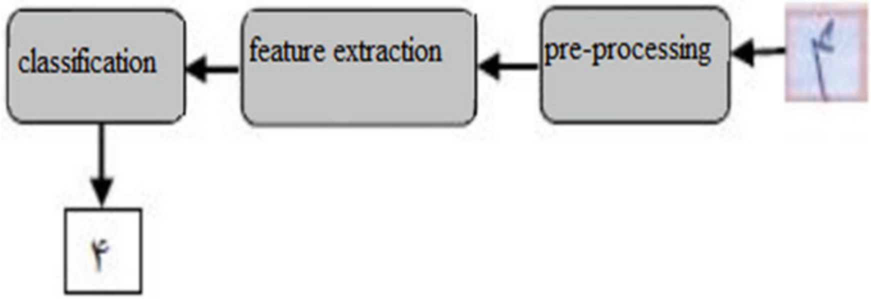


Fig. 1 The main three stages of the handwritten numeric character recognition system

2 A Review of the Research on Farsi Numeric Character Recognition Methods

A method was presented for recognizing Farsi handwritten numeric characters that is adaptable to rotation and scale variation of characters to an acceptable extent [1]. This method was implemented on a database of 8600 numerals, i.e. 860 samples for each of the digits 0–9. In this method, 30% of the numerals were randomly rotated at different angles ranging from 10 to 40 degrees clockwise or counterclockwise. The recognition rate obtained was not reduced significantly compared to the non-rotational state. In this method, k-means clustering method was first used and then fuzzy SVM for classification. Feature extraction was carried using two methods, i.e. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA).

In order to improve the recognition rate, the images of the numerals were first upgraded in the preprocessing stage and then the slope in the image was corrected [2]. The database used included 4096 training numerals and 1532 test numerals written by 500 people in a number of forms. The preprocessing increased the recognition rate by 3.3%. In this method, intersection features, chain codes, and a SVM classifier were used.

Gradient histogram features and the developed characteristic locus method were used in combination [3]. The SVM classifier was also used. In this method, the features that affect the recognition rate among all extracted ones were selected using an Improved New Binary Particle Swarm Optimization (INBPSO) algorithm in order to improve the recognition rate. The extracted initial feature vector included 400 features which were reduced to 64 after feature selection. HODA Farsi Digit Dataset was used, which is a large dataset of handwritten Farsi digits. It consists of 102,352 digits, including 60,000 training, 20,000 test digits, and 22,352 remaining samples. In this methods, a good recognition rate is 99.40%.

In order to improve the recognition rate, effective features were selected among all the features [4]. For feature selection, an evolutionary genetic algorithm was used. In this method, the gradient features and multi-layer perceptron classifier (MLPC) were employed. A good recognition rate of 98.85 was obtained using the HODA dataset. In this research, 40,000 training samples and 20,000 test samples were used.

In order to increase the speed of feature extraction and the recognition rate, two new features called improved gradient and gradient histogram were introduced [5]. The two new features are based on the gradient feature for brightness and for the two-level and gray-scale images. Using a neural network classifier and the HODA dataset, a good recognition rate of 99.02 (improved gradient) and 98.80 (gradient histogram) was obtained. The feature extraction speed was increased 2 and 10 times for the gradient feature and gradient histogram compared to the brightness gradient, respectively.

In another method, binary SVM classifier and the HODA dataset were used for handwritten Farsi digit recognition [6]. In SVM, the One-vs-All (OvA) strategy was adopted. The feature extraction was performed using a two-dimensional wavelet transform and then feature reduction by PCA. A good recognition rate was obtained, 91.75.

Handwritten Farsi digit recognition was done using the features extracted from the image gradient [7]. In the method, the image was first normalized and the gradient was calculated. After that, the gradient angle was calculated for each of the image points, and converted to 4 or 8 standard angles. Using a gradient image, 4 or 8 separate images were created. Each of the images includes the gradient related to each angle. Hidden features were extracted by sampling the above images. The machine vector support was used as a classifier. In the SVM, One-vs-the-Rest (OvR) method was employed. A good recognition rate of 99.59 was obtained by using a database of 3939 test digits. This recognition rate was related to the 8-way gradient using the RBF kernel for the SVM classifier. The database used included 4974 training digits. However, the initial numbers were 5000 and 4000 for training and test samples, respectively. And, they were again reduced to 4974 and 3939 after the removal of samples that were badly written and difficult to recognize even by human. The samples had been written in forms randomly distributed among 90 high school and university students. The samples written by an individual were only in one of two sets, i.e. training or test set.

A method was proposed to improve the fuzzy recognition of handwritten Farsi digits. The initial fuzzy method was based on a fuzzy rule for each digit (a total of 10 simple fuzzy rules) [8]. In the improved fuzzy method, the training data for each digit were split into several clusters, and a fuzzy rule was extracted according to the data of each cluster. The number of optimized clusters was selected randomly by the particle swarm optimization algorithm. A good recognition rate obtained was 96.20. In this article, the HODA dataset was also used. The features used in this method were zoning, characteristic location, and Zernike moment.

A very large dataset called HODA was presented for handwritten Farsi numeric and alphabetic characters recognition [9]. This dataset was extracted from about 11,942 registration forms filled out by undergraduate and postgraduate students. The degree of sample resolution was 200 DPI (dots per inch). Using the forms, two large databases for handwritten Farsi numeric and alphabetic characters that have been used widely. The total number of digits is 102,352, of which 60,000 are training samples, 20,000 test samples, and 22,352 remaining samples. The remaining samples can be used in various problems.

A new method was introduced for improving the recognition rate of handwritten Farsi digits using a combination of different classifiers [10]. Digit recognition is a ten-class problem, which was converted into ten simple two-class problems. Each two-class classifier distinguishes one digit from others. It then recognized by using the combining rule on maximum final output. The database used included 8600 samples, and the recognition rate obtained was 96.3% using 600 test samples.

In order to improve the recognition rate, another method was based on the selection of effective features among all other features [11]. This will increase the recognition rate and reduce computational costs. Population-based algorithms, including binary particle swarm and genetic algorithms, were used for feature selection. The classifier used was a simple fuzzy classifier with no pre-processing and post-processing. The HODA dataset was also used for system evaluation.

A novel smart handwritten Persian digit recognition method was proposed [12]. Using the smart method in the feature selection problem, the recognition rate was increased to an acceptable extent. A fitness function, optimized and minimized by Gravitational Search Algorithm (GSA), was the number of fuzzy classifier errors. The good recognition rate obtained was 84.55% for test digits and 90.01% for training digits, without any preprocessing and post-processing.

To improve the recognition rate, a combination of descriptors HOG and LBP was used [13]. One advantage was that the information and features related to the image texture had been recorded. In addition, the length of the feature vector was short with high computation speed. Using the HODA dataset, the recognition rate was 99.3%.

There is another handwritten Farsi digit recognition method [14], based on which point, local, and global features are extracted for recognition. The SVM classifier was used with the One-vs-All (OvA) approach. Using the TMU-Online database, a recognition rate of 95.99% was obtained.

Two methods for improving the recognition rate in handwritten Farsi digit recognition were presented [15]. In the first method, superior features were selected using the Binary version of the Gravity Search Algorithm (BGSA) from all the extracted features for. In addition to the reduced computation speed, the recognition rate was increased. In the second method instead of feature selection, one weight was assigned to each feature using Real-valued Gravity Search Algorithm (RGSA). A new feature vector was obtained by multiplying the initial vector multiplication by the weight vector. The fitness function was the number of simple fuzzy classifier errors in both of the methods. And, the goal was to minimize this function to increase the recognition rate that was significantly increased using the two methods.

3 A Review of the Research on Farsi Alphabetic Character Recognition Methods

A relatively new handwritten Farsi distinct letter online recognition method was proposed [16]. This method has a pre-processing and post-processing stage. This method included preprocessing and postprocessing stages. In the preprocessing stage, the dimensions of the extracted features were equalized. The recognition stage consisted of two steps. In the first step, the main body of the input letter was assigned to one of the 18 groups of the main body of letters. And in the second step, the final letter was recognized based on the location, shape, and number of micro movements, such as point etc. For example to recognize the letter “ت”, the body group “ب،پ،ت،ث” was first determined, and then it was recognized for the “two points” micro movements above the letter. The classification was performed using support vector machine. In the postprocessing stage, the possible errors of the previous steps were corrected and the recognition rate was increased by matching the information of the main body and micro movements. For example, if the classifier detected the letter “ل” with a point above, the system correct and convert it to the letter “ن” in the postprocessing stage. The database used was the Online-TMU dataset. This dataset is provided by the Electrical Engineering Department, Tarbiat Modares University [17]. 70% of the samples were used in the training and 30% in the test stage. The best recognition rate obtained through this method was 98%. The implementation was done using the MATLAB software and the package LIBSVM as a library for support vector machines.

A simple method for recognizing distinct printed Farsi letters was proposed [18]. Through the method, the letters were divided into nine groups based on points and signs up or down. Using the neural network classifier, the points and signs were recognized and the corresponding group was recognized. At this stage, three simple features were used for feature recognition (the ratio of black to white pixels of the symbol frame, the ratio of the length to width of the symbol frame, and the number of horizontal crossings through the black parts of the symbol frame). If the recognized group includes just one letter, the same letter is assigned to the unknown letter, otherwise the minimum distance classifier would be used, and the final letter was recognized by comparing the body of the unknown letter with the body of letters from the same group. At this stage for recognition, the characteristic location features were extracted. To test the recognition system, fonts like Mitra, Lotus, Zar, and Nazanin were used in different sizes. The recognition rate was 100% using this method.

The classifier combination was used to improve the handwritten Farsi letter recognition rate [19]. The classifier combination was used with the purpose of creating different data for the training process to obtain a different classifier by dividing the input features into each step and to obtain better results by combining the results obtained by the classifiers.

In the method, the input data was first randomly divided into several classes, Principal Component Analysis was implemented on each class, and the features were extracted. By combining these features, the final feature vector was created and the training process was carried out using the SVM classifier. The method was evaluated by 10 datasets. Each dataset included 3200 samples of handwritten Farsi letters (100 samples per letter). 70 samples of each letter (a total of 2240) were used at the training stage, and 30 samples of each letter (a total of 960) were used at the testing stage. In this method, there was the preprocessing stage. Each image had a white background with letters in the middle. To enhance the quality, the images were converted to binary images. To reduce the computational load, the letters were separated from the background, and all the images were normalized to a standard size.

The advantage of this method over the combined methods, is the dispersion of samples at each stage and the increased accuracy of the base classifier. However, the runtime is longer than of other methods. And, the increased recognition rate was 82.51%.

A method for the online recognition of separate handwritten Farsi letters was presented [20]. In this method, the main body and letter micro movement information were used simultaneously. The letters were divided into 18 groups based on the similarity of the main body and into 11 groups based on the similarity of the micro movements. In this method, the main body group and the micro movements were first recognized for unknown letters, and the character recognition would be done if the recognized groups matched. If the recognized groups did not match, the error correction algorithm was used in the post-processing stage. In this method, there were preprocesses like the removal of bracket and duplicate points, point refinement, dimensional alignment, transition to origin coordinates, and the uniformity of the number of points and the spacing between them. For more information, see the reference. The preprocesses were necessary due to the fact that online data were input with an optical pen on a touch sensitive screen, so the number and spacing between the points and the dimensions of the sampled data varied greatly. Therefore, the uniformity was needed. To classify the main body, a series of global and point features were used, and a series of structural features and several features extracted from the first and second micro movements were used to classify the micro movements.

In addition, the feature vector dimensions of the main body were reduced from 102 to 17 features with the aim of increasing the resolution of the features and reducing the computational rate by using the principal component analysis (PCA) and linear separator analysis (LDA) algorithms. The classifier used for the main body of letters and SVM micro movements were based on One-vs-One approach (OvO). In this method, the optimal recognition rate was 98%. The online-TMU database was used with 4022 distinct letters.

A fuzzy method was used for distinct Farsi letter online recognition [21]. In the method, a hierarchical algorithm was proposed. A fuzzy classifier was also used to recognize the body of the letters. And to create this fuzzy classifier, expert knowledge and automated learning were combined. After the body recognition, the secondary symbols of the letters were recognized through a set of fuzzy rules. Using the method, the recognition rate was 90.3 for test samples of 54 individuals.

For Farsi manuscript online recognition, a database was introduced with digits, letters, and 1000 subwords, which were used most in the texts [22]. This database is fully functional for research on the recognition of Farsi letters, digits, and words.

For distinct Farsi letter recognition, the letters were first divided into 12 groups based on the points and signs below or above in the main body [23]. The points and signs of each letter were recognized, and each letter was assigned to one of the 12 groups. If the group included one letter, the unknown word would be recognized. Otherwise, the final recognition was carried by comparing the body of the letter with the body of letters from the same group using the minimum distance classifier.

Using a hidden Markov model (HMM), a method for Farsi letter online recognition was presented [24]. In this method, the number of letter components was obtained and the small components of the letter was recognized. After that, the body and the small components of the unknown word were preprocessed. Feature extraction was thus performed with greater accuracy. The features used were local and structural features. For the training phase, the Baum-Welch algorithm was used. This method also included a postprocessing operation. The good recognition rate was 97.22 for training samples and 94.9 for test samples.

4 Conclusion

According to the literature review, including 15 cases for digit recognition and 8 cases for letter recognition), it is clear that the resent research studies focused mostly on digit recognition in recent years. Table 1 includes the results of various methods used in the recognition of letters and digits, some of which were reviewed in the present paper. According to Table 1, it is evident that both the digit and letter recognition methods showed high recognition rates. Therefore, it is necessary to develop methods for commercialization. Moreover, excellent databases for digits and letters were developed during recent years, most notably the HODA Dataset for digits and the Online-TMU Database for distinct letters, both of which used frequently in the reviewed studies.

Table 1 The results obtained by different character recognition methods

Row	Reference	Feature extraction methods	Classification method	The number of character in the database	Recognition rate
1	Nafisi and Kabir (1994) [12]	Diagonal characteristic location	5-nearest neighbors	12,778	95.5–85.07
2	Masrouri (1998) [12]	DTW	Minimum distance	–	75
3	Mansouri (1988) [12]	Geometric moments and Fourier transform	Neural network	–	60
4	Razavi and Kabir (1997) [12]	3-tuple	Minimum distance	500	81
5	Johari (2000) [12]	Blocking	Fuzzy method	500	83
6	Razavi (2001) [12]	Feature selection with inheritance algorithms	Fuzzy method	500	86.2
7	Darvish (2002) [12]	Side information around contour points	Minimum distance	1288	89.9

(continued)

Table 1 (continued)

Row	Reference	Feature extraction methods	Classification method	The number of character in the database	Recognition rate
8	Nabavi (2004) [12]	Characteristic location	Combined 3 neural networks	2430	91
9	Ketabdar (1998) [12]	Structural features	Binary decision tree	3200	93
10	Soltanzadeh (2004) [12]	Combined 4 features	SVM	8913	99.57
11	Khosravi (2005) [12]	Combined 3 features	Neural network	80,000	99.33
12	Nahvi et al. (2008) [12]	PCA and combined 10 2-class classifiers	Neural network	8600	96.3
13	Moradi et al. (2010) [12]	A new FPGA-based method	Neural network	80,000	96
14	Ebrahimpoor et al. (2009) [25]	Characteristic location	Neural network and mixture of experts	80,000	97.52
15	Harifi et al. [12]	Shadow coding and a proposed non-symmetric fragmentation model	Neural network	–	97.6
16	Ebrahimpoor and Ahmadi (2009) [12]	PCA and characteristic location	Neural network and decision template	80,000	97.99
17	Ebrahimpoor and Sharifizadeh (2009) [12]	PCA	Neural networks and classifier Integration with class-conscious and class-conscious methods	8600	91.98
18	Borji et al. (2008) [12]	Features inspired by anterior visual pathway	KNN, ANN, SVM	80,000	99.63
19	Salehpoor (2012) [1]	K-means clustering and fuzzy SVM	PCA & LDA	8600	–
20	Akbari et al. (2012) [2]	SVM	Chain-code intersect	4096 training and 1532 test digits	–
21	Sedighi Nav et al. (2014) [3]	SVM	Gradient histogram and characteristic location	HODA digits	99.40

(continued)

Table 1 (continued)

Row	Reference	Feature extraction methods	Classification method	The number of character in the database	Recognition rate
22	Khosravi and Kabir (2006) [5]	Neural network	Improved gradient and gradient histogram	HODA digits	99.02–98.80
23	Karimzadeh and Mohammadi (2016) [6]	BSVM	Two-dimensional wavelet transform and PCA	HODA digits	91.75
24	Miri et al. (2013) [8]	Fuzzy classifier	Zoning, characteristic location, and Zernike moment	HODA digits	96.20
25	Ghanbari et al. (2011) [12]	Fuzzy classifier	Zoning	HODA digits	90.01
26	Talebian et al. (2014) [13]	Combined descriptors HOG and LBP	–	HODA digits	99.3
27	Marzani et al. (2013) [14]	SVM with OvA approach	Point, local, and global features	Online-TMU 4022	95.79
28	Ghanbari et al. (2010) (2009) [11]	Fuzzy classifier	Zoning	HODA digits	89.39
29	Mehralian and Fouladi (2012) [16]	SVM	–	Online-TMU 4022	98
30	Alibeigi et al. (2012) [18]	Neural network and minimum distance	Three simple features and characteristic location	Mitra, Lotus, Zar and Nazanin fonts	100
31	Kazemi et al. (2014) [19]	SVM	PCA	3200	82.51
32	Marzani et al. (2015) [14]	SVM with OvO approach	Point, local, global, and structural features, and PCA & LDA	Online-TMU 4022	98
33	Soleimani et al. (2006) [21]	Fuzzy method	–	–	90.3

References

1. Salehpour M (2012). Recognition of handwritten Farsi digits resistant to rotation and scale variation by fuzzing SVM classifier based on K-means clustering. In: The first international conference on farsi script and language processing. Semnan University, Faculty of Electrical and Computer Engineering, Semnan, Iran
2. Akbari Y, Jalili MJ, Foruzandeh A, Sadri J (2012) Examining the effect of image upgrading and gradient correction on improving the recognition rate of digits in Farsi manuscripts. In: The first international conference on farsi script and language processing. Semnan University, Faculty of Electrical and Computer Engineering, Semnan, Iran
3. Sadighi Nav M, Soleimani Eivari A, Khosravi H (2014) Feature reduction by binary particle swarm optimization for handwritten Farsi digit recognition. J Intell Syst Electr Eng 5:1
4. Yasayi S, Hatam A (2016) Selection of effective features in handwritten Farsi digit recognition using genetic algorithm. In: The first international conference on the new achievements in electrical engineering and computer sciences
5. Khosravi H, Kabir E (2006) Introduction of two fast and effective features for handwritten Farsi digit recognition. In: The 4th Iranian conference on machine vision and image processing. Ferdowsi University of Mashhad, Mashhad, Iran
6. Karimzadeh Bejestani S, Mohammadi Anbaran A (2016) Recognition of Farsi handwritten digits using SVM. Sci J Comput Sci Res 2:29–36
7. Soltanzadeh H, Rahmati M (2004) Recognition of Farsi handwritten digits using gradient and SVM classifier. *SID.ir*
8. Miri A, Razavi SM, Sadri J (2013) Improving the recognition rate of a fuzzy classifier for handwritten Farsi digit recognition with K-means clustering and particle swarm optimization. J Soft Comput Inf Technol 2:1
9. Khosravi H, Kabir E (2007) Introducing a very large dataset of handwritten Farsi digits and a study on their varieties. Pattern Recogn Lett 28(10):1133–1141
10. Nahvi M, Rafiee M, Ebrahimpoor R, Kabir E (2008) Combining 2-class classifiers for handwritten Farsi digit recognition. In: The 16th Iranian electrical engineering conference
11. Ghanbari N, Razavi SM, Ghanbari S (2010) Optimization of the handwritten Farsi digit recognition system based on the conscious selection of features affecting the increase of the fuzzy classifier recognition rate using two binary population-based algorithms. In: The sixth Iranian conference on machine vision and image processing, pp 364–371
12. Ghanbari N, Razavi SM, Nabavi Kerizi SH (2011) An intelligent feature selection method based on binary gravitational search algorithm in the handwritten Farsi digit recognition system. Iran J Electr Comput Eng 9:1
13. Talebian R, Mohammadpour M, Khosrow Abadi MM (2014) Handwritten Farsi digit recognition using the HOG-LBP descriptor. In: The first national conference on the development of civil engineering. Architecture. Electrical Engineering. and Mechanics in Iran
14. d. Marzani M, Razavi SM, Taghipour M (2013) Online handwritten Farsi digit recognition using SVM. In: The 8th conference on visual machine and image processing. University of Zanjan
15. Ghanbari N, Razavi SM, Nabavi Kerizi SH (2010) Using the GSA Algorithm in two feature selection methods and weighting the features to improve the recognition rate using fuzzy classifier. In: The 16th national conference by computer society of Iran. Sharif University, Computer Engineering Faculty, pp 356–362
16. Mehralian MA, Fouladi K (2012) Online handwritten Farsi distinct letter recognition based on the main body group recognition using SVM. J Data Symbol Process 1(17):59–66
17. Khosravi H, Kabir E (2006) Introducing two fast and efficient features for recognition of Farsi handwritten digits. The 4th conference on machine vision and image processing, pp 25–26
18. Alibeigi M, Razavi SM, Sadri J (2012) A simple method for handwritten Farsi distinct letter recognition. In: The first conference on pattern recognition and image analysis in Iran. University of Birjand

19. Kazemi M, Yousef Nejad M, Nourian S (2014) Handwritten Farsi letter recognition using a combination of feature extraction based SVM classifiers. In: The 12th national conference on intelligent systems and information and communication technology
20. Marzani M, Razavi SM, Taghipour M (2015) A practical approach to online handwritten Farsi distinct letter recognition using a combination of knowledge of the main body and micro movements. J Comput Intell Electr Eng 6(2):87–100
21. Soleimani Baghshah M, Bagheri Shurki S, Kasmaei S (2006) Online handwritten Farsi distinct letter recognition using fuzzy methods. The 14th Iranian conference on electrical engineering. Amir Kabir University of Technology, Tehran
22. Razavi SM, Kabir E (2004) A Database for online handwritten Farsi manuscript recognition. The 6th conference on smart systems. Shahid Bahonar University, Kerman
23. Razavi SM, Kabir E (2004) Online handwritten Farsi distinct letter recognition. The 6th conference on smart systems. Shahid Bahonar University, Kerman, Azar
24. Faraki M, Palhang M Online handwritten Farsi letter recognition based on the hidden Markov model. J Electr Eng 40:1. Tabriz University
25. Nahvi M, Kiani K, Ebrahimpoor R (2010) Improving gradient feature extraction based on discrete cosine transform towards recognition of Farsi handwritten digits. The 18th Iranian conference on electrical engineering, pp. 3067–71

A New Model for Iris Recognition by Using Artificial Neural Networks

Mina Mamdouhi, Manouchehr Kazemi and Alireza Amoabedini

Abstract Today, with advances in the science of machine vision, wide dimensions have been opened in the field of identity identification (biometric) in people's lives. With the increase in inspection, surveillance and security centers, biometric systems are more crucial than ever. Iris recognition is one of the main biometric identification approaches in human's identity recognition which has become a very functional and attractive subject in the research and practical applications. Due to the unique features of the iris, this kind of recognition is highly effective to identify a person. In literature, many researches have been done with regard to locating, image description and iris recognition. In this paper, a solution is provided for extracting features of iris that generated datasets can be analyzed. Given that the neural network uses this data set, iris patterns classification is done. Adaptive learning strategy is used to train the neural network. The simulation results show that the proposed system of identification of individuals can offer 95% accuracy in normal conditions and 88% in noisy conditions.

Keywords Iris recognition · Feature extraction · Multilayer perceptrons
Gaussian noise · Walsh Hadamard transform · Neural network

M. Mamdouhi · M. Kazemi (✉)
Department of Computer Engineering, Ashtian Branch,
Islamic Azad University, Ashtian, Iran
e-mail: univer_ka@yahoo.com

M. Mamdouhi
e-mail: mina.mamdouhi@gmail.com

A. Amoabedini
Department of Computer Engineering, Safadasht Branch,
Islamic Azad University, Tehran, Iran
e-mail: a.amoabedini@ut.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_14

1 Introduction

Biometric recognition is one of the tools that will lead to the identification in the world today. Given that millions of people daily pass border crossings and checkpoints and since inspection is an exhausting process and full of errors, biometric systems can be used. Biometric measures biological characteristics, such as fingerprints, iris pattern, facial structure, hand or behavioral characteristics such as walking or sound. Biometric uses these features automatically to identify people, in general, ideally, characteristics should be limited to a person and they are constant over time and can be easily measured. The human's eye iris recognition is a biometric process. You can also take advantage of iris recognition such fast matching and strong resistance against being fake and iris stability as one of human organs. Iris tissue, as a biometric feature to identify, is useful even without examining its internal variability. The eye is an internal organ that is visible externally, and we can get its pictures easily. Compared with fingerprints it should be mentioned that they are in outer organ of skin which are more susceptible to damage and change. Also a change in the physical texture of iris includes risk of harm to a person's vision. So, as long as the person does not use lenses with specific characteristics, iris tissue that has been given to the system can be considered valid. Aging effect on iris a biometric, is an active research section [1, 2] but outside the scope of this research. Primarily methods of iris recognition work by focusing iris and marking any artificial obstructive effects. In the image of an eye, the iris is a circular section between the sclera and pupil. Eyelids, eyelashes and bright lights may block parts of this region. So the iris should be normalized in a way that it can be more appropriate to specify the characteristics of iris. Properties often are created through image filtering and normalized shape is a rectangular image that includes the iris without any obstruction. By using neural network (NN: Neural Network); this study offers a prototype of a powerful iris pattern recognition system for classification of a number of different people. At the end the proposed system is carefully evaluated with and without noise.

2 Copyright Form

Surveillance cameras transmit images to a system and these images after pre-processing, are ready for feature extraction. There are many ways to extract features from images. The most common methods of image feature extraction are accelerating the powerful features, Haar wavelet, color histogram, the local binary pattern, swirl gradient histogram techniques, principal component analysis. In this section, we explore existing systems in the field of biometric iris recognition. For example, detection of HT circular methods [3], integro-differential operator [4] and in the field of iris feature extracting methods such as small two-dimensional Haar wavelet transform, Gabor filters, small Daubechies wavelet transform and in the field of recognition, classification algorithms such as NN, Bayesian, deep learning and SVM: Support Vector Machine were presented. In Umer et al. [5], by using a new Multiscale way to extract features an iris recognition system (IRS: Iris Recognition System) with improved performance was introduced. This system is able to evaluate and identify people efficiently and effectively in the four stages. In the first phase, a rapid method for determining the position of the iris was adopted. Secondly, only part of the picture of the iris to avoid the problem of occlusion was used for authentication. Third, Multiscale Morphologic features of segmented image were extracted from iris, and finally in the last stage, SVM was used as a classifier. And according to the results of Falohun's et al. [6] research, we found that Quadtree segmentation in the IRS by which is prepared by NN, in division, training and detection stages act faster than Hough transform (HT) in the same NN environment. Nabti et al. [7] offered a multi-resolution method for extracting iris features. In order to do feature extraction, they applied a specific Gabor filter collection on normal iris image. Khedkar et al. [8] after initial operation of locating and Hadamard feature extraction method, different NN configuration including Multi-Layer Perceptron (MLP), Radial basis function and SVM were applied after changing parameters of respective networks also Gaussian noise and uniform noise have been injected to all input features and it was found that the MLP with one hidden layer in terms of performance acts better compared to all other networks. Mira et al. [9] offered eye IRS on the basis of mathematical morphology operators (MM: Mathematical Morphology). MM is directly related to the information in the form of digital images in the environmental domain. The main advantage of MM is the ability to extract essential texture, structure information and characteristics of the image in the IRS. Liu et al. [10] is able to acquired different domains in different directions and then combine them to make a coherent three-dimensional image of the sample in a process referred to as 4D-OCT, and of course Optical Coherence Tomography (OCT) requires large and expensive pieces of equipment and trained operators to run them. In addition, the sensors act very close to the sample and the slightest movement of the sample will add significant noise. Albadarneh et al. [11] evaluated a hybrid Gabor approach with initiatives such as Discrete Cosine Transform (DCT), Histogram of Oriented Gradients (HOG). HOG accuracy in identifying iris images reaches to 20% and accuracy of a Gabor combined method reaches to 76%. Tallapragada et al. [12] in order to o identify offered a framework with SVM and Hidden Markov Model (HMM) that performs well, but only used 100 samples that reach to the accuracy of 93.2%.

3 Research Methodology

In this section, the proposed layout describes the eye IRS structure in several steps. The proposed plan offers simplified procedure for the iris image processing, feature extraction and training by the MLP for iris recognition. Partitioning the iris image data and a data conversion is offered and studied.

Figure 1 shows the overall architecture of the proposed IRS. First, the proposed system performs pre-processing on the iris image and then detects the iris of the eye through the category of extracted image data. Obtaining iris image includes lighting systems, positioning systems and physical imaging systems, and results are described in the discussion section.

Iris recognition includes pre-processing, feature extraction and information classification by NN. When obtaining an iris, iris picture in the input sequence must be clear. The resolution of points and tiny elements and clarity of boundary between the iris and the pupil and the boundary between iris and sclera affect iris image quality. High-quality pictures should be chosen for iris recognition. In preprocessing, the iris picture is entered, resize and taken into the gray mode. Image enhancement process can be obtained by applying the combined filters. Image is

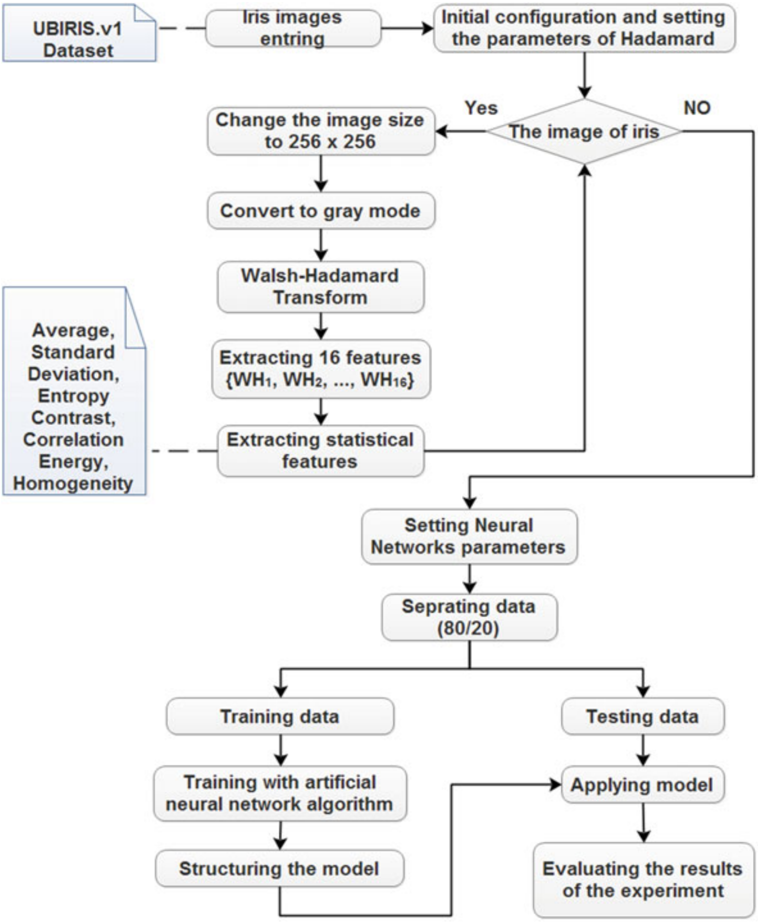

UBIRIS.v1 Dataset
Iris images entring
Initial configuration and setting the parameters of Hadamard
Yes
NO
The image of iris
Change the image size to 256 x 256
Convert to gray mode
Walsh-Hadamard Transform
Extracting 16 features {WH1, WH2, ..., WH16}
Average, Standard Deviation, Entropy Contrast, Correlation Energy, Homogeneity
Extracting statistical features
Setting Neural Networks parameters
Seprating data (80/20)
Training data
Testing data
Training with artificial neural network algorithm
Applying model
Structuring the model
Evaluating the results of the experiment

Fig. 1 The overall architecture of proposed iris recognition system

displayed by a matrix that black and white values in gray scale, describe the iris image. This matrix is converted to a set of data in line with NN training. IRS includes two functions: Training mode and test mode. Firstly, detection system training is done by using black and white values of iris images. NN is trained with all iris images. After the training, in test mode, NN does the classification and recognizes iris patterns that belong to a particular person.

In Fig. 2, the proposed system uses the NN to detect iris patterns. In this method, that standardized and modified image iris is displayed by a two-dimensional array. This array includes gray scale values of iris pattern texture. Characteristics and attributes that are extracted from these values are the input signals for NN.

Two hidden layers can be used in NN. In this structure, $x_1, x_2, \ldots, x_m$ are input array values characteristics that describe the iris texture information, $p_1, p_2, \ldots, p_n$ are output patterns that describe the iris. kth output of NN is determined by Eq. 1 [13].

$$P_k = f_k\left(\sum_{j=1}^{h2} v_{jk} \cdot f_j\left(\sum_{i=1}^{h1} u_{ij} \cdot f_i\left(\sum_{l=1}^{m} w_{li} x_l\right)\right)\right) \tag{1}$$

where v_{jk} are the weights of output and hidden layers of the network, u_{ij} are the weights of the hidden layers, w_{li} is the eight of input and the first hidden layers, F is the activation Function in neurons and x_l is the input signal. Here k = 1, …, n, j = 1, …, h2, i = 1, …, h1, l = 1, …, m that m is the number of neurons in the input layer, n the number of neurons in the outer layer, h1 and h2 are the number of neurons in the first and second hidden layers, respectively. In Eq. 1, P_k determines NN output signals as follows [13]:

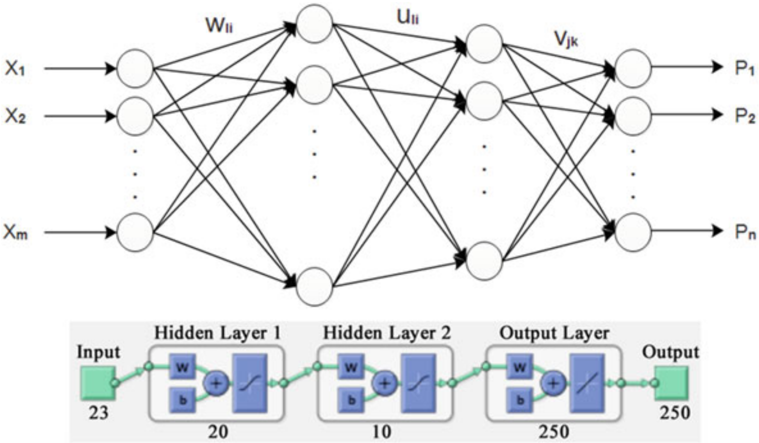


Fig. 2 The structure of the neural networks in the proposed plan

$$
\begin{aligned}
&\mathrm{y}_j = \frac{1}{1+e^{-\sum_{i=1}^{h1} u_{ij} y_i}}, \ \mathrm{y}_i = \frac{1}{1+e^{-\sum_{l=1}^{m} w_{li} x_i}} \\
&p_k = \frac{1}{1+e^{-\sum_{i=1}^{h2} v_{jk} y_j}}
\end{aligned}
\tag{2}
$$

Here y_i and y_j are respectively equal to the first and second hidden layers output signals. After NN activation, the NN parameters training starts. Then the trained network is used for iris recognition in test system. In this section, a gradient-based learning algorithm with adaptive learning rate is used. This ensures convergence and accelerates learning processes. In addition, a momentum is also used to speed up the learning process.

Initially, NN parameters are randomly generated. u_{ij}, v_{jk} and w_{li} parameters that are related to the NN are equal to the weight coefficients of the first, second and third layers. Here, $\mathrm{k} = 1,\ldots,\mathrm{n}$, $\mathrm{j} = 1,\ldots,\mathrm{h2}$, $\mathrm{i} = 1,\ldots,\mathrm{h1}$, $l = 1,\ldots,\mathrm{m}$. To produce model for NN detection, Training of the weight coefficients u_{ij}, v_{jk} and w_{li} is carried out. During the training cost function value is calculated as follows [13]:

$$
E = \frac{1}{2}\sum_{k=1}^{n} (\mathrm{P}_k^d - \mathrm{P}_k)^2 \tag{3}
$$

Here n is the number of output signals and P_k^d and P_k are respectively equal to the current and desired network output values. Parameters u_{ij}, v_{jk} and w_{li} that are related to NN are regulated by the following formula [13].

$$
\begin{aligned}
w_{li}(t+1) &= w_{li}(t) + \gamma \frac{\vartheta E}{\vartheta w_{li}} + \lambda(w_{li}(t) - w_{li}(t+1)) \\
u_{ij}(t+1) &= u_{ij}(t) + \gamma \frac{\vartheta E}{\vartheta u_{ij}} + \lambda(\mathrm{u}_{ij}(t) - u_{ij}(t+1)) \\
v_{jk}(t+1) &= v_{jk}(t) + \gamma \frac{\vartheta E}{\vartheta v_{jk}} + \lambda(\mathrm{v}_{jk}(t) v_{jk}(t+1))
\end{aligned}
\tag{4}
$$

Here γ is the learning rate, and λ is momentum. Adaptive learning rate is used to increase learning speed and ensure convergence. The following strategy is used for any given number of new courses. NN learning parameters begins with a small amount of γ learning rate. During the training, if the amount of the error $\Delta E = E(T) - E(T+1)$ is positive, the learning rate increases and If changes in the error $\Delta E = E(T) - E(T+1)$ is negative, the learning rate decreases. After determining the derivatives, NN parameters updating is performed.

4 Discussion and Results

To evaluate the proposal, a dataset is needed that contains images of human iris. Therefore, we extract 250 images from UBIRIS.v1 dataset that we have put them beside the executable file.

UBIRIS dataset is composed of 241 people in various states and has 1877 photos. This collection is standardized and recorded with the minimal noise and noise factors, brightness, contrast and reflectivity are minimized. All images were recorded by Nikon E5700 camera in RGB mode with 300 dpi image quality and focal length of 71 mm at a depth of 24 bits [14].

Dataset contains several images of the human iris is that each belongs to a human. All pictures must be in standard mode and the four directions (up, down, left and right) and the image should start from cornea and space should not be empty. First of all human iris images were converted to the same size and 23 features were extracted from any image. The vector of these features is as follows, that all of which have been obtained from the images of the iris and at the end of each image has 23 features.

Person1= [WH1, WH2… WH16, Contrast, Correlation, Energy, Homogeneity, Average, Standard Deviation, Entropy]

WH1…WH16: Sixteen Walsh features have been achieved from Walsh-Hadamard transform. To find 16 features of Walsh-Hadamard, each iris picture will be converted into 16 blocks and Walsh- Hadamard transform on each block is a feature of 16 features [15].

$$W(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\left[\prod_{i=0}^{n-1}(-1^{(b_i(x)b_{n-1-i}(u)\,+\,b_i(y)b_{n-1-i}(v))})\right] \tag{5}$$

Walsh function has a unique sequence value. This function can be produced in different ways. Hadamard function can be used in MATLAB to create this function. Here the length of Walsh function is considered as 64. Rows and columns in this matrix are symmetrical and comprise Walsh functions. Walsh functions in this matrix are not considered in increasing order of their sequence or the number of zero-crossing. But they are arranged according to Hadamard order. Then, one dataset contains 250 images; each with 23 features will be prepared. Of course in the initial assessment this number was obtained and the original volume of this dataset is greater than this number of image. Now data is ready to enter into the neural network. NN is used for the production of IRS. Eighty percent of data will be used for NN training in order to train the system to identify the future cornea and 20% of data will be used to evaluate the training system. Output matrix specifies that any input image of the iris is related to which person. In fact, in NN, each neuron can identify a person so in this project 250 irises were selected. Thus the structure of neural network has 250 neurons in the output layer. Therefore, the output matrix should also have 250 rows that each row represents an individual. In NN with Iris input and output matrix, the number of person is designed with two hidden layers that there are 20 neurons in the first and 10 neurons in the second layer. Neural network output matrix shows the most similarity to the cornea.

Program testing was conducted at two conditions of with and without noise. In no noise condition, accuracy has been validated between 95 and 100% and in terms of noise the accuracy was 88%. A noisy condition was applied $\sigma = 4$ by Gaussian Blur filter on the 20% of the images. Figure 3 shows the acceptable level of mean square error. This chart reports the mean square error in 10 epochs for a test, training, and validation collections. As it is evident in the fourth repetition the lowest error has been achieved for the validation operation. Figure 4, shows the peak signal to noise ratio with respect to the $\sigma = 4$ value for Gaussian filtering for 20 sample pictures in the dataset. In classification operation, the accuracy of a class equals to the number of items that are to be labeled correctly as diagnosed normal class with regard to the total number of items labeled normal class. Here, after evaluating 180 cases properly diagnosed by normal label and given that the amount of wrong negative is equal to 20, we have:

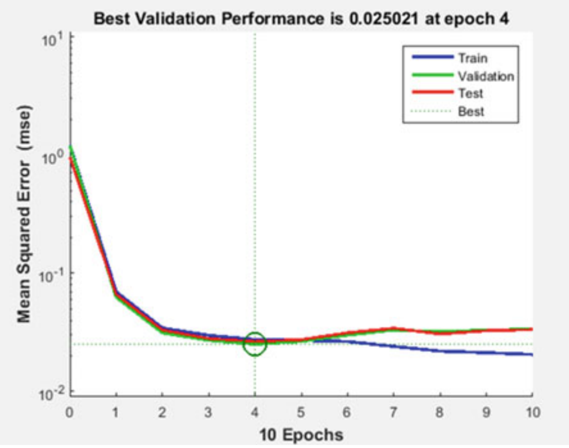
$$TPR = \frac{\sum True\,Positive}{\sum Condition\,Positive} \tag{6}$$

And also for the false noise level, given that the rate of 10 cases were misidentified as false positive class, we have:

$$FPR = \frac{\sum False\,Positive}{\sum Condition\,Negative} \tag{7}$$

According to the above results, positive likelihood ratio ($LR+$) can be simply calculated.

Fig. 3 The average error in 10 epochs



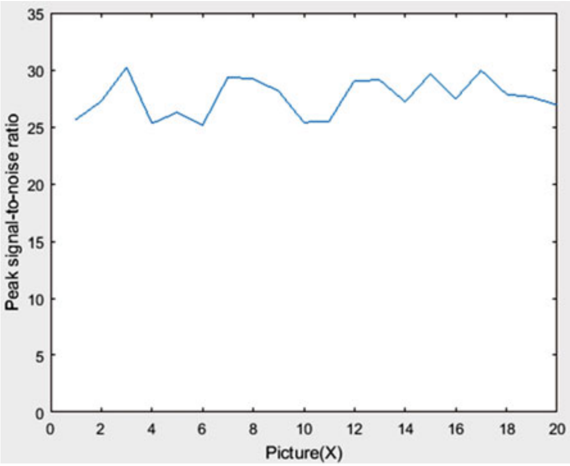


Fig. 4 The peak signal-to-noise ratio for each picture by applying a Gaussian filter

$$LR+ = \frac{TPR}{FPR} \tag{8}$$

For items rate, class lost items can be calculated from total negative items that have been misdiagnosed based on positive number of items.

$$Miss\,Rate = \frac{\sum False\,Negative}{\sum Condition\,Positive} \tag{9}$$

Calculating sensitivity (SPC) shows us the rate of negative stories that have been diagnosed correctly.

$$SPC = \frac{\sum True\,Negative}{\sum Condition\,Negative} \tag{10}$$

To calculate SPC, the negative likelihood ratio (−LR) can be obtained.

$$LR- = \frac{FNR}{TNR} \tag{11}$$

Diagnostic odds ratio (DOR) according to positive and negative likelihood ratios has been achieved from the following equation.

$$DOR = \frac{LR+}{LR-} \tag{12}$$

The resulting number shows the odds ratio for non-noise level is high. To calculate the overall accuracy according to the following equation we have.

$$ACC = \frac{\sum True\,Positive + \sum True\,Negative}{\sum Total\,Population} \quad (13)$$

All methods that can be used in feature extraction are not applicable for biometric iris recognition. Methods such as gradient HOG which are used in Biometric facial recognition have low accuracy in iris recognition. Table 1 shows the proven results of above relationships.

Finally in Fig. 5, a comparison of research that has been proposed as the frequency shows the accuracy of proposed scheme.

In this paper it has been shown that Walsh-Hadamard transform method is an effective way to extract features from pictures. Also applying the NN classifier is an effective detection tool in noisy environments. Initiatives filters to increase the resolution of very small elements in the resolution of the boundary between the iris and the pupil and boundary between iris and sclera that is effective on the iris picture quality, will be studied.

Table 1 The results of the classification process

Metric	Results (%)
FPR	16.67
TPR	94.74
SPC	66.67
LR−	7.89
LR+	5.68
DOR	71.99
ACC with noise	88
Miss rate (with out noise)	5.26
ACC (with out noise)	95

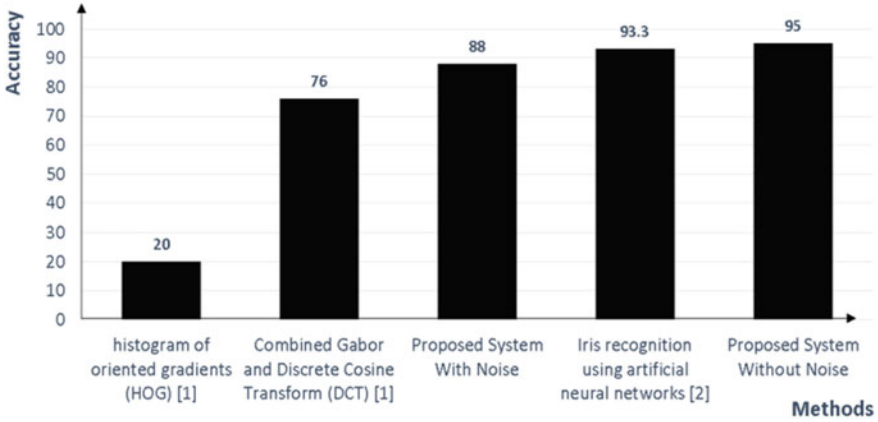


Fig. 5 Comparing the proposed scheme in noise and without noise condition with other previous projects on the UBIRIS dataset

5 Conclusion

In this article, we described a multi-step plan. We have shown the suggested solutions with the feature extraction approach of Walsh-Hadamard transfer with other statistical characteristics and classification of neural network can identify iris pictures with high precision. Simulations show that up to 95% recognition accuracy can be achieved with testing the same irises. The Gaussian blur filter was applied to 20% of records. The results show that proposed scheme with an accuracy of 88% is still good diagnostic tool. In terms of the future researches.

References

1. Fenker SP, Bowyer KW (2011) Experimental evidence of a template aging effect in iris biometrics. In: IEEE computer society workshop on applications of computer vision, pp 232–239
2. Sazonova N, Hua F, Liu X, Remus J, Ross A, et al. (2012) A study on quality-adjusted impact of time lapse on iris recognition. In: Sensing technologies for global health, military medicine, disaster response, and environmental monitoring II; and biometric technology for human identification IX, vol 8371, pp 83711–83719
3. Djekoune O, Messaoudi K, Amara K (2017) Incremental circle Hough transform: an improved method for circle detection. Optik Int J Light Electr Opt 133:17–31
4. Kumar AA, Gupta A (2015) Iris localization based on integro-differential operator for unconstrained infrared iris images. In: 2015 international conference on signal processing, computing and control (ISPCC), Waknaghat, pp 277–281
5. Umer S, Dhara BC, Chanda B (2015) Iris recognition using multiscale morphologic features. Pattern Recogn Lett 65:67–74
6. Falohun AS, Ismaila WO, Adeosun O (2015) Performance evaluation of quadtree & hough transform segmentation techniques for iris recognition using artificial neural network (ANN). Int J Comput Trends Technol (IJCTT) 25(1):18–22
7. Nabti M, Bouridane A (2008) An effective and fast iris recognition system based on a combined multiscale feature extraction technique. Pattern Recogn 41(3):868–879
8. Khedkar MM, Ladhake SA (2013) Robust human iris pattern recognition system using neural network approach. In: 2013 international conference on information communication and embedded systems (ICICES), Chennai, pp 78–83
9. Mira Jd Jr, Neto HV, Neves EB et al (2015) Biometric oriented iris identification based on mathematical morphology. J Sig Process Syst 80(2):181–195
10. Liu JJ, Grulkowski I, Potsaid B et al (2013) 4d dynamic imaging of the eye using ultrahigh speed ss-oct. In: Proceedings of SPIE, vol 8567
11. Albadarneh IA, Alqatawna J (2015) Iris recognition system for secure authentication based on texture and shape features. In: 2015 IEEE Jordan conference on applied electrical engineering and computing technologies (AEECT), Amman, pp 1–6
12. Tallapragada VVS, Rajan EG (2012) Improved kernel-based IRIS recognition system in the framework of support vector machine and hidden markov model. IET Image Process 6(6):661–667
13. Abiyev RH, Altunkaya K (2008) Personal iris recognition using neural network. Int J Secur Appl 2(2):41–50

14. Proenca H, Alexandre LA (2005) UBIRIS: a noisy iris image database. In: Springer lecture notes in computer science—ICIAP 2005: 13th international conference on image analysis and processing, Cagliari, Italy, vol 1, pp 970–977
15. Bell DA (1996) Walsh functions and Hadamard matrixes. Electr Lett 2(9):340–341

Designing a Fuzzy Expert Decision Support System Based on Decreased Rules to Specify Depression

Hamed Movaghari, Rouhollah Maghsoudi and Abolfazl Mohammadi

Abstract Depression is a psychological disorder, if it doesn't be diagnosed and cured in time, can effect on quality of humans' life in wide dimensions. Thus, diagnosis easily and quickly is necessary need of sociality's generally healthy. The aim of this paper is designing an fuzzy decision support system to implement BDI-II. Questions for BDI-II are grouped into multiple factors. In medical sciences, disorders, and diseases that lack high confidence and complexity in diagnosis, intelligent systems have better confidence capacity in diagnosis. In this investment, the structure is designed in form of two factors and five factors. The results show that designed system with two factors compared to five factors structure with 94.2% diagnostic power has implementations train data of BDI-II. Hence, in future, the psychologist can use this system as a decision support system of decision support in clinical and hospital diagnoses.

Keywords Fuzzy logic · ANFIS · Depression · BDI-II

1 Introduction

Depression as one of the common psychiatry diseases is a major threat to the general health [1]. This thread with vanity sense and physical and cognitive changes has the effect on one's performance ability [2]. Bad sense and disappointment continue until the person cannot act at all [3]. It reduces the quality of social,

H. Movaghari · R. Maghsoudi (✉)
Department of Computer Engineering, Islamic Azad University,
Mahmudabad Branch, Mahmudabad, Iran
e-mail: r.maghsoudy@srbiau.ac.ir

H. Movaghari
e-mail: hmovaghari@iaumah.ac.ir

A. Mohammadi
Hospital Rouzbeh, Tehran Medical Sciences University, Tehran, Iran
e-mail: a-mohammadi@tums.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_15

interpersonal, occupational or educational performance [4]. Providing guides that can effect while helping to clinic judgment in diagnosis and diagnostic information, and as established diagnostic system upgrade value and decision reliability, is a psychiatry need [5]. Reliable diagnosis is a basic need in curing mental disorders, based on which, many manuals are offered in responsibility such as DSM,[1] BDI[2] [6]. According to concepts of fuzzy sets and fuzzy logic in numerous articles and related questions: if disease in BDI-II[3] has scored such as 46 between interval 29–63 that is in severe depression range. But does this person belong to this set? This set represents that all diagnoses and symptoms in emotional-cognitive and physical axes influenced by disease. But the person has score 46, in fact, is semi-depressed and severe semi-depressed. In other words, if we suppose severe depression as black and depression as white, that person's depression with score 46 has the spectrum of gray. In fact, there is no boundary between a patient and healthy [7, 8]. Therefore, we need a hypothesis that can formulate human science systematically and put it with other mathematical models in engineering systems. In medical sciences also ambiguity and uncertainty are effects. This problem is obvious and related to medical nature. Because diseases appear as various forms and with various severities. In fact, the best definitions of symptoms and illnesses are presented using vague and inaccurate language terms [9]. In [10] the author claimed, fuzzy systems model vague phenomena and basically, the fuzzy theory that defined by Professor Zadeh is an accurate theory. There are two types of justification for fuzzy theory: first, due to the complexity of the real world, one cannot offer exact description. Hence, fuzzy has introduced that has the power of analysis. Second, due to the importance of human science and knowledge, we need a hypothesis that can be formulated systematically.

In this relation, in [11], to specify the validity and reliability of BDI-II scores in determining the severity of depression, through a fuzzy logic model of the Chinese version were used. The method has been fuzzy logic. Samples of this study were 204 (123 women and 81 men) patients. The results showed that the reliability of the Chinese version of the BDI is higher than the original version. The analysis also showed that Both versions have the ability to recognize of the clinical and nonclinical (Chinese version 80.3% and original version 73.2%). Finally propose, Due to the better diagnostic power of the Chinese version, Besides its use in medical diagnosis used in engineering research.

In [12] the authors, applied a score of BDI-II by clustering technique FCM[4] and compared it with k-means clustering technique. All of the samples in this study were 559 (240 clinical, 319 nonclinical respectively). The results showed that clustering FCM has less cognitive cost in healthy observances.

[1] Diagnostic and statistical manual of mental disorders.

[2] Beck Depression Inventory.

[3] Beck Depression Inventory Second Version.

[4] Fuzzy C-means.

In [13] the researches, divided 21 questions of BDI-II with five factors structure and implemented it with Sugeno fuzzy inference system. The results showed that designed system has high value in psychiatry diagnosis.

In [14] the authors, due to the ambiguity of depression (mild, moderate and severe) for physicians, BPNN[5] offered an appropriate way to determine the exact scores. 124 samples were selected to answer 26 questions. 70% data for train and 30% were used in the test. The results showed that The propagation neural network (100% mild depression, 77% moderate and 90% severe depression) to detect. Finally, a suitable model was implemented with R = 94%.

In [15] the researches, according to the association between depression and heart rate variability. Data of 10 depressed patients and 10 healthy controls heart rate variability were collected. Changes in heart rate with her ECG[6] was recorded for a period of 800s. 6 features of heart rate changes were extracted. A cluster-based neuro-fuzzy system with fuzzy membership function was implemented. The results showed that there is a obvious difference between heart rate and depression. Therefore, this method can be used in mobile phone applications, as a regulatory system in an emergency conditions to help people.

In [16] the authors, for screening and diagnosing of telling states of depression (mild, moderate and severe), two supervised learning algorithms (classification with BPNN and ANFIS[7] classifiers) and unsupervised (clustering technique with SOM[8]) used both methods was given. The performance of two methods was compared. The results showed that a hybrid system in ANFIS compared with BPNN classifier has better performance and help diagnose depression.

In [17] the researches used BDI-II with five factors structure with effective parameters on depression such as age, systolic hypertension, diastolic, and body mass index in the investigation and simulated a fuzzy system through the neuro-fuzzy system and ANFIS inference system. The results showed that using BDI-II with physiological parameters effective on making depression by using expert intelligent system is an effective factor in depression diagnosis. Hence, they suggested that due to being effectiveness above way, we develop this method in heart, pulmonary and cancer diseases.

In [18] the author to model the process of diagnosing depression, two fuzzy clustering techniques (FCM and FkNN[9]) used. The study samples were 302 persons. Information was collected through questionnaires in India. Data analysis was performed using Cronbach's alpha test. Results showed that both methods have correct diagnostic power with the reliability of 0.98. nevertheless proposed GA[10] to optimize the purpose of this study.

[5] Back propagation neural network.

[6] Electro Cardio Graphy.

[7] Adaptive network-based fuzzy inference system OR Adaptive neuro-fuzzy inference system.

[8] Self-organizing map.

[9] Fuzzy k-Nearest Neighbour.

[10] Genetic Algorithm.

In [19] the researches due to differences of opinion physicians in determining the severity of depression, expressed a combination of expert tools. The instrument used a combination of (1) information confirmed by experts and specialized resources of the signs and symptoms of depression and (2) multilayer feedforward backpropagation neural network. The samples were 302 adults. Each of the samples had 16 symptoms of depression, such as sadness, pessimism and some things like them. Signs and symptoms were used as input. 70% data and 30% were used for training and testing respectively. Results showed that proposed tools have a 98.96% average diagnosis. And with all signs and symptoms of accuracy is 98.91%.

In [20], a fuzzy logic expert system to determine the risk levels of depression in consultation with psychiatrists and psychologists from the group consisting of two hospitals Nigeria, were implemented. The consultation, in order to select three physical symptoms (age, blood pressure and body mass index) and a predictor of psychological symptoms (GHQ[11]), respectively. 125 depressed adults were enrolled. Results showed that the implemented system has the power to detect the risk of depression as same as experts psychiatry and psychology expert.

In [21], the author offers a mathematical model to understand symptoms and depression diagnosis by using clinical psychiatrists. This model is revised completely with DSM-IV-TR.[12] In this model, with the help of clinical psychiatrists, 14 symptoms of adult depression are considered by the American Society of Psychiatrists in the last resort. Each of these symptoms is measured according to power and pathogenicity severity by expert psychiatrists. Then, by using principal component analysis, 7 Factors are derived from 14 Factors. Then by using these 7 Factors as the input of a controlling system, Then, using these 7 factors as the input of a fuzzy Mamdani composite controller system in the multi-layered back propagation of the neural network was used. The output of this controlling system is regulated by algorithm after publishing neural networks. Finally, by using 302 adult depressed patients and 50 healthy persons, it is r concluded that hybrid controller system can diagnose depression in adults with 95.50% accuracy.

In [22] the researches represented two types of intelligent neural network (Propagation neural network and neural network with radial basis function) to simulate the data examined depression. This research has been done during 2004–2005 with a total of 300 samples (which were first referred to the hospital and did not consume any antidepressant medication and had suicidal thoughts), was selected by 3 expert psychiatrists (average experience of 10.4 years). The results showed that both intelligent neural network considerably very low error, and both instruments are the right techniques in automatic detection of depression. But the neural network with radial basis function compared with propagation neural network is better. Finally, proposed to reach a final diagnosis of depression, increased the number of test samples.

[11] General Health Questionnaire.

[12] DSM IV Text Revise.

In [23], the author to mechanize the process of diagnosing depression Psychiatric Association's DSM-IV-TR America, with Purposive sampling 270 samples from each of the samples East Indian hospital has seven clinical depression factors were collected over two years. Then five expert psychiatrists and psychologists in consultation with the validity and reliability (statistically) examined data collected. This data was simulated by a fuzzy neural model. The neuro-fuzzy system, the functions of the Gaussian triangular membership and membership functions in two modes were considered mild depression and moderate depression. The results show a neuro-fuzzy system with Gaussian membership functions, is best model for automating and the accurate prediction of depression is 94.4%.

In [24] the researches by using EEG data evaluated by experts and choosing 65 patients by convenient sampling and applying it with two algorithms artificial neural and neuro-fuzzy network concluded that neuro-fuzzy algorithm is suitable for depression diagnosis with 76.88% accuracy.

This study has checked designing fuzzy decision support system based on data with decreased rules to diagnose depression severity. To design this system, first, we produce data set of inputs and output related to BDI-II by experts and use it to design an ANFIS with decreased rules in five factors and two factors structure.

2 BDI-II

BDI-II has 21 self-reported multi choices to measure depression severity in adults and adolescences with age 13 and more. To use this inventory, we must answer 21 questions, a number from 0 to 3 to each choice. Sum of numbers related to selected choices equals depression severity that psychiatrists refer it as depression score. This score in number between 0 and 63, by which we can determine depression in four sets. These four sets are determined as:

(1) 0–13, healthy or minimal depression
(2) 14–19 mild depression
(3) 20–28 moderate depression
(4) 29–63 serve depression [13, 25]

These sets can be shown as the diagram in Fig. 1.

According to psychological expert persons, we can divide 21 questions of BDI-II in various structures [26]. In this investigation, both five factors and two factors structures are checked. Five factors structures of 21 questions are grouped into five Factors: emotional, cognitive, motivational, physical, and delusional, as are shown in Table 1 [13, 26].

The structure of the two factors of 21 questions is grouped into two cognitive-affective and somatic Factors and is shown in Table 2 [26, 27].

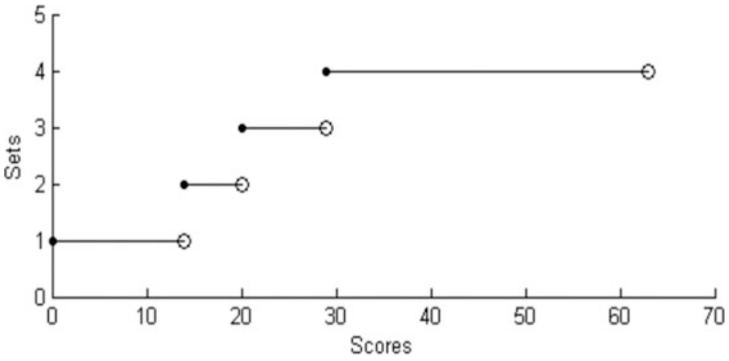


Fig. 1 Diagram of depression severity in BDI-II

Table 1 BDI-II with five factors structure

Factors	Test
Emotional	Sadness
	Crying spells
	Agitation
	Loss of Interest
	Irritability
Cognitive	Pessimism
	Sense of past failure
	Guilt
	Punishment feelings
	Self-dislikes
	Self-criticalness
	Indecisiveness
	Worthlessness
	Concentration difficulty
Motivational	Change in sleep pattern
	Changes in appetite
	Fatigue
	Loss of interest in sex
Physical	Loss of pleasure
	Loss of energy
Delusional	Suicidal thoughts

3 Fuzzy Logic

Fuzzy in glossary means vague, inexact, unknown, ambiguous, dizzy and chaotic [28]. First, following setting up fuzzy sets has introduced by Professor Zadeh in 1965 [29]. Fuzzy sets were founding a successful method to modeling uncertainty

Table 2 BDI-II with 2 factor structure

Factors	Test
Cognitive-affective	Sadness
	Crying spells
	Sense of past failure
	Guilt
	Punishment feelings
	Self-dislikes
	Self-criticalness
	Pessimism
	Loss of pleasure
	Loss of interest
	Worthlessness
	Suicidal thoughts
	Indecisiveness
	Irritability
	Agitation
	Loss of interest in sex
Somatic	Fatigue
	Loss of energy
	Concentration difficulty
	Changes in appetite
	Change in sleep pattern

and ambiguous [30]. Kosko in fuzzy thinking writes, there has been a wrong in science that all scientists have committed. According to classic logic (Aristotelian, two values and boolean), everything proved by a constant principle, that thing is correct or false. The mistake of science is the generalization of this issue to all phenomena [7]. Any kind of Fact Expression is not entirely correct or false. Their truth is that between complete and complete incorrectness [9]. In other words, the expression of reality is not correct or incorrect. The fact of them is something between correctness and incorrectness. This is, something between one and zero if we suppose correctness as white and incorrectness as black. Fuzzy logic is something between black and white, gray [7]. Fuzzy logic is a form of logic that is used in some expert systems and artificial intelligence programs, and in which, variables can show degrees of "correctness" or "incorrectness" by the wide range between 1 (correct) and 0 (incorrect) [31]. One of the fuzzy logic advantages is this that one can refrain complex calculations by using linguistic variables, and also these variables make better understanding [32]. Professor Zadeh, first, offered the concept of calculation in 1996 that based on this, fuzzy logic almost equals with words calculations [33, 34]. The structure of fuzzy systems is defined in two types Mamdani and Sugeno. If the conditional rules of the fuzzy system in Sections Antecedent and Consequent are expressed in terms of fuzzy sentences and variables, this system is defined in Mamdani kind. And if the conditional rules of the system phase in Section antecedent are expressed as sentences and fuzzy variables, and in Section Consequent, it is expressed as a function of inputs or a fixed function, a fuzzy system is defined in Sugeno kind. In the fuzzy Sugeno systems, of the section Consequent of the function is constant, the degree is zero, and if section Consequent is defined as a function of degree n, it is called the degree n [35]. To implement a fuzzy system, we must have input and output linguistic variables and expert person's knowledge. If in the description at least one of these cases, there is no enough knowledge or interference, we must use metaheuristic algorithms and soft computing and methods such as Lookup Tables, clustering, GD[13] learning, RLS[14], ANFIS and design fuzzy system from input-output data [10, 36]. In another hand, some time may our peripheral problems become do wide and complicated that increases number of rules. Choosing rules is important, because increasing number of rules may reach cardinal numbers that will complicate the system and decrease speed in intelligent programs [10, 37]. For example, in BDI-II, we have 21 multi-choice questions. If we design this problem without using a certain tool and only by using fuzzy logic, the number of rules in this issue will be computed by relation (1).

$$m^n \tag{1}$$

In relation (1), m represents the number of conditions of any input (number of choices of any question) and n is a number of inputs (number of questions) [38]. And by replacing and computing, we reach the amount of 4,398,046,511,104 of the rule. If we divide 21 questions of BDI-II by five factors and two factors structures, the number of above rules in five factors and two factors structure decreased as 1024 and 16, respectively. In such issues, we also can use clustering technic to decrease the number of rules. So a number of rules equal with the number of clusters [10].

4 ANFIS

ANFIS, architecture to implement non-linear issues with the framework of the system is educable and adaptable that is hybrid of the fuzzy inference system and artificial neuro-fuzzy networks in designing intelligent systems [35, 36]. In this method, by using learning algorithms and data sets, parameters related to a Sugeno fuzzy inference system such as membership functions and rules are produced [39, 40]. Fuzzy inference system and artificial neural networks are supplementary. Because when there are no enough samples to test or direct method to extract rules

[13] Gradient Descent.

[14] Recursive Least Squares.

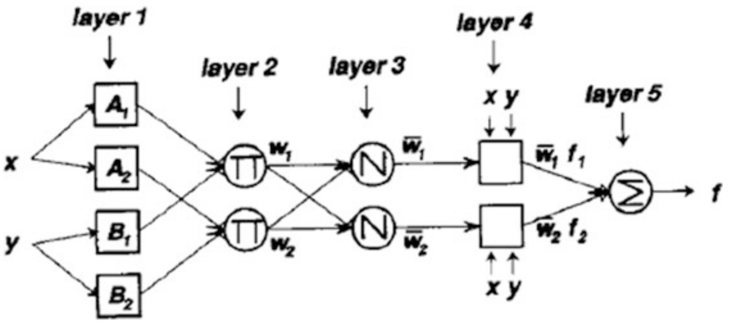


Fig. 2 Structure equals to ANFIS [35]

or have interference, we can use artificial neural networks [36]. Indeed, ANFIS is a logical system called as fuzzy logic to compute hidden uncertainty in data exactly. This work is done by fuzzification input through membership functions that depict diagram-like relation between input amount and interval (1, 0). One can gain these components in input such as membership functions through algorithms such as post-publication or least squares. Then, unlike multilayer perceptron in which weights are updated, in ANFIS, fuzzy linguistic rules or if-then conditional rules orders are determined for the learning process. Figure 2 shows the general structure of an ANFIS. The main problem in ANFIS to modeling is selecting a fuzzy inference system. In this part, the fuzzy inference system is a linear equation and one can compute its parameters by using a simple method of least squares [40, 41].

For example, suppose we have a fuzzy inference system with two inputs x, y and output y. One can use ordinary rules according to relation (2) and (3) with if-then conditional rules to define a fuzzy model of first order Sugeno.

$$\mathrm{Rule(1)}:\ \mathrm{if\,x\,is\,A_1\,and\,y\,is\,B_1\,Then\,f_1} = \mathrm{p_1x + q_1y + r_1} \tag{2}$$

$$\mathrm{Rule(1)}:\ \mathrm{if\,x\,is\,A_2\,and\,y\,is\,B_2\,Then\,f_2} = \mathrm{p_2x + q_2y + r_2} \tag{3}$$

where A_1, A_2, B_1, B_2 are membership functions of fuzzy sets that are considered as input for x and y. in other words, very $\mu_{Ai}(\mathrm{x})$ and $\mu_{Bi}(\mathrm{x})$. Term of $\mu_{Ai}(\mathrm{x})$ is degree of input membership x in A_i set. There are membership functions that can be divided to groups such as Gaussian shape, Triangle, Trapezoid and general bell.

In fact, A_i and B_i convert crisp values to fuzzy kind. Parameters r_1, r_2, p_1, p_2, q_1 are related to output function parameters. ANFIS architecture such as that shown in Fig. 2. All nodes of layers related to ANFIS perform equally. ANFIS performance follows as:

Layer 1: each of nodes of this layer creates membership degree of one of the input components; this is, determines fuzzy parts of input space.

$$\begin{aligned} op_i^1 &= \mu_{A_i}(x) for\, i = 1,2\, or \\ op_i^1 &= \mu_{B_i}(y) fori = 3,4 \end{aligned} \tag{4}$$

In relation (4), x or y are inputs of any node. A_i or B_{i-2} is fuzzy set that interacts with this node and defines as membership functions. Such as functions can be any kind of proper, continued fragment, a derivable function such as Gaussian, general bell, Triangle, and Trapezoid.

Layer 2: nodes send input signals of this layer to each other. This process is shown as p.

op_i^1 That is basic motivate force, forms as relation (5):

$$op_i^2 = w_i = \mu_{A_i}(x).\mu_{B_i}(y), i = 1,2 \tag{5}$$

In other words, this layer is a place of applying operators T-norms, such as product or minimum.

Layer 3: with I^s in this layer (is shown as N), normalizing input signals is done:

$$op_i^3 = \bar{w} = \frac{w_i}{w_1 + w_2}, i = 1,2 \tag{6}$$

Layer 4: with node I^s, in this layer, the participation rate of any rules is computed in output:

$$op_i^4 = \bar{w}f_i = \bar{w}(p_i x + q_i + r_i) \tag{7}$$

Layer 5: this layer computes final ANFIS layer [40]:

$$op_i^5 = \sum_i \bar{w}f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{8}$$

To product, primarily Sugeno fuzzy system, one can use grid partition, subtractive clustering, and clustering FCM methods. Since grid partition method in this issue, while doesn’t decrease the number of rules, increases the number of output membership functions, and the number of rules. Therefore, one can use subtractive clustering and FCM. In this article, we use subtractive clustering. Based on this, the number of rules equals with a number of determined clusters [10]. In this paper, a hybrid learning algorithm is used to train the fuzzy inference network. This algorithm first has introduced by Jang in 1997. In this kind of algorithm, to find optimum parameters of the Sugeno fuzzy inference system, we use gradient descent method and least squares estimate. One of the basic advantages of this algorithm is its speed that is used more in online work [35].

5 Proposed Method

To implement BDI-II by software MATLAB and ANFIS, we need a data set. This data set included the knowledge of expert psychiatry and psychology (10 Persons) and confirming and reliability by them.

6 Implementation and Experimental Results

After receiving data set related to five factors and two factors structure, first divided to 70% for train data, 15% for validation data and 15% for test data. Then clustered train data with subtractive clustering that because the number of output sets of BDI-II has four sets, the number of four clusters devoted to them. Allocating 4 clusters by regulating radial of clustering for train data related to five factors and two factors structure are regulated equally to 0.9222 and 0.7212 respectively. After clustering, two Sugeno fuzzy inference systems are produced. Because of using subtractive clustering, the number of rules becomes equal to a number of determined Sugeno fuzzy inference system; Thus, the number of rules for five factors structure decreased from 1024 to 4 and for two factors structure from 16 to 4. Amounts of parameters of these two Sugeno fuzzy inference systems are shown in Table 3.

The structure of ANFIS for five factors (right) and two factors (left) is shown in Fig. 3.

Sugeno fuzzy inference systems of two factors and five factors structure are trained by using ANFIS and hybrid learning algorithm. Amounts of parameters related to training these two Sugeno fuzzy inference systems are shown in Table 4.

Diagrams of decreasing error related to train five factors (right) and two factors (left) structure have been shown in Fig. 4.

The final amount of decreasing error in this train with four decimal digits, for five factors structure, is 0.0238 for train data, 0.0241 for validation data and for two factors structure is 0.0166 for train data, 0.0169 for validation data. These amounts show that two factors structure has better ability than five factors structure in implement train data related to BDI-II.

Table 3 Parameters of Sugeno fuzzy inference system

Amount	Parameter
Prod	And
Probor	Or
Prod	Implication
Max	Aggregation
Wtaver	Defuzzification
Gaussmf	Input membership function
Linear	Output membership function

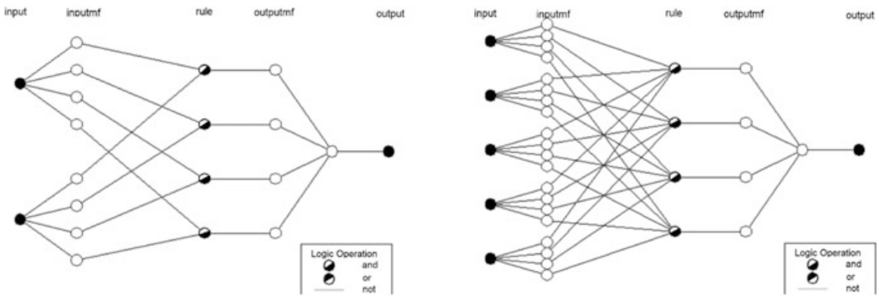


Fig. 3 Five factors and two factors structure of ANFIS

Table 4 Parameters of train

Parameter	Amount
Epoch	1000
Error tolerance	0
Initial step size	0.001
Step size decrease	0.9
Step size increase	1.1

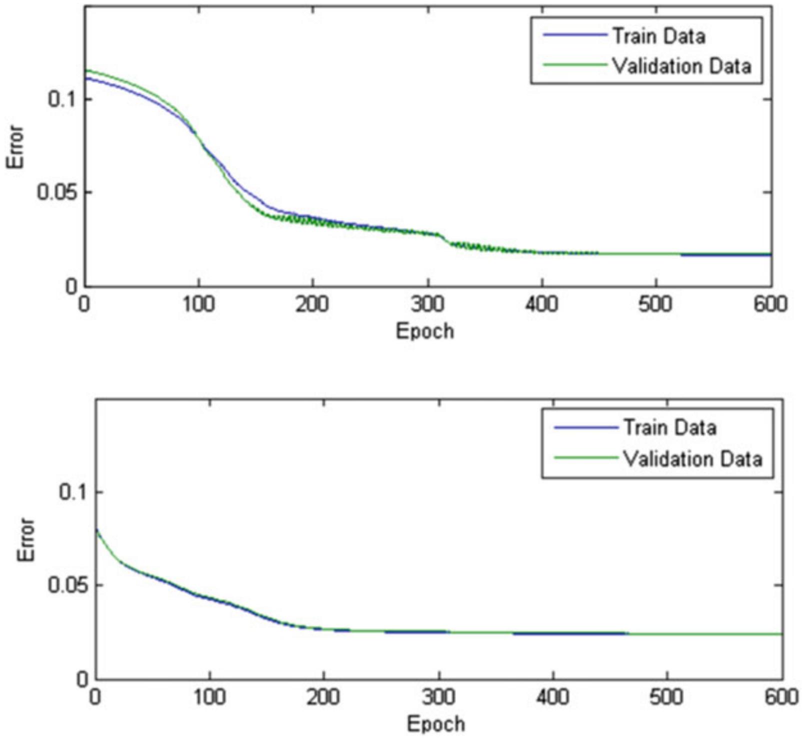


Fig. 4 Diagram of decreasing error in train and validation data

7 Reliability and Validity

Firstly, parameters and diagram of test data error calculated by MATLAB are shown. Then, using software SPSS was used to calculate the Pearson correlation coefficient and diagnostic power.

Error parameters of the test data for two factors and five factors structure in Table 5 shown.

Also, a diagram of error deviation, error histogram, and error regression related to two designed systems related to test data, for five factors structure (right) and two factors structure are shown (left) in Figs. 5, 6 and 7 respectively.

The Pearson correlation coefficient table for five factors and two factors structure by the SPSS software is presented in Table 6.

According to Table 6, two factors have more complete correlation than five factors. To calculate Diagnostic power, used from Cross tabulation in SPSS

Table 5 Comparing error of two designed systems related to test data, with five factors and two factors structure

Parameter	Five factors	Two factors
STD error[a]	0.0233	0.0169
MSE[b]	5.4470e-04	2.8518e-04
RMSE[c]	0.0233	0.0169

[a] Standard deviation error

[b] Mean squared error

[c] Root mean square error

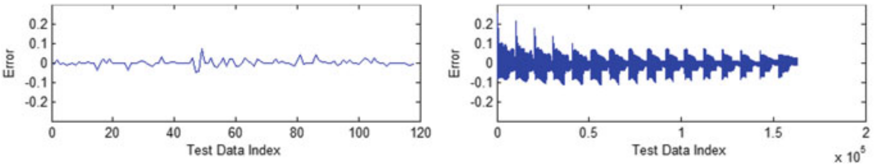


Fig. 5 Diagrams of error deviation related to test data for five factors structure and two factors structure

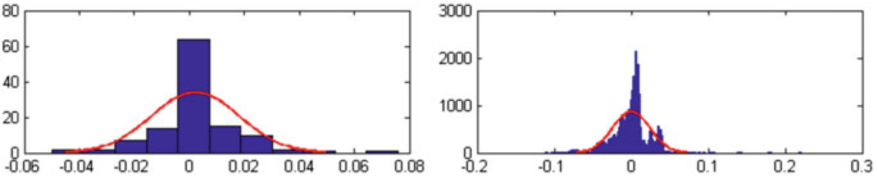


Fig. 6 Diagrams of error histogram related to test data for five factors structure and two factors structure

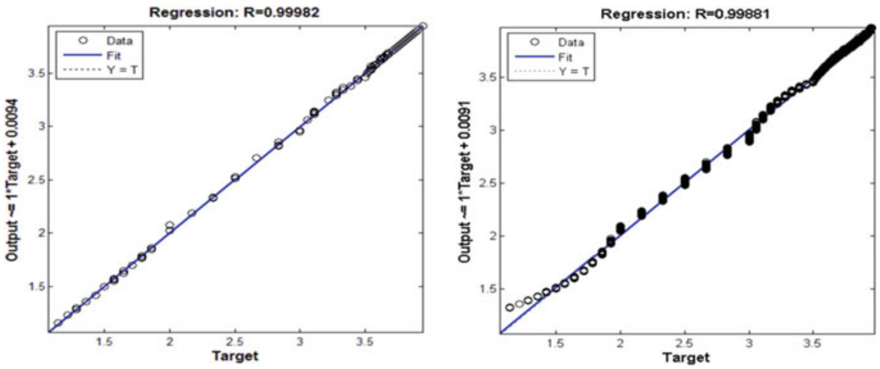


Fig. 7 Diagrams of error regression related to test data for five factors structure and two factors structure

Table 6 Pearson correlation coefficient for five and two factors

	Five factors	Two factors
Pearson coefficient	0.978[a]	1.000[a]
Sig. (2-tailed)	0.000	0.000

[a] Correlation is significant at the 0.01 level (2-tailed)

Table 7 Cross tabulation in five factors structure

Five factors BDI-II	1 (%)	2 (%)	3 (%)	4 (%)
1	3.8	0.0	0.0	0.0
2	0.0	9.5	0.0	0.0
3	0.0	2.3	24.0	0.0
3	0.0	0.0	6.9	53.5

Table 8 Cross tabulation in two factors structure

Two factors BDI-II	1 (%)	2 (%)	3 (%)	4 (%)
1	13.4	0.0	0.0	0.0
2	0.0	12.1	0.0	0.0
3	0.0	2.0	16.3	0.0
3	0.0	0.0	3.7	52.4

software. Diagnostic power for five factors in Table 7 and for two factors in Table 8 shown.

The result of Tables 7 and 8 shown that by the sum of the elements of the main diagonal of the Cross tabulation can calculate the Diagnostic power. Diagnostic power of the five factors is equal to 90.8% and for two factors is equal to 94.2%, Which indicates the higher Diagnostic power of two factors.

8 Discussion

The results of this investigation according to Tables 5, 6, 7 and 8, and Figs. 4, 5, 6 and 7 represent that two factors structure, having advantages such as Diagnostic power 94.2%, Pearson coefficient 1 with Correlation significant at the 0.01 level and correlation coefficients R in regression 0.97724 has implemented BDI-II better than two factors structure. Therefore, designing an ANFIS decision support system with subtractive clustering and hybrid algorithm for BDI-II with two factors structure works better than five factors structure. In [11] compared BDI-II and Chinese version by fuzzy logic. The result showed the Chinese version with diagnostic power 80.3% is better. Yu and Lin [12] implemented score of BDI-II by clustering FCM technic and compared it with clustering k-means. The results showed that clustering FCM has less recognizing cost in recognizing [12]. In Ref. [13], by designing and implementing BDI-II with five factors structure by Sugeno fuzzy inference system, showed that the above system has high value in diagnosis. On the other hand in [17] have used BDI-II with five factors structure with parameters affecting on depression in the research and simulated it through the neuro-fuzzy system and ANFIS. The results showed that using BDI-II with parameters affecting on creating a depression by using expert intelligent system is an effective factor in recognizing depression. In this paper, The results showed that the structure of the two factors had a higher diagnostic power than the five factors. Therefore, it can be concluded that the design of the expert system has been shown to be more effective in determining the severity of depression with ANFIS. In general, it can be admitted that the two factors structure is better able to detect the severity of depression than previous studies.

9 Conclusion and Suggestions

Considering the superiority of the structure of two factors (94.2%), versus to five factors (90.8%) is suggested. To recognize depression in patients or the impact of factors affecting depression, it is better to group the parameters associated with depression into two factors structure. Also, by designing psychological inventories through related methods in soft computing and artificial intelligence, one can design recognizing the online hospital soon and in all regions, also used in rural and remote parts, where there is no need for access to psychology and psychiatry experts, and Introduce patient to equipment centers after specifying depression intensify.

Acknowledgements We are grateful to all our participants that support us to do the best. These grand persons are:

President, secretaries, and manager of computer engineering group of Islamic Azad University of Mahmoudabad, honorable chairman of hospital Rouzbeh, the sincere cooperation of Sir Dr. Hamid Haghani, Dr. Homayoun Amini and personnel of clinic of hospital Rouzbeh dependent on Tehran medical sciences university.

References

1. Yach D, Stuckler D, Brownell KD (2006) Epidemiologic and economic consequences of the global epidemics of obesity and diabetes. Nat Med 12(1):62–66
2. Association AP (2013) Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub
3. Organik D, Tarigan CJ, Penelitian Alb (2003) Perbedaan depresi pada pasien dispepsia fungsional dan
4. Nussbaum AM (2013) The Pocket Guide to the DSM-5® Diagnostic Exam. American Psychiatric Pub
5. Association AP. DSM 5: American Psychiatric Association (2013)
6. Fried EI, Epskamp S, Nesse RM, Tuerlinckx F, Borsboom D (2016) What are 'good' depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. J Affect Disord 189:314–320
7. Kosko B (1993) Fuzzy thinking: the new science of fuzzy logic. Hyperion Books
8. Wang Y-P, Gorenstein C (2013) Assessment of depression in medical patients: a systematic review of the utility of the Beck depression inventory-II. Clinics 68(9):1274–1287
9. Torres A, Nieto JJ (2006) Fuzzy logic in medicine and bioinformatics. BioMed Research International
10. Wang L-X (1999) A course in fuzzy systems. Prentice-Hall press, USA
11. Yu S-C, Yu M-N (2007) Fuzzy partial credit scaling: A valid approach for scoring the Beck Depression Inventory. Soc Behav Pers: Int J 35(9):1163–1172
12. Yu S-C, Lin Y-H (2008) Applications of fuzzy theory on health care: an example of depression disorder classification based on FCM. WSEAS Trans Inf Sci Appl 5(1):31–36
13. Ariyanti RD, Kusumadewi S, Paputungan IV (eds) (2010) Beck Depression Inventory Test Assessment Using Fuzzy Inference System. In: 2010 international conference on intelligent systems, modelling and simulation. IEEE
14. Chattopadhyay S, Kaur P, Rabhi F, Acharya R (eds) (2011) An automated system to diagnose the severity of adult depression. In: Emerging applications of information technology (EAIT), 2011 second international conference on. IEEE
15. Zhang Z-X, Tian X-W, Lim JS (eds) (2011) New algorithm for the depression diagnosis using HRV: a neuro-fuzzy approach. In: Bioelectronics and Bioinformatics (ISBB), 2011 international symposium on. IEEE
16. Chattopadhyay S, Kaur P, Rabhi F, Acharya UR (2012) Neural network approaches to grade adult depression. J Med Syst 36(5):2803–2815
17. Ekong VE, Inyang UG, Onibere EA (2012) Intelligent decision support system for depression diagnosis based on neuro-fuzzy-CBR hybrid. Mod Appl Sci 6(7):79
18. Chattopadhyay S (2013) Mathematical modelling of doctors' perceptions in the diagnosis of depression: a novel approach. Int J Biomed Eng Technol 11(1):1–17
19. Misra K, Chattopadhyay S, Kanhar D (2013) A hybrid expert tool for the diagnosis of depression. J Med imaging Health Inf 3(1):42–47
20. Ekong VE, Ekong UO, Uwadiae EE, Abasiubong F, Onibere EA (2013) A fuzzy inference system for predicting depression risk levels. Afr J Math Comput Sci Res 6(10):197–204
21. Chattopadhyay S (2014) A neuro-fuzzy approach for the diagnosis of depression. Appl Comput Inf 13
22. Mukherjee S, Ashish K, BaranHui N, Chattopadhyay S (2014) Modeling depression data: feed forward neural network vs. radial basis function neural network. Am J Biomed Sci 6 (3):166–174
23. Chattopadhyay S (2014) Neurofuzzy models to automate the grading of old-age depression. Expert Syst 31(1):48–55
24. Mohammadzadeh B (2016) Comparing diagnosis of depression in depressed patients by EEG, based on two algorithms: artificial nerve networks and neuro-fuzy networks. Int J Epidemiol Res 3

25. Beck AT, Steer RA, Brown GK (1996) Manual for the beck depression inventory-II. San Antonio, TX: Psychol Corporation 1:82
26. Wang Y-P, Gorenstein C (2013) Psychometric properties of the Beck depression inventory-II: a comprehensive review. Revista Brasileira de Psiquiatria 35(4):416–431
27. Storch EA, Roberti JW, Roth DA (2004) Factor structure, concurrent validity, and internal consistency of the beck depression inventory—second edition in a sample of college students. Depress Anxiety 19(3):187–189
28. Stevenson A (2010) Oxford dictionary of English. Oxford University Press, USA
29. Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353
30. John R, Coupland S (2007) Type-2 fuzzy logic: a historical view. IEEE Comput Intell Mag 2 (1):57–62
31. Aiken P (2002) Microsoft computer dictionary. Microsoft Press
32. Yu TH-K, Wang DH-M, Chen S-J (2006) A fuzzy logic approach to modeling the underground economy in Taiwan. Physica A: Stat Mech Appl 362(2):471–479
33. Zadeh LA (1996) Fuzzy logic = computing with words. IEEE Trans Fuzzy Syst 4(2):103–111
34. Zadeh LA (1999) Fuzzy logic = computing with words. Computing with words in information/intelligent systems 1. Springer, pp 3–23
35. Jang J-S (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23(3):665–685
36. Abraham A (2005) Adaptation of fuzzy inference system using neural learning. Springer, Fuzzy systems engineering, pp 53–83
37. Ledeneva Y, Gelbukh A, García CAR, Hernandez RAG (eds) (2007) Automatic determination of parameters for rule base reduction of complex fuzzy control systems. In: 8th conference on computing CORE-2007. Mexico City, Mexico
38. Ledeneva YN, García CAR, Méndez JAD (eds) (2007) Automatic estimation of parameters for the hierarchical reduction of rules of complex fuzzy controllers. ICINCO-ICSO
39. Jang J-SR (ed) (1991) Fuzzy modeling using generalized neural networks and Kalman Filter algorithm. AAAI
40. Rahmani Seryasat O, Haddadnia J, Ghayoumi Zadeh H (2016) Assessment of a novel computer aided mass diagnosis system in mammograms. Iran J Breast Dis 9(3):31–41
41. Maghsoudi R, Moshiri B (2017) Applying adaptive network-based fuzzy inference system to predict travel time in highways for intelligent transportation systems. J Adv Comput Res 8 (3):87–103

Part II
Control Engineering

Self-tuning PD2-PID Controller Design by Using Fuzzy Logic for Ball and Beam System

Milad Ahmadi and Hamed Khodadadi

Abstract Nowadays, the science of aircraft equipment has made remarkable progress. However, due to the heavy costs and sensitivity, study on a laboratory scale for this equipment is impossible. Therefore, equivalent systems are used. The ball and beam is one of the systems which is used to simulate the aviation sciences. This system is composed of one motor, one ball and one beam. The purpose of this system is controlling and balancing the ball position on the beam at a desired value. There are several challenges like sensors noise and servo motor nonlinearities in the controlling this system. To solve these problems, a new control method is presented in this paper. This approach is made from the combination of two different methods consist of PD2-PID and fuzzy logic controller. The main purpose of the presented approach is realizing the best performance of the system and locating the ball position in desired value in the lowest time, steady state error and overshot. The obtained simulation results indicate the proposed method has better performance in controlling the ball and beam system compared to the other techniques like traditional PD2-PID.

Keywords Ball and beam · PD2-PID controller · Fuzzy logic
Fuzzy logic self-tuning PD2-PID controller

1 Introduction

The ball and beam system has a simple structure. By controlling the dc motor, the system can balance the ball at the desired position. This system is highly nonlinear and unstable, makes several challenges in its control. This system has been considered by many researchers who have applied various control methods to improve

M. Ahmadi · H. Khodadadi (✉)
Department of Electrical Engineering, Islamic Azad University,
Khomeinishahr Branch, Isfahan, Iran
e-mail: khodadadi@iaukhsh.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_16

system performance. The purpose of all of these methods are balancing the ball in a specified position with the minimum error in the fastest time. Some of the applied methods in the literature are feedback linearization considering system singularities [1], PID-controller [2], optimal PID controller [3], model based and non-model based control approaches [4], Coefficient Diagram Method (CDM) based PID controller [5], Extended State Observer (ESO) based LQR controller [6], State Dependent Riccati Equation (SDRE) controller [7], neural network [8] and fuzzy logic control [9]. In addition, some more advanced methods such as fuzzy PID controller [10] have been applied for ball and beam system.

In this paper, a Fuzzy Logic Self-Tuning (FLST) PD2-PID is designed for the nonlinear mathematical model of the system. This structure composed from two PD controllers cascaded by a PID controller. Using the PD2 controller create more flexibility for the system and increase the Degree of Freedom (DOF) of the controller. In addition, combining the PID controller with the fuzzy logic caused the controller can compensate the system uncertainties. Comparison between the obtained results of the proposed controller and PD2-PID controller show the ability of the FLST PD2-PID controller in yielding to the desired objectives.

The remainder of this paper is organized as follows. In Sect. 2 the mathematical model of the ball and beam system is presented. Section 3 is dedicated to the principle of PID controller. Section 4 describes the fuzzy Logic controller. Section 5 indicates the FLST PD2-PID as the proposed approach of the paper. Section 6 is dedicated to presenting the simulation results of designing the mentioned controllers for the ball and beam system and the resultant discussion. This part comprises the ability of the proposed method and the other controllers At the end, the article is concluded in Sect. 7.

2 Mathematical Modeling of the Ball and Beam

The ball and beam system schematic is presented in Fig. 1. This structure is very common for deriving the system model. When the ball rolls on the beam freely, the beam makes an angle with horizontal axis (θ). With this angle, the ball faces with an incline. This slope makes the force that pulls to the ball. The equation of this force is given as:

$$F = mg\,\sin(\theta) \tag{1}$$

In Eq. (1), m is ball mass and g is the gravity constant. The main purpose of the ball and beam system is standing against this force when the ball is getting out of the system. The rod accomplishes this by reducing its angle with the horizontal axis. Using the second Newton's laws yields:

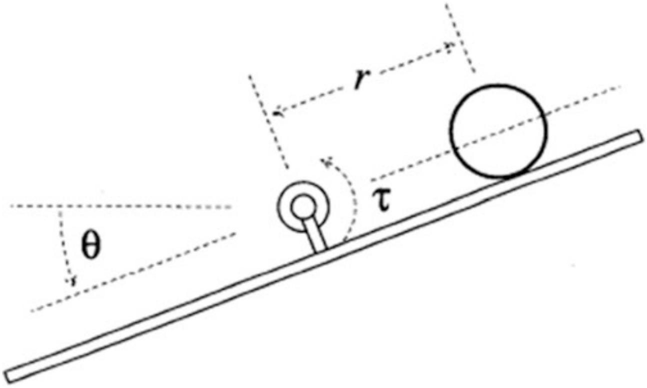


Fig. 1 A ball and beam system [1]

$$\sum f - \sum R = a \sum m \tag{2}$$

where $\sum f$ should be stated as sum of forces in same direction with movement and $\sum R$ is sum of forces opposite with the movement. When the ball faces a slope, Eq. (1) works as $\sum f$ and beam friction as $\sum R$. Because the friction of the beam is very low, a DC motor can control the force in Eq. (1).

The state space model of the ball and beam system can be describe as (3). In this equation, $x_1 = r$ is the distance between the ball and the pivot; $x_3 = \theta$ is the created angle between the beam and the horizontal axis; B and g are some constant numerical values and g describes the gravity constant; the input u is a nonlinear transformation of the torque τ [1].

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} \dot{x}_2 \\ B\left(x_1 x_4^2 - g \, \sin x_3\right) \\ x_4 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = f(x) + g.u \tag{3}$$

where $x = (x_1, x_2, x_3, x_4)^T = (r, \dot{r}, \theta, \dot{\theta})^T$ and $y = h(x) = r$.

3 PID Controller

PID controller is one of the common controller computes the control signal as follows [11]:

$$u(t) = k_p e(t) + k_i \int\limits_0^t e(t) + k_d \frac{d}{dt} e(t) \tag{4}$$

However this controller performs well for a variety of system, it hasn't good performance for all systems. Adding the PD controller to the mentioned structure can improve the system performance.

4 Fuzzy Logic Controller

Intelligent methods can be used to improve the system performance. One of these methods is fuzzy logic can be applied for various unstable and controllable systems. The inputs of fuzzy logic controller for this system are the ball position error and ball velocity error. After the inputs are inferenced by the rule database of fuzzy controller, the appropriate outputs will be generated and applied to the system [12]. For inputs of the fuzzy controller, five levels are considered as follows: Negative Large (NL), Negative Small (NS), Zero (ZE), Positive Small (PS) and Positive Large (PL). For the output variables seven levels are considered as Positive Very Small (PVS), Positive Small (PS), Positive Medium Small (PMS), Positive Medium (PM), Positive Medium Large (PML), Positive Large (PL) and Positive Very Large (PVL). As an example, the membership function of the system outputs is presented in Fig. 2 [12]. Moreover, the fuzzy rules are presented in Table 1.

5 Fuzzy Logic Self-tuning PD2-PID

To achieve the best performance and the most robust controller, the FLST PD2-PID controller is designed for the non-linear system. Against classical PID controller, in which the PID coefficients are constant, the parameters of FLST PD2-PID are continuously changing during the process.

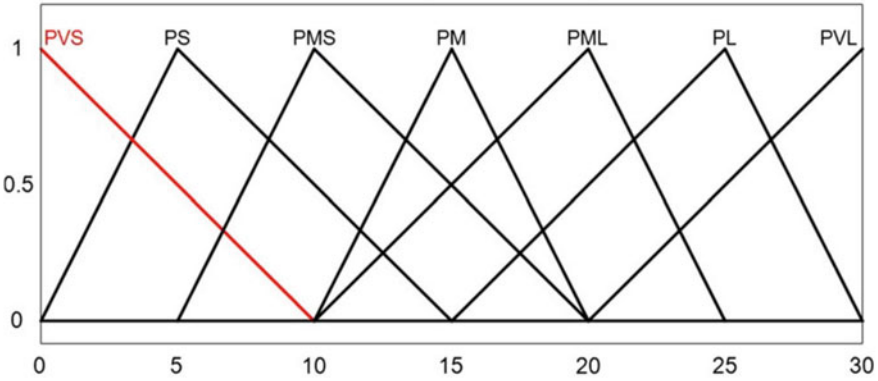


Fig. 2 Membership functions of the output variable [12]

Table 1 Fuzzy rules

ve/ce	NL	NS	ZE	PS	PL
NL	PVL	PVL	PVL	PVL	PVL
NS	PML	PML	PML	PML	PML
ZE	PVS	PVS	PS	PMS	PMS
PS	PML	PML	PML	PM	PM
PL	PVL	PVL	PVL	PVL	PVL

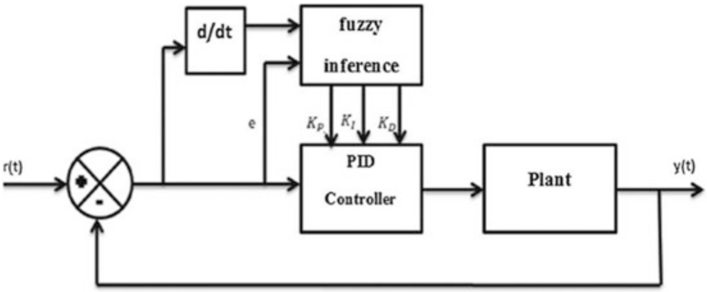


Fig. 3 The structure of FLST-PID controller [13]

Due to the combination of two different structure of the controller, this method is very resistant to system uncertainties and can attenuate the effects of disturbances, noise and changing the operating conditions on the system response. These benefits of FLST PD2-PID are evident in the simulation results. In this paper, by changing the operating condition, the PID's coefficients will change. The base of variation in PID's parameters is employing the fuzzy logic in a determined range. Fuzzy logic controller take the position error and ball velocity error as its inputs and appear the PD2-PID terms in the output. Finally, by tuning the controller parameters and applying to the system, a fast, stable, and robust response is achieved. The schematic of a FLST PID controller is illustrated in Fig. 3.

6 Simulation Results

In this section, the simulation of ball and beam system is done based on the proposed structure of controllers. By applying the PD2-PID and FLST PD2-PID controllers on the system mathematical model the below results are yielded. In Fig. 4 the step response of the ball and beam system to the both of proposed controllers are illustrated. In addition, the control signals of the controllers are shown in Fig. 5.

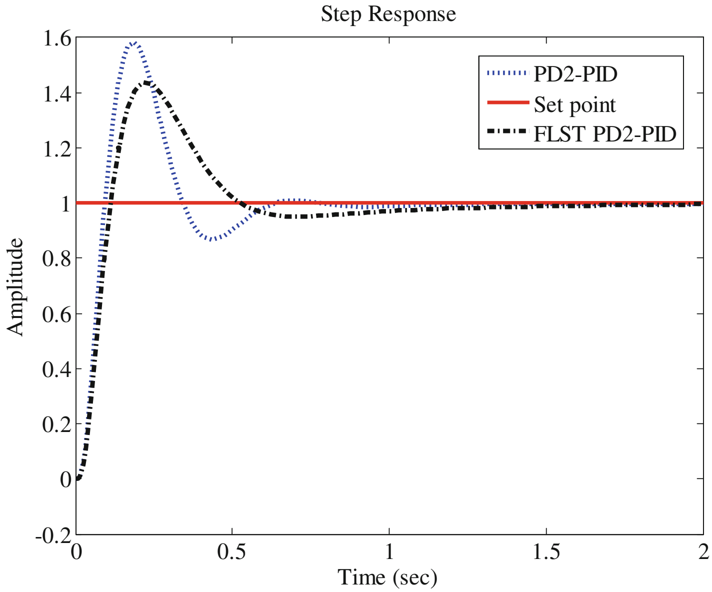


Fig. 4 Step response of the system

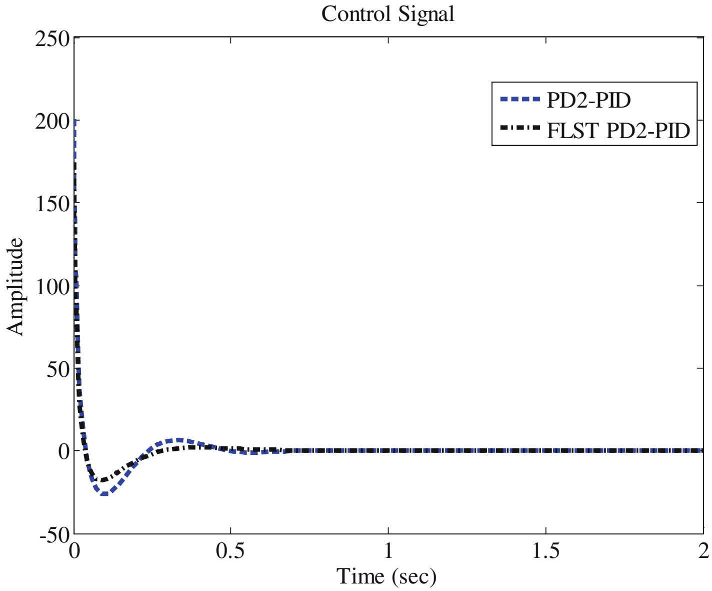


Fig. 5 Control signal for both controllers

Table 2 Comparison based on step response

	Overshoot (%)	Rise time (s)	Settling time (s)
PD2-PID	59	0.14	1.9
FLST PD2-PID	42	0.13	1.7

Table 2 illustrate the performance of the proposed controllers evaluated from step response. As can be comprehend, although both of the proposed controllers have a good performance in balancing and locating the ball in the desired position, the FLST PD2-PID has the lowest rise time and overshoot. In addition, the settling time and control signal of this controller are lower than PD2-PID.

For evaluating the ability of the proposed controllers in disturbance attenuation, by assuming a pulse disturbance applied in second 4, the step responses and control signals of the system are shown in Figs. 6 and 7, respectively. The effect of disturbance on system response are decreased for two controllers, albeit the disturbance attenuation is faster in FLST PD2-PID.

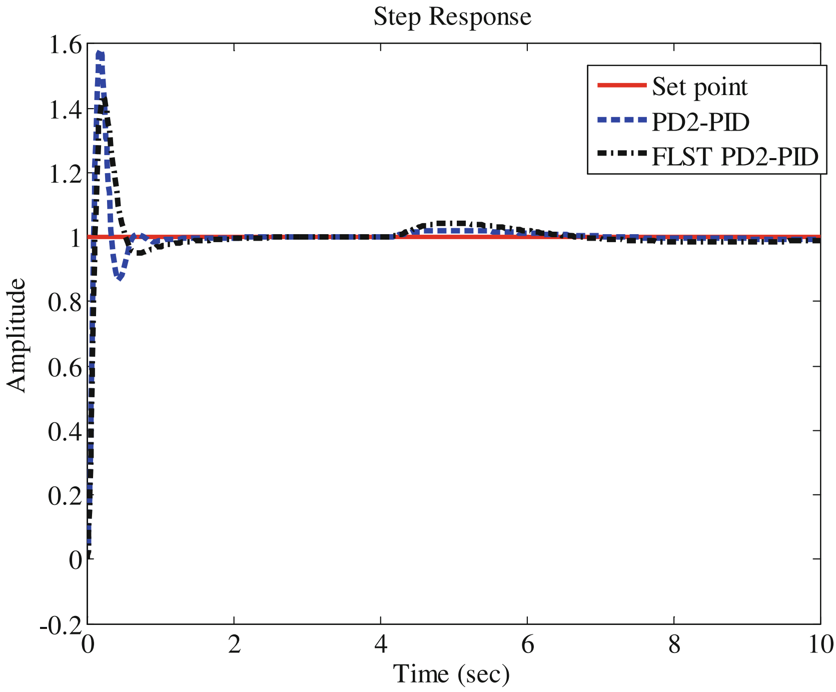


Fig. 6 Step response of the system in presence of disturbance

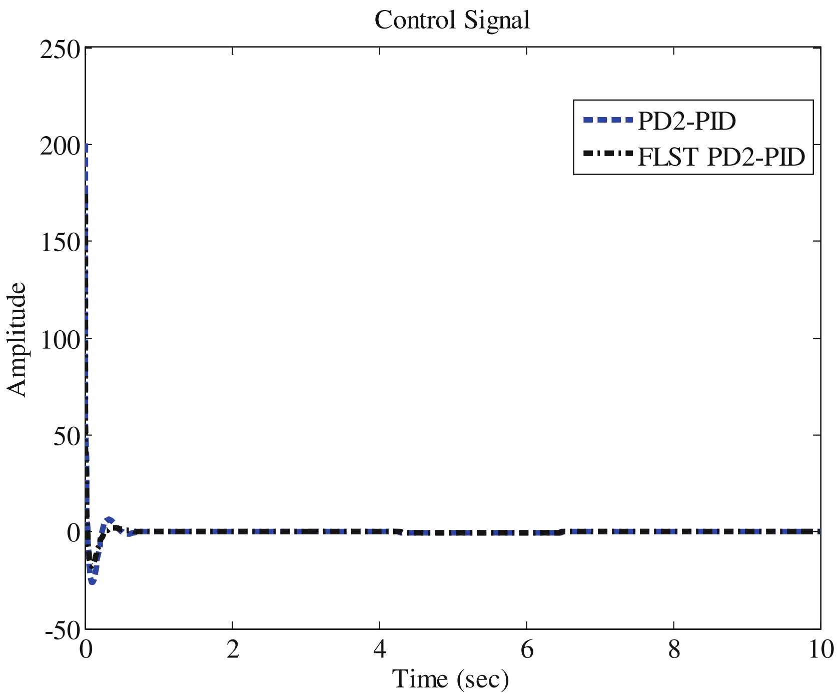


Fig. 7 Control signal for both controllers in presence of disturbance

7 Conclusion

In this paper, a STFL PD2-PID controller is suggested for ball and beam system. The main purpose of controlling this system is locating the ball in its desired position on beam in the lowest time and overshot. Moreover, the effect of uncertainty and disturbances on system performance should be decreased. Simulation results indicate STFL PD2-PID has a better performance than the other controller and show the ability of proposed controller in achieving to the best response.

References

1. Zhang F, Fernndez-Rodriguez B (2006) Feedback linearization control of systems with singularities. In: The 6th international conference on complex systems (ICCS). Boston, MA
2. Mana Maalini PV, Prabhakar G, Selvaperumal S (2016) Modelling and Control of Ball and Beam System using PID Controller. In: International conference on advanced communication control and computing technologies (ICACCCT)
3. Prasad KT, Hote YV (2014) Optimal PID Controller for Ball and Beam System. In: IEEE international conference on recent advances and innovations in engineering (ICRAIE). Jaipur, India, 09–11 May 2014

4. Keshmiri M, Fellah Jahromi A, Mohebbi A, Amoozgar MH, Xie WF (2012) Modeling and control of ball and beam system using model based and non-model based control approches. Int J Smart Sens Intell Syst 5(1), March 2012
5. Meenakshipriya B, Kalpana K (2014) Modelling and control of ball and beam system using coefficient diagram method (CDM) based PID controller. In: Third international conference on advances in control and optimization of dynamical systems (ACODS). Kanpur, India, 13–15 March 2014
6. Choudhary MK, Naresh Kumar G (2016) ESO based LQR controller for ball and beam system. In: 4th IFAC conference on advances in control and optimization of dynamical systems (ACODS). Tiruchirappalli, India, 1–5 Feb 2016
7. Vinodh Kumar E, Jerome J, Raaga G (2014) State dependent Riccati equation based nonlinear controller design for ball and beam system. In: 12th global congress on manufacturing and management (GCMM)
8. Rahmat MF, Wahid H, Wahab NA (2010) Application of Intelligent Controller in a Ball and Beam Control System. Int J Smart Sens Intell Syst 3(1), March 2010
9. Amjad M, Kashif MI, Abdullah SS, Shareef Z (2010) Fuzzy logic control of ball and beam system. In: 2nd international conference on education technology and computer (ICETC)
10. Aziz NSA, Adnan R, Tajjudin M (2017) Design and evaluation of fuzzy PID Controller for ball and beam system. In: IEEE 8th control and system graduate research colloquium (ICSGRC). Shah Alam, Malaysia, 4–5 Aug 2017
11. Khodadadi H, Dehghani A (2015) Fuzzy logic self-tuning PID control for a single-link flexible joint robot manipulator in the presence of uncertainty. In: IEEE, 15th international conference on control, automation and systems (ICCAS), pp 186–191
12. Khodadadi H, Ghadiri H (2018) Self-tuning PID controller design using fuzzy logic for half car active suspension system. Int J Dyn Control 6(1):224–232 https://doi.org/10.1007/s40435-016-0291-5
13. Akbari-Hasanjani R, Javadi S, Sabbaghi-Nadooshan R (2014) DC motor speed control by self-tuning fuzzy PID algorithm. Trans Inst Meas Control 37(2):164–176

Design of Automatic Gain Control (AGC) Circuit for Using in a Laboratory Military Submarine Sonar Systems Based on Native Knowledge

Davood Jowkar, Mohammad Reza Bahmani, Mohammad Bagher Jowkar, Ali Shourvarzi and Ameneh Jowkar

Abstract Some of the applications of submarine communication are collecting the oceanography data, marine contaminations control, marine studies, military systems like relations between submarines and ships, and identification of submarine targets and torpedo. The principle of operation in submarine communication is based on sound acoustic emission which is utilized in a military sonar system in this paper. AGC is one of the most important parts of a sonar receiver. In this paper, designation of a circuit (based on a new topology) high quality separate automatic gain control (AGC) in a sonar receiver system is proposed and the results are compared with the previously designed circuits for RF and acoustic systems. The most important use of this circuit is in military industry, submarine communication, and ships.

Keywords AGC · Submarine communication · Acoustic emission

1 Introduction

Submarine communication after the second war world on the year 1945, in which the submarine telephone for communications between submarines was designed, became significant. In submarine communication networks, sound transmission is a physical layer usual technology. In fact, to avoid the electromagnetic wave losses, it is necessary for the waves to emit in low frequencies (30–300 Hz) which brings the

D. Jowkar (✉) · M. B. Jowkar
Young Researchers and Elite Club, Islamic Azad University, Dariun Branch, Dariun, Iran
e-mail: Engjowkar@gmail.com

M. R. Bahmani
Young Researchers and Elite Club, Islamic Azad University, Shiraz Branch, Shiraz, Iran

A. Shourvarzi
Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

A. Jowkar
Department of Physics, Islamic Azad University, Marvdasht Branch, Marvdasht, Iran

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_17

needs to long antennas and high transmission power [1]. Compare to radio frequency waves, light waves has got lower attenuation, but dispersion influences their emission. Furthermore, transmitting light signals needs precise directing of narrow laser beam, Therefore, the links in submarine communication are based on basic challenges in acoustic wireless communication [1–5].

Submarine communications have numerous applications in military industries such as relations between submarines and ships, and identification of submarine targets and torpedo. Also, the military submarine equipments which are equipped with some special sensors are capable of collecting the required information. To achieve these goals it is necess. In the first years of radio circuits' invention, because of the need to maintain a relative constant output signal, fading "defined as slow variations in the amplitude of the received signals" required continuing adjustments in the receiver's gain. The result of this state was the design of circuits, which primary ideal function was to maintain a constant signal level at the output, regardless of the signal's variations at the input of the system. Mainly, those circuits were defined as automatic volume control circuits, a few years later they were distributed under the name of Automatic Gain Control (AGC) circuits [6, 7].

During the second half of the 20 century by growth of communication systems, selectivity and good control of the output signal's level became a vital need in the design of any communication system. Nowadays, AGC circuits are available in any device or system which their wide amplitude variations in the output signal could result in a lost of information or to an insufficient performance of the system. AGC's usual application is in receiver for regulating an arbitrary input-signal to some specified power lever. Different AGCs have different advantages and disadvantages [8, 9].

2 SONAR[1] Systems

During World War I the need for submarines detection resulted in more research about the use of sound. The earliest studies and utilizing of underwater hydrophones were done by the British, while the French physicist Paul Langevin, cooperated with a Russian immigrant electrical engineer, Constantin Chilowski, on the development of active sound devices in order to detect submarines in 1915 by means of quartz. Later, the electrostatic transducers superseded by the piezoelectric and magnetostrictive sensors. But, their work had a great effect on the future designs. The main usage of lightweight sound-sensitive plastic film and fibre optics was in hydrophones, while Terfenol-D and PMN[2] goal to be developed was for their application in projectors.

[1] Sound Navigation and Ranging.

[2] Plead Magnesium Niobate.

A sound transmitter and a receiver are in use by active sonar. Monostatic operation is the case in which both of them (transmitter and receivers) are in the same place. When the transmitter and receiver are separated it is called as bistatic operation. Multistatic operation is the condition in which more transmitters (or more receivers) are used, again spatially separated. Most of sonar systems work in monostatical state and they have the same array often being used for transmission and reception. The operation of active sonobuoy fields may be in the multistatical condition.

Active sonar produces a pulse of sound, often called a "ping", and then it is ready to detect reflections (echo) of the pulse. The creation of this pulse of sound is usually performed electronically via a sonar projector which is organized by a signal generator, power amplifier and electro-acoustic transducer/array. A beam former is usually utilized to concentrate the acoustic power into a beam, which may be swept to fulfil all the required search angles. The electro-acoustic transducers are usually of the Tonpilz type and their design may be optimised to have maximum efficiency over the widest bandwidth which can bring the optimise performance of the overall system. Sometimes, some other ways may be utilized to create the acoustic pulse, e.g. (1) utilizing explosives (chemically), or (2) airguns or (3) plasma sound sources.

3 Basic Principles in a Typical Active Sonar System

Figure 1 shows the schematic of the basic principles in a sonar system. In this structure, an oscillator starts to produce and amplify the wave. Sonar projector emits this acoustic wave under water. This wave will be received by the hydrophone after coming back. The sonar receiver is a device which applies some changes on the received signal shape. A sample for the sonar receiver is shown in Fig. 2. It is possible to add or eliminate some parts to or from this sample scheme. The received signal from the sonar sensor (hydrophone) after a pre-amplification will experiment a filtering.

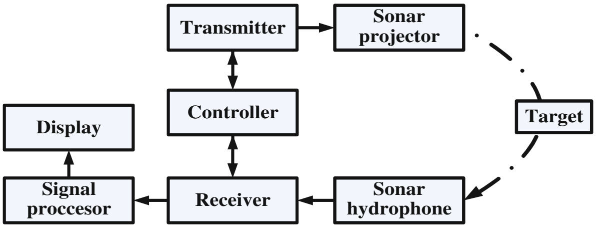


Fig. 1 A typical structure for a sonar active system

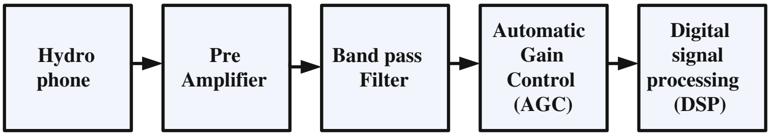


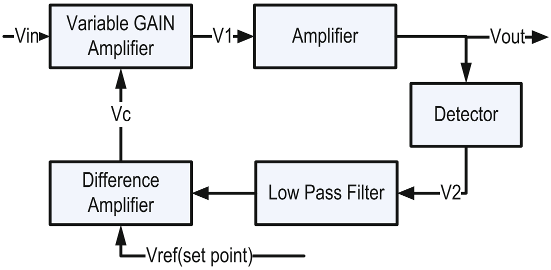
Fig. 2 The whole structure for a typical sonar receiver

The filter is designed as a pass-band one. So, the required frequency range in processing operations will pass through the filter and the remained ranges will be blocked. In the AGC circuit the signal amplitude will be corrected to maintain the amplitude. If the amplitude is large the amplification coefficient will be reduced and vice versa. By adjusting the signal amplification coefficient, a signal with proper amplitude for ADC operations is accessible in a wide range from the signal domain. The wider amplitude of the DSP entered signal leads in higher resolution. Of course, if the signal magnitude exceeds the saturation voltage the signal will be cut and ADC cannot perform its converting operations correctly. So, the high quality level of the AGC can play basic role to increase the whole system quality.

4 The Basic Theory of the AGC Circuits

The whole block diagram for an AGC circuit is shown in Fig. 3 [10]. A variable gain amplifier (VGA) which its gain is controlled by an external signal VC is used to amplify the input signal. To generate a sufficient level of Vout, the output from the VGA can be amplified more by a second stage. The detector senses some the output signal's parameters, like amplitude, carrier frequency, index of modulation or frequency. Any undesired component is blocked via a filtering and a comparison between the remaining signal and a reference signal is performed. The result of such comparison is the main importance in generating control voltage (VC) and adjusting the gain of the VGA [11].

Fig. 3 AGC block diagram



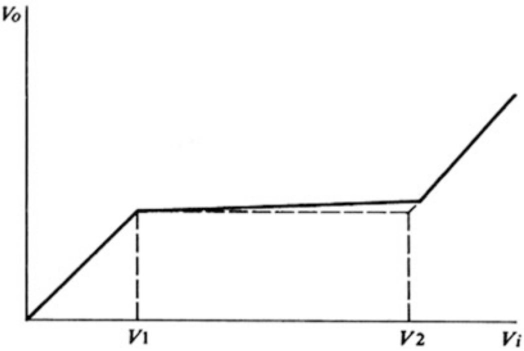


Fig. 4 AGC's ideal transfer function [10]

An AGC is basically a negative feedback system. So it can be defined via its transfer function. Figure 4 shows the idealized transfer function for an AGC system. The AGC is disabled when the input signal is low, and the output dependence on the input is based on a linear function. The AGC enters the operative situation when the output reaches a threshold value (V1) and the output remains constant till a second threshold value (V2) is reached. This is the time in which the AGC becomes inoperative again. This condition is usually occurs to prevent stability problems when the gain is in a high level.

5 The Proposed Idea

In the previous works [4, 10], the AGC circuit is utilized in as an analogue part in the RF or acoustic receiver stage. As it is mentioned in Fig. 3, it is tried to keep the output stable against the input fluctuations. In such circuits there is no need to have a large transfer function. In the other words, low differential voltage between V1 and V2 Fig. 4 is needed. So this circuits works in small voltage ranges. Also, due to the high working frequency of the circuits, the AGC must work in a high speed level. Vice versa, in sonar receiver circuits the differential voltage between V1 and V2 is so large. The reason is the large power decrease of the acoustic waves in water in comparison to the distance. In the other words, with an increase in the distance between the sonar projector and hydrophone the induced voltage in hydrophone will be reduced seriously.

This voltage changes slowly in the time domain. System working frequencies are in the acoustic or ultra sonic (20 Hz–100 kHz) ranges and frequency changes are very small. Due to these requirements, appropriate AGC circuit is needed.

Besides, in the designed AGC circuits for the other mentioned applications, because of their use in the RF or acoustic systems, it is not important to define the primary voltage domain of the input signal and the main point is to have a signal with proper voltage domain in the output. While, in sonar circuits there is important information in the input signal amplitude.

The AGC circuit must meet these needs:

Appropriate output amplitude despite the fast changes in the input amplitude.

Transferring the output voltage range to the DSP, so by multiplying it to the amplitude of the received signal from the AGC the computer can monitor the output signal.

Better operation speed in comparison with the AGCs which have other applications.

6 The Proposed Circuit Structure

The block diagram for the utilized AGC circuit is shown in Fig. 5. In this structure, the input analog voltage enters the divider section after primary amplification. In this section, the voltage will be lowered due to the magnitude of the impedance. Greater amounts of impedance cause more reduction in the output magnitude. These two stages in company play the role of the AGC. The output of the divider stages is in a connection with the DSP system. A sample is taken from the this output and after the transmitting through a low pass filter, it is fed to the input of an ADC. In fact, this ADC has the mission to produce an n-bit digital number from the output peak voltage. A processor take this number and compare it with the set point voltage. Set point voltage is the needed voltage value in the DSP system which is given to the AGC by the DSP system. The AGC processor compares the set point and ADC output number thousands times in a second and after some arithmetic operations produces a digital number as the output and fed this number back to the divider circuit. This number is also given to the DSP circuit. In the DSP circuit through some arithmetic operations and multiplying the mentioned number to the output of the ADC, the input signal voltage will be resulted.

The principles of the operation in the AGC circuit are as follows. The input voltage after the primary amplification and dividing arrives at the output. Output

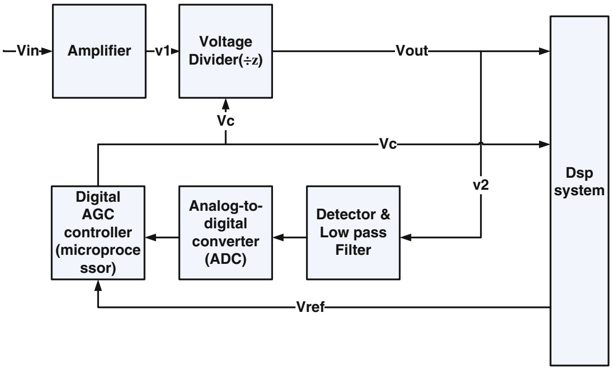

Vin
Amplifier
v1
Voltage Divider(÷z)
Vout
Vc
Vc
v2
Dsp system
Digital AGC controller (microprocessor)
Analog-to-digital converter (ADC)
Detector & Low pass Filter
Vref

Fig. 5 The block diagram of the proposed structure

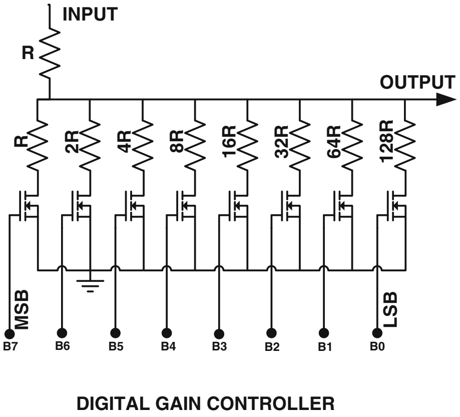


Fig. 6 The practical circuit for step-by-step gain control

control is performed by the divider number. This number is set in the AGC processor in a way that the AGC output remains fixed despite the input fluctuations. The different sections of Fig. 5, are going to be described as circuits in the following:

The step-by-step gain reduction circuit (divider):

In the circuit which is shown in Fig. 6 a series of MOSFET switches are set in parallel with the resistor. As the switches are closed the resistors will be in the current path cause the gain reduction. The peak gain is equal to unit (one). In general, the gain can be obtained from the Eq. (1).

The parameters B0–B7 are the digits that fed to the step-by-step gain control circuit through the processor. B0 which is the list significant bit (LSB) has the list effect on gain and B7 which is the most significant bit (MSB) has the most effect on the gain. This binary number can vary between 00000000 and 11111111. LSB is related to the unit gain and MSB is a sign of the maximum gain (0.33).

$$AV = \frac{R}{R + R[B0/128 + B1/64 + B2/32 + B3/16 + B4/8 + B5/4 + B6/2 + B7]} \quad (1)$$

Figure 6 shows a practical circuit utilizing switching MOSFETs which can reduce the gain step-by-step through changing the value of the gates.

Peak detector circuit and the low pass filter (integrator):

This circuit detects the signal peak voltage. Performing such application can be available utilizing an active diode. After that, a capacitor filter changes the peak voltage to a dc voltage with a limited varying range. The output dc voltage changes proportional to the input voltage variations Fig. 7.

The overall scheme of the designed AGC circuit:

In the circuit shown in Fig. 8, first the input signal is amplified by the Opamp(1). The circuit gain can be obtained from: Av = −R2/R1.

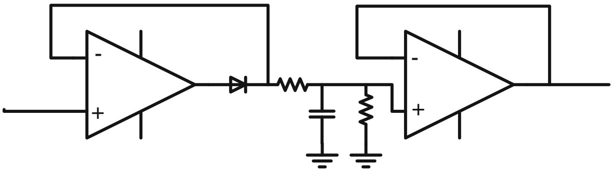
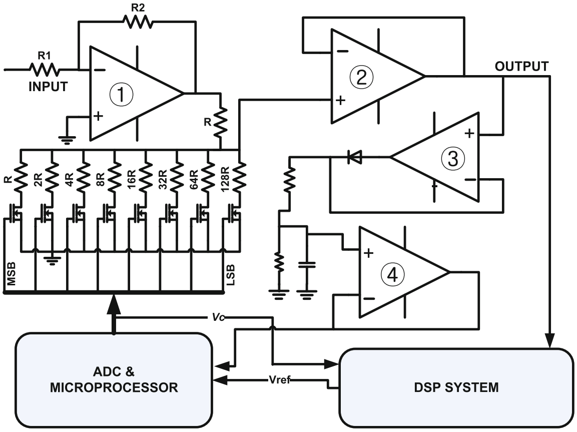


Fig. 7 Peak detector circuit and the low pass filter (integrator)

Fig. 8 Overall AGC circuit designed specifically to work in sonar systems



The gain is reduced through resistance dividing. When the MOSFETs are turned on through the control circuit, the resistors which are in series with MOSFETs will enter the circuit function and proportional to the added resistance the gain will be reduced. A buffer (Op-Amp 2) has the duty of impedance matching. This buffer applies the divided input voltage to the output. The circuit output is connected to the DSP circuit. The operational amplifiers 3 and 4 supply the needed dc voltage adequate to the output ac voltage. This voltage enters an ADC circuit and the resulted digital number will be compared to the set point, which is applied by the DSP, in the processor. Finally, through the online processing by the system processor the gain voltage of the MOSFETs from switching the parallel resistors will be applied. The digital number which is applied to the gated enters the DSP system, too. In this system, the DSP processor can simulate the hydrophone output signal through instantaneous multiplying of the entered digital number to its output.

7 Simulation and Results

The proposed AGC circuit is fabricated and tested experimentally. This circuit shows these advantageous comparing with the traditional AGC circuits:

Large changes of the output voltage toward the input voltage.

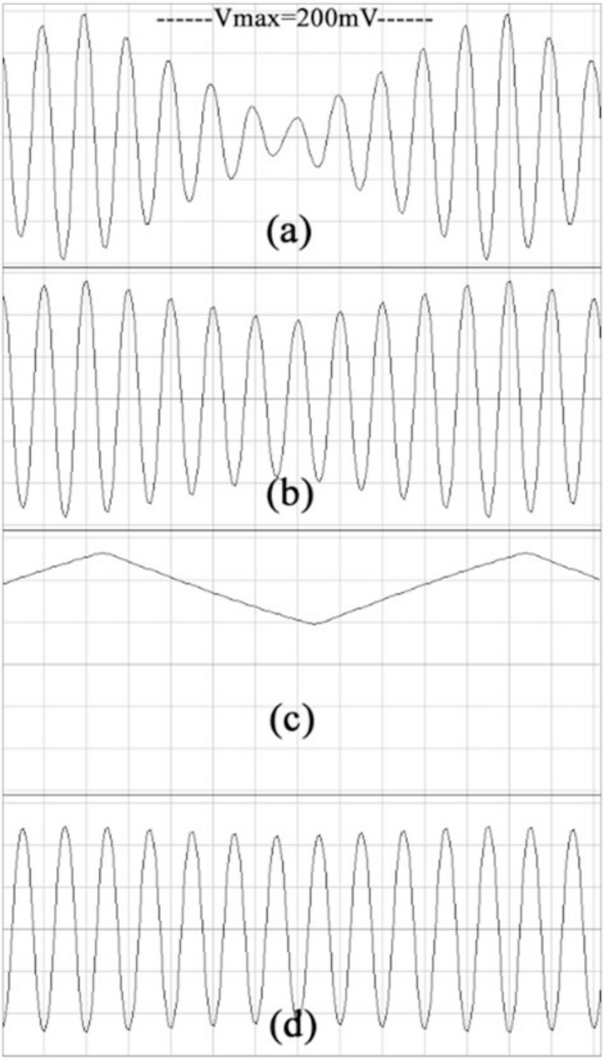


Fig. 9 The comparison of input and output signals: **a** input voltage **b** the output voltage of a typical AGC **c** this diagram shows the binary number applied to the MOSFETS **d** the output of the proposed AGC circuit

Great frequency response because of the utilized step-by-step gain reduction circuit and using resistor in voltage dividing.

The existence of the gain change rate in output through which the DSP can simulate the input signal amplitude.

Simple design and mass production availability.

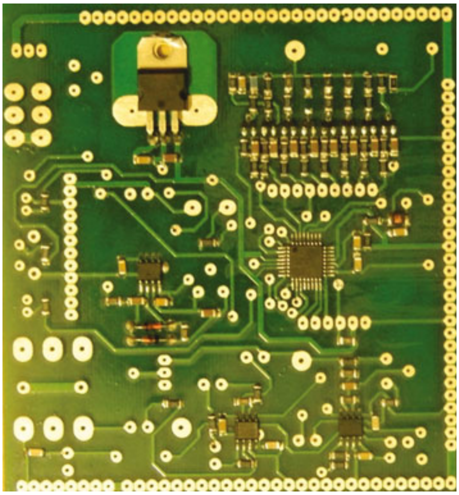


Fig. 10 Practical circuit, made for tests and applications

The possibility to bit number development by using divider circuit with more bit numbers. In this paper 8bit divider was utilized and by using circuits with more bit numbers more resolution and stability domain are accessible.

Utilizing the fabricated circuit and in a comparison with the similar circuits the diagram in Fig. 10 is resulted. In this figure the diagram (a) is the input signal for the AGC circuit. Maximum voltage is 200 mv. The frequency is approximately 40 Hz and the maximum domain changes slowly.

Diagram (b) is the output of an AGC circuit testing sample that is shown for better comparison. As it can be seen, as a result of large input voltage changes, the output could not eliminate input voltage changes.

So, the output involves voltage changes. Diagram (d) is the output of the AGC circuit shown in Fig. 9. It can be observed that despite the large changes in the input, the output is stable. Diagram (c) shows the binary number applied to the Mosfets. In this figure, due to the increase in the peak input voltage, the binary number applied to the Mosfets increases and gain reduction is the result. Vice versa, by a decrease in peak voltage, the binary number applied to the Mosfets decreases and the gain will be increased. As a result, the output will remain stable. This circuit is used as a practical. Figure 6 shows this circuit (Fig. 10).

8 Conclusion

AGC is one of the most important parts of a Sonar receiver. In this paper design of an AGC circuit based on a novel topology for use in a Sonar receiver system was discussed. In such topology, the amplitude is set to have a process over the sensor signal shape automatically. Nevertheless, the amplitude of the initial wave shape is accessible in the DSP processing system. This procedure can be done via multiplying the resulted number from the AGC circuit to the amplitude.

In this paper, the design levels of the AGC circuit are studied. This circuit is first simulated theoretically and then fabricated.

References

1. Akyildiz IF, Pompili D, Melodia T (2005) Underwater acoustic sensor networks: research challenges. Elsevier, Feb
2. Partan J, Kurose J, Levine BN (2006) A survey of practical issues in underwater networks. IEEE
3. Heidemann J et al (2006) Research challenges and applications for underwater sensor networking. IEEE
4. Cui JH et al (2006) The challenges of building scalable mobile underwater wireless sensor networks for aquatic applications. IEEE Network
5. Kilfoyle DB, Baggeroer AB (2000) The state of the art in underwater acoustic telemetry. IEEE J Oceanic Eng OE- 25(5):4–27
6. Smith JR (1998) Modern communication circuits. In: McGraw Hill electrical and computer engineering series, 2nd edn. New York
7. Rohde UL, Bucher TTN (1988) Communication receivers: principles and design. McGraw Hill, New York
8. Wenzhao W, Yaqin C, Qi Z (2004) Implementation of mixed feedback feedforward analog and digital AGC. In: International conference on microwave and millimeter wave technology proceedings
9. Walker S (1994) A high speed feed forward pseudo automatic gain control circuit for an amplifier cascade. CH3389-4/94/$\sim$ -0941$01. (XIO1994 Leee)
10. Isaac Martinez G (1997) Automatic gain control (AGC) circuits theory and design. ECE1352 Analog Integrated Circuits. University of Toronto
11. Schilling DL (1989) Electronic circuits: discrete and integrated. McGraw Hill, New York

Control of Robot Manipulators with a Model for Backlash Nonlinearity in Gears

Soheil Ahangarian Abhari, Farzad Hashemzadeh, Mehdi Baradaran-nia and Hamed Kharrati

Abstract This paper presents a model for backlash nonlinearity in gears based on torque input-output equation. By combining the robot dynamic model with backlash, a stable sliding mode controller is developed and the asymptotic stability of the closed-loop system is shown using Lyapunov method and Barbalat's Lemma. The proposed method is produced no chattering in the control torques and the tracking performance is desirable. Effectiveness of the controller is verified by comparative studies with numerical simulation. Finally, experimental results are presented to demonstrate the efficiency and capability of the proposed controller in dealing with backlash nonlinearities in gears of a five-bar manipulator.

Keywords Nonlinear control · Robot manipulator · Modeling
Backlash nonlinearities

1 Introduction

Industrial robot manipulators are dynamically coupled multi-axis electro-mechanical manufacturing machines that have been widely used in industrial automation. Each axis of robot has nonlinear characteristic due to transmission components of joints such as gear reducers, belts, or shafts. In engineering applications, an important nonlinearity that complicates control systems performance is

S. A. Abhari (✉) · F. Hashemzadeh · M. Baradaran-nia · H. Kharrati
Department of Control Engineering, Faculty of Electrical and Computer Engineering,
University of Tabriz, Tabriz, Iran
e-mail: sahangarian@tabrizu.ac.ir

F. Hashemzadeh
e-mail: hashemzadeh@tabrizu.ac.ir

M. Baradaran-nia
e-mail: mbaradaran@tabrizu.ac.ir

H. Kharrati
e-mail: kharrati@tabrizu.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_18

backlash. These faults can be diverse and, depending on the situation, can affect the performance of the system considerably. Consequently, it results in the degradation of the manipulator performance, reduces the positioning accuracy and even may lead to instability. From the control point of view, there exist many models of backlash compensation [1–4]. The spacing between gears results in a hysteresis type function that relates the input and the output angles. The effect of nonlinear friction and backlash of harmonic drive transmissions on robotic systems performance was studied in [5]. Ruderman et al. [6–10] described a dynamical model in which the hysteresis was related to the torsional displacement and velocity. Acho and some authors presented an application of the dead-zone model of backlash [11]. In [12–14], the authors have built a combination of flexible bodies and used a detailed simulation to develop and test the effectiveness of the proposed detection approach and correlated them to the parameters of the nonlinear system. The effect of backlash in electric actuator transmission system is evidencing the goodness of this model. Other authors proposed some control techniques to reduce effects of backlash with hysteresis model. It is important to find a model describing the nonlinear behavior and to utilize this model for controller design. Various models have been proposed to describe the hysteresis [15]. In several papers, authors presented a common feature of those schemes that rely on the construction of an inverse hysteresis to reduce the effects of the hysteresis [16]. If backlash spacing in gears is not accounted for, it will result in the degradation of the system performance, reduce positioning accuracy and even may lead to instability. The widely used methods include computed torque control, robust control, and optimal control [17–21]. For stabilizing a closed-loop system that contains a backlash, some papers propose an adaptive control law [22, 23]. A large number of publications have focused on the adaptive dead-zone inverse strategy for the systems with backlash. In such methods, the bounded output tracking error has been successfully achieved [22–24]. One of the main drawbacks of this technique is its necessity to rely on an inverse model to compensate the backlash nonlinearity. This technique depends on the model and is very complicated [11]. Su, Oya and Chen proposed a new approach control synthesis using the properties of the hysteresis model. To deal with the nonlinearities in the plant, a universal function approximator was adopted for compensation of plant nonlinearities, where parameters of the approximator were adapted using a Lyapunov-based design [25]. Therefore fast and accurate control of a robot is a challenging task and a nonlinear robust scheme is necessary to be used in the robot manipulator controller [26–28]. Chang and Yen in [29] designed a reduced-order observer to estimate the velocity signals and then proposed an observer-based robust position tracking controller without velocity measurements, and developed this method using the back-stepping technique. He et al. have discussed the input and output constraints and dead-zone model in the inputs and have proposed a robust adaptive vibration control scheme [30, 31]. Wang in [32] have presented a traditional robust adaptive control design technique which can be applied to achieve the bounded tracking errors without explicitly exploring the detailed dead-zone characteristics rather than the fact that the dead-zone effect can always be treated as a bounded input disturbance.

In this paper we proposed an approach for robust controller by using for a new model of the backlash with the properties of the gears. To improve the backlash effect in the dynamic of robot manipulator, a robust controller is designed, where parameters of the controller are adapted using a Lyapunov-based method. The proposed controller significantly improves the tracking performance and also provides less control effort with no chattering effect in input torques. The simulation results which are obtained by a two axis, five-bar parallel robot manipulator are given to illustrate the effectiveness of the proposed method. Finally, experimental results are illustrated to show the effectiveness of Proposed adaptive robust controller, with a real-time five-bar motion control setup.

2 Problem Statement

In this section, the robot manipulator dynamic model with the backlash nonlinearity is discussed. First we consider the backlash model and its properties:

2.1 Backlash Model and Its Properties

Consider model of a mechanical system with backlash. This model combines with a robot manipulator dynamic model.

Traditionally, backlash nonlinearity can be described by [15, 22, 33],

$$q(t) = \begin{cases} p(t) - b & \textit{if} \quad \dot{p}(t) > 0 \quad \textit{and} \quad p(t) \geq q(t) + b \\ p(t) + b & \textit{if} \quad \dot{p}(t) < 0 \quad \textit{and} \quad p(t) \leq q(t) - b \\ q(t^{-}) & \textit{otherwise} \end{cases} \tag{1}$$

where $q(t)$ and $p(t)$ are output and input angles respectively, and $\dot{p}(t)$ is input angular velocity of a gearbox with backlash. The parameter $b > 0$ is the backlash distance (Fig. 1). When backlash is presented in the joint gearbox, output angle follows input angle with a constant distance. If there is no contact in gears, output angle $q(t)$ will be between $p(t) - b$ and $p(t) + b$. In the proposed model the mentioned problem is avoided by the proposed adaptive robust controller. In almost all researches, input and output angles of the gears have been used in backlash modeling. However, in the proposed model, the input and output torques of the gear are used to simplify the overall model and the design procedure. The interval of the gear is estimated using two external encoders which are assembled on both sides of the gearbox. Further details on the experimental setup, are given in part 5. For further details on the experimental setup, we refer to experimental validations section.

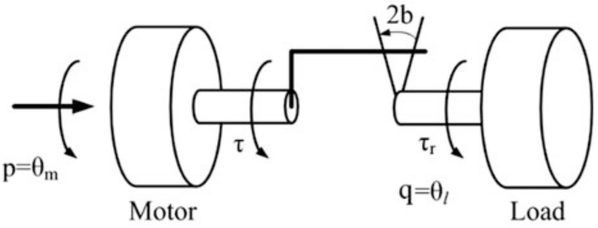


Fig. 1 A schematic diagram of the mechanism with backlash

$$\tau_r(t) = \begin{cases} \tau(t) & if \quad \dot{p}(t) > 0 \quad and \quad p(t) \geq q(t) + b \\ \tau(t) & if \quad \dot{p}(t) < 0 \quad and \quad p(t) \leq q(t) - b \\ 0 & otherwise \end{cases} \tag{2}$$

Here $\tau_r(t)$ and $\tau(t)$ are the output and input torque of each joint gearbox of the robot, respectively. Now, a discontinuous function is defined for the backlash model as follows:

$$\begin{gathered} f(q,p,\dot{p}) = \begin{cases} 1 & if \quad \dot{p}(t) > 0 \quad and \quad p(t) \geq q(t) + b \\ 1 & if \quad \dot{p}(t) < 0 \quad and \quad p(t) \leq q(t) - b \\ 0 & otherwise \end{cases} \\ \tau_r(t) = f(q(t),p(t),\dot{p}(t))\tau(t) \end{gathered} \tag{3}$$

where *f* is a function of input and output angles. Another discontinuous function for the backlash model has been defined which is used to show the relation between the input torque and the output torque as:

$$\begin{gathered} h(q,p,\dot{p}) = \begin{cases} 0 & if \quad \dot{p}(t) > 0 \quad and \quad p(t) \geq q(t) + b \\ 0 & if \quad \dot{p}(t) < 0 \quad and \quad p(t) \leq q(t) - b \\ 1 & otherwise \end{cases} \\ h(q(t),p(t),\dot{p}(t)) = 1 - f(q(t),p(t),\dot{p}(t)) \end{gathered} \tag{4}$$

Here *h* is a function of input and output angles similar to *f*. Their difference is that when *f* is one, *h* is zero and vice versa.

$$\tau_r(t) = \tau(t) - h(q(t),p(t),\dot{p}(t))\tau(t) \tag{5}$$

Let us define some functions to describe *h*.

$$
\begin{aligned}
u(t) &= \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases} \\
v(t) &= \begin{cases} 0 & t > 0 \\ 1 & t \leq 0 \end{cases}
\end{aligned}
\tag{6}
$$

It is easy to find that $h_i(q_i(t), p_i(t), \dot{p}_i(t))$ can be described for each axis i using functions u(t) and v(t) as follows.

$$
\begin{aligned}
h_i(q_i(t), p_i(t), \dot{p}_i(t)) = v((u(\dot{p}_i(t)) * u(p_i(t) - q_i(t) - b_i)) \\
+ (u(-\dot{p}_i(t)) * u(-p_i(t) + q_i(t) - b_i)))
\end{aligned}
\tag{7}
$$

For independent gearboxes of robot manipulator joints, the input and output torque, and input and output angle vectors are defined as follows:

$$
\begin{aligned}
\tau_r &= \begin{bmatrix} \tau_{r1} & \tau_{r2} & \ldots & \tau_{rn} \end{bmatrix}^T \\
\tau &= \begin{bmatrix} \tau_1 & \tau_2 & \ldots & \tau_n \end{bmatrix}^T \\
q &= \begin{bmatrix} q_1 & q_2 & \ldots & q_n \end{bmatrix}^T \\
p &= \begin{bmatrix} p_1 & p_2 & \ldots & p_n \end{bmatrix}^T
\end{aligned}
\tag{8}
$$

where τ_r is the input torques vector of the joints of the robot and the output torques vector of gearboxes. τ is the input torques vector of the gearboxes and is the output torques vector of the actuators. The parameter p is a vector of input angles of joints and q is a vector of output angles. Finally, for input and output torques of the robot actuators, a backlash model can be defined as:

$$
\begin{aligned}
H(q, p, \dot{p}) &= diag(\, h_1(q_1, p_1, \dot{p}_1) \quad \ldots \quad h_n(q_n, p_n, \dot{p}_n)\,) \\
\tau_r(t) &= \tau(t) - H(q(t), p(t), \dot{p}(t))\tau(t)
\end{aligned}
\tag{9}
$$

where $H(q(t), p(t), \dot{p}(t))$ is a diagonal matrix with zero or one elements.

2.2 Robot Manipulator Dynamic Model

The basic dynamic model of the serial robot manipulator can be expressed by following second order nonlinear vector differential equations defined in the joint space of the robot [34],

$$
B(q(t))\ddot{q}(t) + C(q(t), \dot{q}(t))\dot{q}(t) + f_v(t)\dot{q}(t) + g(q(t)) = \tau_r \tag{10}
$$

where $q(t)$, $\dot{q}(t)$ and $\ddot{q}(t) \in R^{n \times 1}$ are the joint angular positions, velocities and accelerations of the robot manipulator, respectively. $B(q(t))$ is the inertia matrix,

$C(q(t), \dot{q}(t))$ is Coriolis and centrifugal force matrix, $f_v(t)$ is the friction matrix, $g(q(t))$ is the gravity torque vector and τ_r is the motor output torque.

We shall propose a Lyapunov based control technique for plants of the robot manipulator (10), preceded by the backlash model described in (9). The proposed controller and backlash model diagram is shown in Fig. 2. Using the solution expression of backlash model (9), the robot dynamic model (10) changes to:

$$\begin{aligned} &B(q(t))\ddot{q}(t) + C(q(t), \dot{q}(t))\dot{q}(t) + f_v(t)\dot{q}(t) + g(q(t)) \\ &= \tau(t) - H(q(t), p(t), \dot{p}(t))\tau(t) \end{aligned} \tag{11}$$

3 Proposed Control Law

In presenting the Lyapunov based control structure, the following definition is required:

$$\begin{gathered} q = [q_1 \quad q_2 \quad \ldots \quad q_n]^T \\ q_d = [q_{d1} \quad q_{d2} \quad \ldots \quad q_{dn}]^T \\ \tilde{q} = q - q_d \end{gathered} \tag{12}$$

where q is the angular position vector of joints, q_d is the desired angular position vector and $\tilde{q}$ represents the tracking error vector. The control object is to design a control law for $\tau(t)$ in (9) to force the robot angular positions of joints q to follow a specified desired trajectory q_d.

Definition 1 A filtered tracking error is defined as [23]:

$$\begin{gathered} S(t) = \lambda\tilde{q}(t) + \dot{\tilde{q}}(t) \\ \lambda = diag[\lambda_1 \quad \lambda_2 \quad \ldots \quad \lambda_n] \\ \lambda_i > 0 \qquad i = 1, 2, \ldots, n \end{gathered} \tag{13}$$

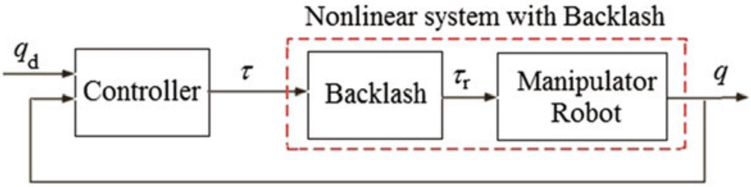


Fig. 2 Dynamic model of robot with backlash

Definition 2 A define a continuous function of filtered tracking error $S(t)$,

$$Sat(S) = \begin{cases} 1 - \exp\left(-\frac{S}{\gamma}\right) & S > 0 \\ -1 + \exp\left(\frac{S}{\gamma}\right) & S \leq 0 \end{cases} \tag{14}$$

where γ is a small positive constant. As $\gamma \to 0$, $Sat(S)$ approaches a step transition from -1 at $S = 0^{-}$ to 1 at $S = 0^{+}$ continuously.

Theorem 1 *The proposed controller is:*

$$\begin{aligned} \tau^{*}(t) = {} & B(q(t))(-KS(t) - k_d Sat(S) + \ddot{q}_d(t) - \lambda\dot{\tilde{q}}(t)) \\ & + (C(q(t), \dot{q}(t))\dot{q}(t) + f_v(t)\dot{q}(t) + g(q(t))) \end{aligned} \tag{15}$$

where K is a constant positive definite matrix and k_d is a positive scalar. The equation $S(t) = 0$ defines a time-varying hyperplane in R^n on which the tracking error vector $\tilde{q}(t)$ decays exponentially to zero.

Applying the nonlinear controller (15) in the dynamic model of the robot with backlash (11) the closed-loop is stable in the sense of Lyapunov and the position error $\tilde{q}(t) = q(t) - q_d(t)$ is bounded for constant backlash distance $b_i > 0$ in the all of joints.

Proof of Theorem 1 The derivative of the filtered tracking error vector (13) can be written as:

$$\dot{S}(t) = \lambda\dot{\tilde{q}}(t) + \ddot{\tilde{q}}(t) = \lambda\dot{\tilde{q}}(t) + (\ddot{q}(t) - \ddot{q}_d(t)) = -\ddot{q}_d(t) + \lambda\dot{\tilde{q}}(t) + \ddot{q}(t) \tag{16}$$

and Eq. (11) becomes,

$$\begin{aligned} \ddot{q}(t) = {} & B(q(t))^{-1}(-C(q(t), \dot{q}(t))\dot{q}(t) - f_v(t)\dot{q}(t) - g(q(t)) + \tau^{*}(t) \\ & - H(q(t), p(t), \dot{p}(t))\tau^{*}(t)) \end{aligned} \tag{17}$$

Using the simplified robot dynamic model (17), the Eq. (16) becomes,

$$\begin{aligned} \dot{S}(t) = {} & -\ddot{q}_d(t) + \lambda\dot{\tilde{q}}(t) \\ & + B(q(t))^{-1}(-C(q(t), \dot{q}(t))\dot{q}(t) - f_v(t)\dot{q}(t) - g(q(t)) + \tau^{*}(t)) \\ & - B(q(t))^{-1}(H(q(t), p(t), \dot{p}(t))\tau^{*}(t)) \end{aligned} \tag{18}$$

By applying the proposed controller (15), we have,

$$\begin{aligned}\dot{S}(t) = {} & -\ddot{q}_d(t) + \lambda\dot{\tilde{q}}(t)\\ & + B(q(t))^{-1}(-C(q(t),\dot{q}(t))\dot{q}(t) - f_v(t)\dot{q}(t) - g(q(t))\\ & + B(q(t))(-KS(t) - k_d Sat(S) + \ddot{q}_d(t) - \lambda\dot{\tilde{q}}(t))\\ & + (C(q(t),\dot{q}(t))\dot{q}(t) + f_v(t)\dot{q}(t) + g(q(t))))\\ & - B(q(t))^{-1}(H(q(t),p(t),\dot{p}(t))\tau^*(t))\end{aligned} \tag{19}$$

Equation (19) can be supposed as:

$$\dot{S}(t) = -KS(t) - k_d Sat(S) - B(q(t))^{-1}(H(q(t),p(t),\dot{p}(t))\tau^*(t)) \tag{20}$$

Define a Lyapunov candidate function:

$$V(t) = \frac{1}{2}S(t)^T S(t) \tag{21}$$

The time derivative of $V(t)$ is:

$$\dot{V}(t) = S(t)^T \dot{S}(t) \tag{22}$$

By applying $\dot{S}(t)$ in (20)–(22), we have:

$$\dot{V}(t) = S(t)^T(-KS(t) - k_d Sat(S) - B(q(t))^{-1}(H(q(t),p(t),\dot{p}(t))\tau^*(t))) \tag{23}$$

Suppose that K_2 is a positive definite matrix such that:

$$K_2 S = k_d Sat(S) + H(q(t),p(t),\dot{p}(t))B(q(t))^{-1}\tau^*(t) \tag{24}$$

Note that H is a diagonal matrix and $B^{-1}H$ is equal to $H^{-1}B$. By applying control low (15) in Eq. (24):

$$\begin{aligned}k_d = {} & (F(q(t),p(t),\dot{p}(t))Sat(S))^T(K_2 S + H(q(t),p(t),\dot{p}(t))(KS(t) - \ddot{q}_d(t) + \lambda\dot{\tilde{q}}(t)\\ & - B(q(t))^{-1}(C(q(t),\dot{q}(t))\dot{q}(t) + f_v(t)\dot{q}(t) + g(q(t))))))/\|F(q(t),p(t),\dot{p}(t))Sat(S)\|^2\end{aligned} \tag{25}$$

Note that when the norm of $F(q(t),p(t),\dot{p}(t))Sat(S)$ is zero, then k_d should be set to zero. Using above computation for k_d can be written as:

$$\dot{V}(t) = -S(t)^T KS(t) - S(t)^T K_2 S(t) = -S(t)^T(K + K_2)S(t) \tag{26}$$

in which K and K_2 are positive definite matrices. We can infer that V is a Lyapunov function which leads to global boundedness of $S(t)$ and based on the theory of Lyapunov stability, it is proved that the closed-loop system is stable in the sense of Lyapunov. Then, the proof is completed.

To complete the proof and establish asymptotic convergence of the tracking error, it is necessary to show that $S(t) \to 0$ as $t \to \infty$. This is accomplished by applying Barbalat's Lemma [33]:

Theorem 2 *In the closed-loop system* (11) *with controller* (15), *the absolute values of the position error* $\tilde{q}(t) = q(t) - q_d(t)$ *tend to zero asymptotically* (*i.e.,* $\tilde{q}(t) \to 0$) *if all conditions in Theorem 1 are satisfied and* $\tau(t)$ *is bounded for constant backlash distance* $\mathrm{b}_i > 0$.

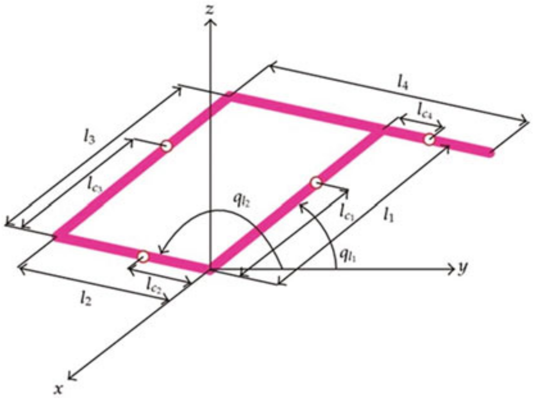
Proof of Theorem 2 Integrating both sides of (26), and based on the result of Theorem 1, where it is shown that $V(t)$ is a lower bounded decreasing function, it is easy to see that, $S(t) \in \ell_2$ and bounded. S(t) and $\tilde{q}(t)$ have a BIBO dynamic system (13) so $\tilde{q}(t)$ and $\dot{\tilde{q}}(t)$ are bounded signals. Using (19) it is possible to see that $\dot{S}(t)$ is bounded ($\dot{S}(t) \in \ell_\infty$). Since $S(t) \in \ell_2$ and $\dot{S}(t) \in \ell_\infty$, using Barbalat's lemma it can be found that $S(t) \to 0$. Applying $S \in 0$ in (13) indicates that $\tilde{q}(t) \to 0$ and $q(t) \to q_d(t)$ as $t \to \infty$. So the position tracking error converge to zero asymptotically. The proof is completed.

4 Simulation Results

Some simulation results are presented in this section to evaluate the performance of the proposed controller. In order to illustrate the effectiveness of the proposed method, an example of a two-axis, five-bar parallel robot manipulator is considered.

Figure 3 shows the schematic of five-bar parallel robot manipulator. The dynamic model of the five-bar manipulator can be expressed by second order nonlinear differential equations defined in the joint space of the robot as follows:

Fig. 3 Schematic of five-bar parallel robot manipulator



$$\begin{gathered}
(M_{11}+I_h^1)\ddot{q}_1+M_{12}\ddot{q}_2+\tfrac{\partial M_{12}}{\partial q_2}\dot{q}_2^2+g(m_1\ell_{c1}+m_3\ell_{c3}+m_4\ell_1)\cos(q_1)=\tau_1\\
(M_{22}+I_h^2)\ddot{q}_2+M_{12}\ddot{q}_1+\tfrac{\partial M_{12}}{\partial q_1}\dot{q}_1^2+g(m_2\ell_{c2}+m_3\ell_2+m_4\ell_{c4})\cos(q_2)=\tau_2\\
M_{11}=I_{11}^1+I_{11}^3+m_1\ell_{c1}^2+m_3\ell_{c3}^2+m_4\ell_1^2\\
M_{22}=I_{11}^2+I_{11}^4+m_2\ell_{c2}^2+m_3\ell_2^2+m_4\ell_{c4}^2\\
M_{12}=M_{21}=(m_3\ell_{c2}\ell_2-m_4\ell_{c4}\ell_1)\cos(q_1-q_2)
\end{gathered} \tag{27}$$

where $q(t)$, $\dot{q}(t)$ and $\ddot{q}(t)\in R^{2\times 1}$ are the joint angular positions, velocities and accelerations of the five-bar manipulator, respectively. In addition I_h is the motor inertia and I_{11} is the link inertia, ℓ is the length of the link and ℓ_c is the length of the center of link, m is the mass of the link, g is the gravity and τ is the motor output torque. The nominal values of the five-bar robot parameters are selected from [35]. The parameters of the dynamic model of the five-bar manipulator are:

$$\begin{aligned}
&I_{11}^1=1, I_{11}^2=2, I_{11}^3=1, I_{11}^4=2, I_h^1=1, I_h^2=1.5, m_1=1.2190,\\
&m_2=0.5534, m_3=1.2190, m_4=1.2771,\\
&\ell_{c1}=0.015, \ell_{c2}=0.00093, \ell_{c3}=0.012, \ell_{c4}=0.018, \ell_1=0.33,\\
&\ell_2=0.12, \ell_3=0.33, \ell_4=0.45, g=9.8
\end{aligned}$$

To show the efficiency of the proposed controller, the backlash is applied in all joints of the robot manipulator. The backlash parameters in joints are set to $b_1=0.1\mathrm{rad}$ and $b_2=0.1\mathrm{rad}$. There are many issues associated with the path planning problem, such as avoiding obstacles and making sure that the planned path does not require the voltage and torque limitations of the actuators. To show the performance of the proposed controller, the desired reference trajectory is chosen as a trapezoidal pulse wave and a sinusoidal function for each joint of the robot manipulator. The sampling time of the simulation is 10 ms and the simulation is executed for two periods of desired reference with zero initial states for joint positions and velocities. The parameter of adaptation is $K_h=1*I_2$ and the parameters of the controller are chosen to be $K=10*I_2$, $K_2=1*I_2$. The parameter λ is set to 1 and γ is set to 0.5.

Figure 4 demonstrates the desired reference trajectory and the tracking results of the joint positions. The tracking errors for the desired trajectory is shown in Fig. 5.

Figure 6 shows the desired sinusoidal reference trajectory and the tracking results of joint positions and Fig. 7 shows the tracking errors for the desired trajectory. It can be easily seen that using the proposed method, robot manipulator tracks the desired trajectory efficiently.

To make a clearer comparison, the results of the proposed method are compared with a PDFF controller (PD with feedforward inverse dynamics of the hysteresis) proposed in [36].

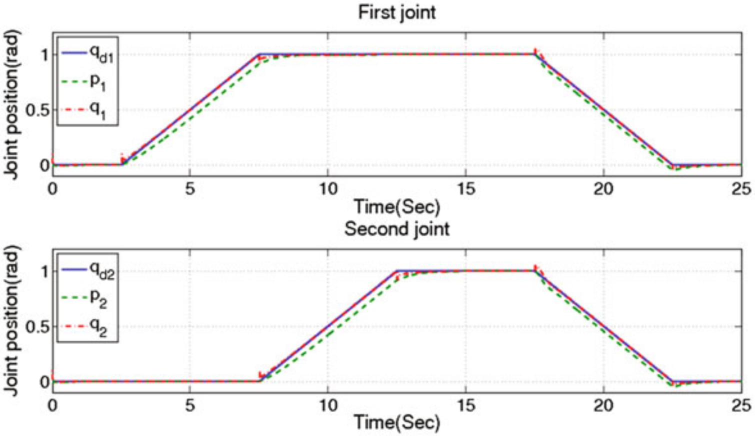


Fig. 4 Joint positions of the robot with desired trapezoidal pulse

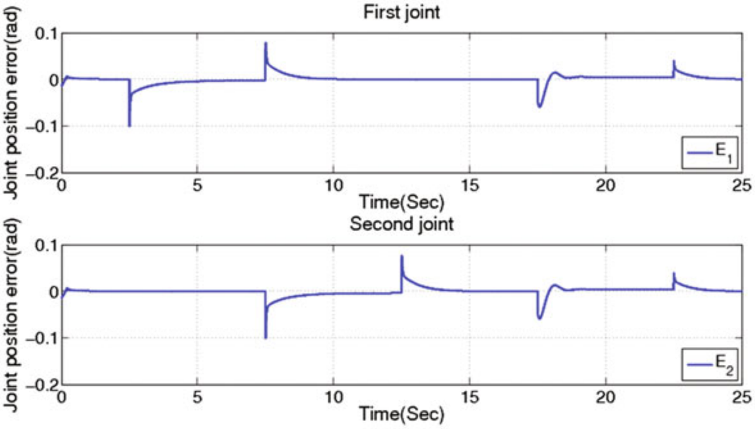


Fig. 5 Joint position errors (trapezoidal pulse)

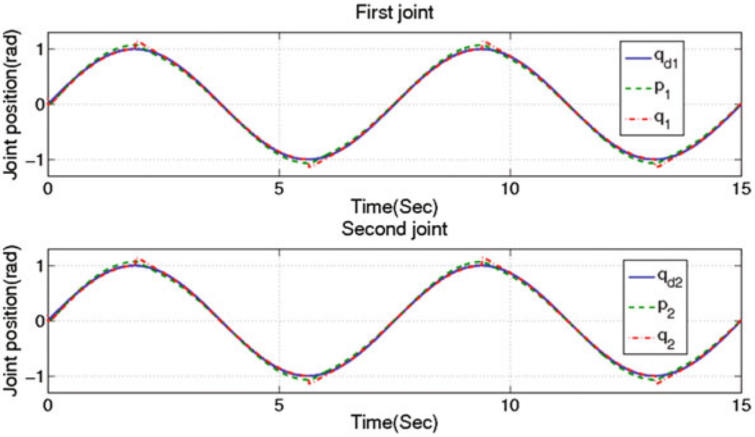


Fig. 6 Joint positions of the robot with desired sinusoidal function

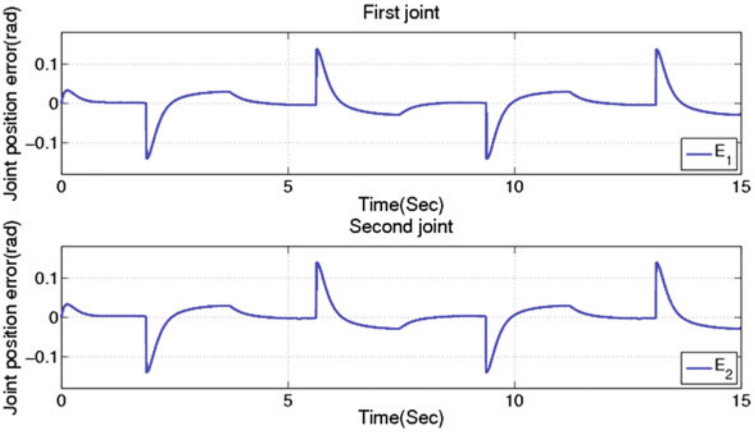


Fig. 7 Joint position errors (sinusoidal function)

Figure 8 demonstrates the tracking errors for the desired trajectory of the joint positions when the proposed controller and the other controller are applied. Considering Fig. 8, it can be easily seen that the proposed controller can effectively handle the backlash effects in the joint gears when it is compered with other method.

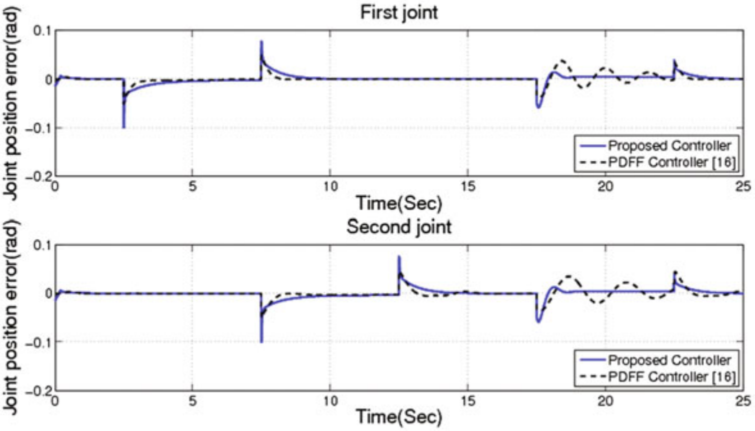


Fig. 8 Joint position errors (sinusoidal function)

5 Experimental Validations

The validity of the proposed algorithm is demonstrated in real-time on a five-bar motion control setup. Figure 9 shows the five-bar manipulator, which has been built in robotics research lab in University of Tabriz. The five-bar manipulator consists of DC motors and high resolution encoders for both joints of the robot. The controller unit contains a real-time national instruments PCI-6601 data acquisition (DAQ) card. This card contains the digital input/output signal ports and four 32-bit fast counters/timers. Two channels of the counters/timers are used for counting high resolution encoders for output angle feedback and the other two channels are used for Pulse Width Modulation (PWM) signals to drive actuators for applying the control law. The backlash interval of the gear is estimated using two external encoders which are assembled on both sides of the gearbox. The proposed controller has been implemented using real time toolbox of the Matlab with one millisecond sampling time.

To show the performance of the developed control approach in experimental results, the desired reference trajectory is chosen to be a trapezoidal pulse wave with constant amplitude and frequency for each joint axis of the five-bar manipulator (a video file from experimental result is attached with paper submission). The sampling time of the real-time motion control setup is 10 ms and executed for one period of the desired reference with zero initial states for each joint position and velocity. The parameters of the proposed controller are chosen as the same as the simulation section to be $K = 10 * I_2$, $K_2 = 1 * I_2$. The parameter λ is set to 1 and γ is set to 0.5.

In Figs. 10, 11 and 12, $q_{d1}(t)$ and $q_{d2}(t)$ are the desired trajectory and $q_1(t)$ and $q_2(t)$ are the joint positions and E_1, E_2 are the first and second joint position tracking errors. The desired reference trajectory and the tracking results of the joint

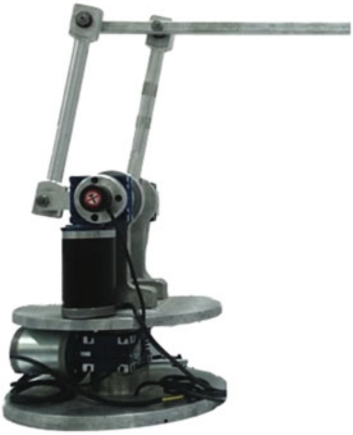


Fig. 9 Five-bar manipulator (built in robotics research lab in University of Tabriz)

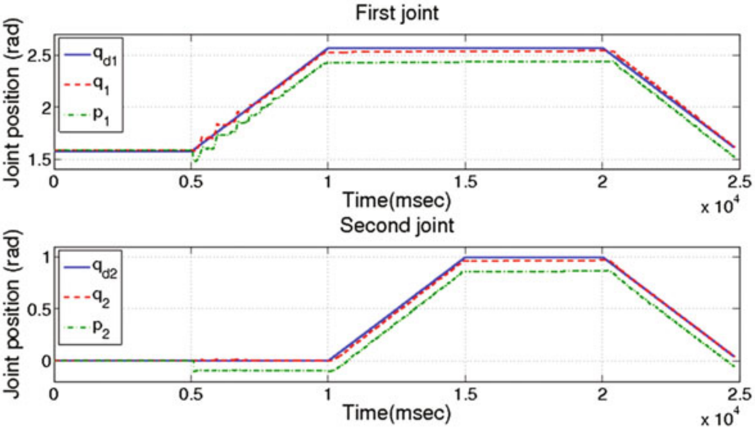


Fig. 10 Five first and second joint positions with desired trapezoidal pulse wave using proposed controller

angles of the five-bar manipulator are shown in Fig. 10. Figure 11 shows the tracking errors for the desired trajectory. To show the performance of the proposed controller, a PDFF controller like simulation controller applied to the robot. Figure 12 shows the tracking results of the desired trapezoidal pulse wave trajectory.

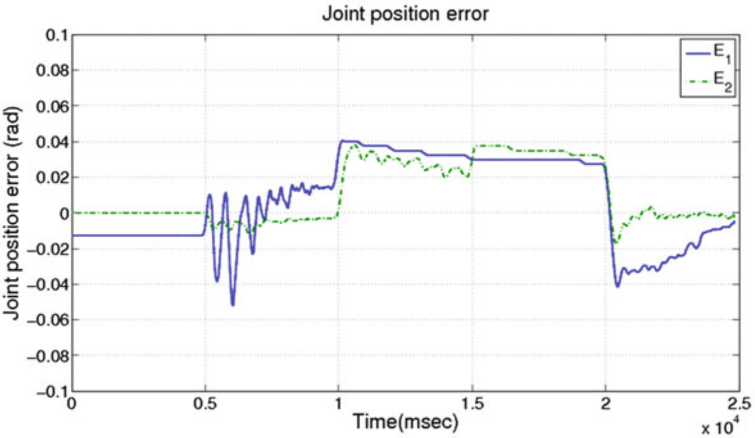


Fig. 11 First and second joint position errors using proposed controller

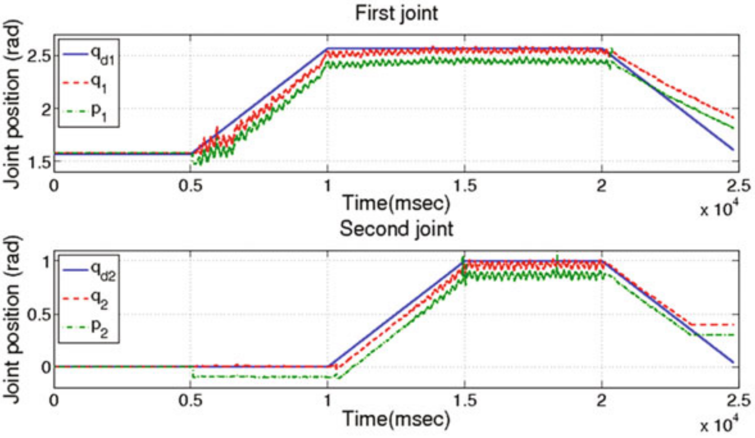


Fig. 12 First and second joint positions with desired trapezoidal pulse wave using PDFF controller

6 Conclusions

In this paper, a new modeling method is proposed for backlash nonlinearity in gears. Combining the backlash model with robot dynamic model, a stable sliding mode controller was developed and the stability of the closed-loop system was shown using the Lyapunov method which guarantees asymptotic stability of the nonlinear system. The effectiveness of the proposed method was confirmed through simulations for a six axis serial robot manipulator. The experiments conducted for a five-bar manipulator and it was successfully demonstrated that the proposed controller is capable of tracking the reference signals with backlash in gears of the robot manipulator.

References

1. Dhaouadi R, Ghorbel FH, Gandhi PS (2003) A new dynamic model of hysteresis in harmonic drives. IEEE Trans Industr Electron 50(6):1165–1171
2. Xiao Y, Du Z, You W, Li R (2010) Modeling and simulating the nonlinear characters of robot joints. In: 2010 IEEE international conference on robotics and biomimetics. IEEE, pp 914–919
3. Lagerberg A, Egardt B (2007) Backlash estimation with application to automotive powertrains. IEEE Trans Control Syst Technol 15(3):483–493
4. Villwock S, Pacas M (2009) Time-domain identification method for detecting mechanical backlash in electrical drives. IEEE Trans Industr Electron 56(2):568–573
5. Zhang H, Ahmad S, Liu G (2015) Modeling of torsional compliance and hysteresis behaviors in harmonic drives. IEEE/ASME Trans Mechatron 20(1):178–185
6. Ruderman M, Hoffmann F, Bertram T (2008) Preisach model of nonlinear transmission at low velocities in robot joints. In: 2008. AMC'08. 10th IEEE international workshop on advanced motion control. IEEE, pp 721–726
7. Ruderman M, Hoffmann F, Bertram T (2009) Modeling and identification of elastic robot joints with hysteresis and backlash. IEEE Trans Industr Electron 56(10):3840–3847
8. Ruderman M, Bertram T (2012) Modeling and observation of hysteresis lost motion in elastic robot joints. IFAC Proc Volumes 45(22):13–18
9. Ruderman M, Bertram T, Iwasaki M (2014) Modeling, observation, and control of hysteresis torsion in elastic robot joints. Mechatronics 24(5):407–415
10. Ruderman M, Iwasaki M (2014) On identification and sensorless control of nonlinear torsion in elastic robotic joints. In: IECON 2014-40th annual conference of the IEEE industrial electronics society. IEEE, pp 2828–2833
11. L. Acho, F. Ikhouane, and G. Pujo, "Robust control design for mechanisms with backlash," *Journal of Control Engineering and Technology*, vol. 3, no. 4, 2013
12. Trendafilova I, Van Brussel H (2001) Non-linear dynamics tools for the motion analysis and condition monitoring of robot joints. Mech Syst Signal Process 15(6):1141–1164
13. Tjahjowidodo T, Al-Bender F, Van Brussel H (2007) Experimental dynamic identification of backlash using skeleton methods. Mech Syst Signal Process 21(2):959–972
14. Dion J-L, Le Moyne S, Chevallier G, Sebbah H (2009) Gear impacts and idle gear noise: Experimental study and non-linear dynamic model. Mech Syst Signal Process 23(8):2608–2628
15. Su C-Y, Stepanenko Y, Svoboda J, Leung T-P (2000) Robust adaptive control of a class of nonlinear systems with unknown backlash-like hysteresis. IEEE Trans Autom Control 45 (12):2427–2432
16. Liu S, Su C-Y, Li Z (2014) Robust adaptive inverse control of a class of nonlinear systems with prandtl-ishlinskii hysteresis model. IEEE Trans Autom Control 59(8):2170–2175
17. Bi S, Deng M, Wang L, Ma S (2013) Operator-based robust control for nonlinear uncertain systems with backlash-like hysteresis. In: 2013 international conference on advanced mechatronic systems (ICAMechS). IEEE, pp 710–715
18. Soltanpour MR, Khalilpour J, Soltani M (2012) Robust nonlinear control of robot manipulator with uncertainties in kinematics, dynamics and actuator models. Int J Innovative Comput Inf Control 8(8):5487–5498
19. Lin F, Brandt RD (1998) An optimal control approach to robust control of robot manipulators. IEEE Trans Robot Autom 14(1):69–77

20. Dong R, Tan Y, Janschek K (2016) Nonsmooth predictive control for wiener systems with backlash-like hysteresis. IEEE/ASME Trans Mechatron 21(1):17–28
21. Tao G, Ma X, Ling Y (2001) Optimal and nonlinear decoupling control of systems with sandwiched backlash. Automatica 37(2):165–176
22. Guo J, Yao B, Chen Q, Jiang J (2009) Adaptive robust control for nonlinear system with input backlash or backlash-like hysteresis. In: 2009. ICCA 2009. IEEE international conference on control and automation. IEEE, pp 1962–1967
23. Hu C, Yao B, Wang Q (2011) Adaptive robust precision motion control of systems with unknown input dead-zones: a case study with comparative experiments. IEEE Trans Industr Electron 58(6):2454–2464
24. Li Y, Wang Q (2017) Adaptive robust tracking control of a proportional pressure-reducing valve with dead zone and hysteresis. Transactions of the Institute of Measurement and Control
25. Nordin M, Bodin P, Gutman P-O (2001) New models and identification methods for backlash and gear play. Adaptive control of nonsmooth dynamic systems, pp 1–30
26. Zheng M, Liao W, Yin C, Wang A (2014) Nonlinear tracking control design of a robot arm using robust right coprime factorization and sliding mode approaches. In: 2014 international conference on advanced mechatronic systems (ICAMechS). IEEE, pp 11–16
27. Khalate A, Dey R, Ray G et al (2014) Robust control of robot manipulator based on estimation of upper bounds on parametric uncertainty. In: 2014 international conference on electrical and computer engineering (ICECE). IEEE, pp 745–748
28. Bechlioulis CP, Liarokapis MV, Kyriakopoulos KJ (2014) Robust model free control of robotic manipulators with prescribed transient and steady state performance. In: 2014 IEEE/RSJ international conference on intelligent robots and systems (IROS 2014). IEEE, pp 41–46
29. Chang Y-C, Yen H-M (2011) Design of a robust position feedback tracking controller for flexible-joint robots. IET Control Theory Appl 5(2):351–363
30. He W, Ouyang Y, Hong J (2017) Vibration control of a flexible robotic manipulator in the presence of input deadzone. IEEE Trans Industr Inf 13(1):48–59
31. He X, He W, Sun C (2017) Robust adaptive vibration control for an uncertain flexible timoshenko robotic manipulator with input and output constraints. Int J Syst Sci 48(13):2860–2870
32. Wang X-S, Su C-Y, Hong H (2004) Robust adaptive control of a class of nonlinear systems with unknown dead-zone. Automatica 40(3):407–413
33. Su C-Y, Oya M, Chen X (2001) Robust adaptive control of nonlinear systems with dynamic backlash-like hysteresis. In: Adaptive control of nonsmooth dynamic systems. Springer, pp 273–288
34. Lewis FL, Dawson DM, Abdallah CT (2003) Robot manipulator control: theory and practice. CRC Press, 2003
35. Badamchizadeh MA, Hassanzadeh I, Abedinpour Fallah M (2010) Extended and unscented kalman filtering applied to a flexible-joint robot with jerk estimation. Discrete Dynamics Nat Soc 2010
36. Ruderman M, Iwasaki M (2016) Sensorless torsion control of elastic-joint robots with hysteresis and friction. IEEE Trans Industr Electron 63(3):1889–1899

Designing an Automatic and Self-adjusting Leg Prosthesis

Vahid Noei and Mehrdad Javadi

Abstract Throughout the whole human history, loss of limbs has been an important issue, such that extensive research has been conducted in designing and controlling above-knee prostheses, which have a long history of application by humans. Over time, with advances in medical sciences and engineering, this auxiliary tool evolved and has always been improving. Nevertheless, some challenges have remained. Therefore, proper and optimal design of knee mechanism for those who are not able to move can be very useful. In this research, first the human motion will be simulated and then prosthesis will be controlled. Received data from the rehabilitation center will be transferred through PLC to micro-prosthesis section. In the processing section, microcontroller will be calculated the information received and will be applied to the ankle and knee actuators. Using sensors embedded in different parts of the prosthesis, the prosthesis will be detected the actual position of the prosthesis and will be controlled the movement of the prosthesis.

Keywords Exoskeleton · Newton-Euler · 8-bar mechanism · Dynamics
Knee rehabilitation

1 Introduction

Jun et al. (2011) presented a smart and active multiaxial four bar mechanism, where a ball-screw causes motion of this artificial knee [1, 2]. They concluded that as use of a lower driving force in comparison with smart bionic leg above knee is required

V. Noei (✉)
Mechanical Group, Technical and Engineering Faculty,
Islamic Azad University, Tehran South Branch, Tehran, Iran
e-mail: st_v_noei@azad.ac.ir

M. Javadi
Department Mechanical Engineering, Islamic Azad
University Tehran South Branch, Tehran, Iran
e-mail: mjavadi@azad.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_19

for complete walking and one of the most advanced smart prostheses. A bionic leg should be able to support and control the body weight and accelerate the body mass, thereby helping walking criteria [3–5].

Primary studies regarding prosthesis of lower limbs for people suffering from amputation initiated since 1958, whereby a number of research groups in the US and former Yugoslavia were active in this area [6]. Chakraborty et al. (1994) presented a six-bar mechanism for an artificial knee, providing the possibility of sitting as cross-legged [7]. By installing LEDs to the thigh, above and below the knee, ankle joint, heel, and toes, and using a camera whose shutter had been kept open in a dark room, they registered the walking pathways of a person in a full cycle of walking. The results suggested a close relationship between experimental and analytical values. The position of the immediate center of the prosthetic knee considering ankle-pelvis reference line which has been plotted across different stages of walking indicates stability of prosthesis in the stage of stance and easy flexion for beginning of swinging. One of the important properties of this prosthesis is its multiplicity, allowing for adjusting it for the height of different amputated individuals.

The opinion of execution is very difficult and humans have only two legs. Therefore, for beginning a motion, the human should use one or both legs. Accordingly, to initiate the motion of this smart structure, engine torch force should be used [8].

Jin et al. (2003) used experimental and computer methods to investigate a six-bar mechanism in knee prosthesis. They concluded that six-bar mechanism, in comparison with four-bar mechanism, performs foot ankle joint path traversal better in the swing stage [9]. They also noted that the six-bar mechanism for greater stability in the stance stage can be designed in a way that it has more passive joints in comparison with the four-bar mechanism. Using compound error function optimization method, they determined it the design parameters and then through dynamic analysis, they determined the site of torque controller for the minimum level of input controlling torque. They also investigated the user on an inclined plane and indicated that under those conditions again the six-bar mechanism was superior to the four-bar mechanism designed with the same methods. Overall, in the study conducted by this group, it was found that the six-bar mechanism develops more real foot ankle joint path traverse and thigh motion, in comparison with the four-bar mechanism.

Sanseis et al. (2009) presented an optimization method for the performance of a four-bar mechanism for artificial knee prosthesis [10]. In that method, using the experimental motion of walking as well as the special conditions in necessities associated with the ability of the patient in controlling and stabilizing the prosthesis, they determined a four-bar mechanism, which had the greatest proportion with the natural state, yet it was in line with the conditions of the patient. Unlike its previous methods, the new method did not need a reference cent rode. This property allows for the possibility of determining a four bar mechanism, which can develop natural knee motions with a suitable and adequate accuracy. Al-Arcadian et al. (2011) presented mechanical design of human walking [11]. Using design rules, they obtained an eight-bar mechanism with one degree of freedom, such that the input member was a dual member, connected to the earth from two points. This mechanism indicates human leg motion. They used MATLAB software, and using a series of primary assumptions based on an adult human leg, they analyzed the mechanism. The design characteristic considered included height and thinness as well as manner of walking. Further, as with the human leg, the mechanism stimulus was considered in the upper part of the mechanism. They also used Working Model suffer for validation of their results and indicated that this mechanism can be used for the walking motion of humanoid bi-legged robots.

Xin et al. (2014) investigated a smart bionic leg, which is an advanced artificial member [12]. They used a four-bar knee mechanism and then using genetic algorithm, they optimized IBL. The results of this investigation indicated that the four-bar mechanism can resent characteristics similar to a real knee.

Hans et al. (2015) worked on controlling a bionic hand, where a complicated method has been designed for controlling it directly [13]. This bionic hand is able to return sensation through an electric interface with nerves.

Forego et al. (2017) designed a flexible robotic leg by imitating human instep, which is compatible across different areas [14].

2 Stages of the Experiment

The most common method for investigating human motion is installing a series of signs on the skin of different parts of the body. Most of the analytical techniques available consider human body members as a rigid body.

In this research, motion analysis and data acquisition were performed by marking method. Figure 1 demonstrates the cycle of walking for the studied person in the laboratory environment by nexus software.

3 Properties and Conditions of the Experiment

Motion analysis test was performed in the research center of smart neurological rehabilitation technologies Javad Movafaghian for 20 min [15]. The tested person had the following specifications (gender: male, weight: 83 kg, height: 182 cm, the number of used markets: 32).

The experiment was performed in three states of normal, fast, and slow, and the motion analysis data was received from the laboratory as an Excel file.

After receiving the data from the laboratory, these data were implemented in a Programmable Logic Controller (PLC). For this purpose, the PLC classified the extracted data as voltage (0–5 v) and sent to the main board Arduino Mega 2560. After this stage, the main board computes the angle of knee and ankle of the prosthesis at any moment considering the received voltage. Control of angles is as

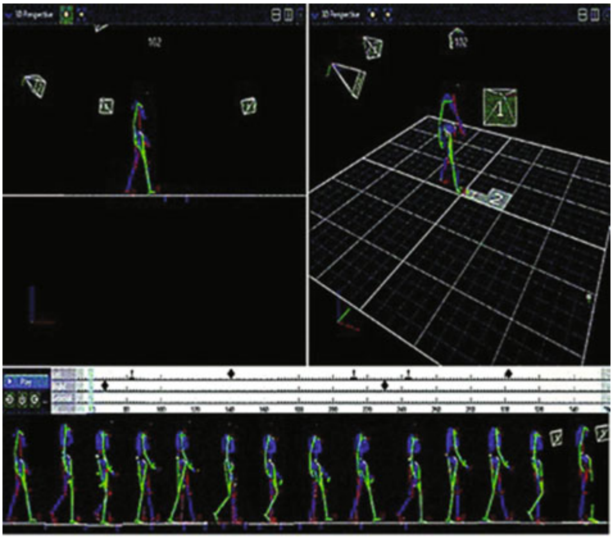


Fig. 1 The cycle of walking by the studied individual in Nexus software

closed-loop and is constantly devised by encoders connected to DC engines. In relation to the sent information, the main board in knee and ankle investigates the status of prosthesis toes. The flow chart of the prosthesis control of the accessories of the main board (Arduino Mega 2560) has been demonstrated in Fig. 2.

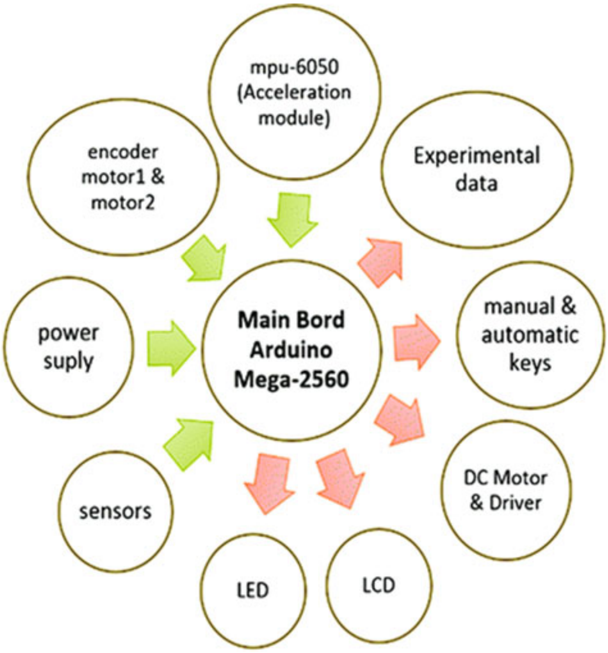


Fig. 2 Prosthesis control flowchart

The mechanism shape

Based on Fig. 3, leg prosthesis consists of the following:

1. DC engine and knee stimulator gearbox, 2. Micro switches constraining prosthesis shin motion, 3. 12-V battery 4. Encoder in DC engine of the knee joint. 5. Power switch and battery charger jack and lights notifying “on service” and “on charging”, 6. Key for motion in the manual state, 7. CLD, 8. Driver of DC engines, 9 Main board, 10. Power, 11. Encoder of DC engine of the prosthesis ankle joint, 12. Micro switches constraining prosthesis ankle motion, 13. DC engine and gearbox for stimulating ankle, 14. Accelerometer sensor, 15. Damper, 16. Sensors detecting the prosthesis bottom position.

In this project, for validation of the motion of knee and ankle joints of the prosthesis, the data received from the laboratory was applied to the prosthesis and the results were compared with the prosthesis output (Figs. 4, 5, 6 and 7). The fabricated prosthesis has been tested across the three stages (normal velocity, slow, fast) (Figs. 8 and 9).

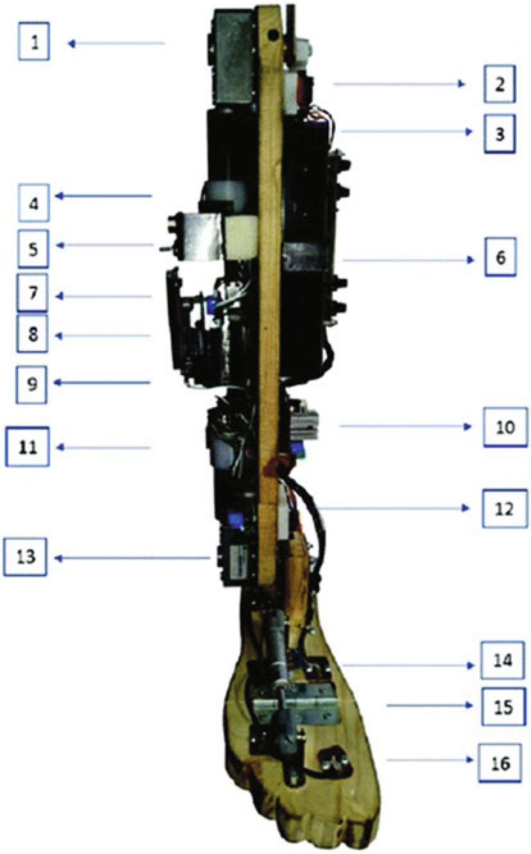


Fig. 3 The prosthesis constituents

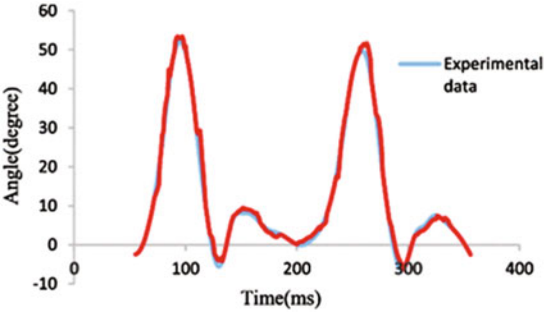


Fig. 4 Comparison of the output of knee joint with experimental data in normal gait

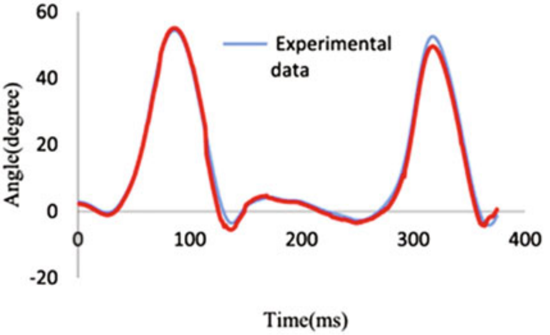


Fig. 5 Comparison of the output of knee joint with experimental data in slow gait

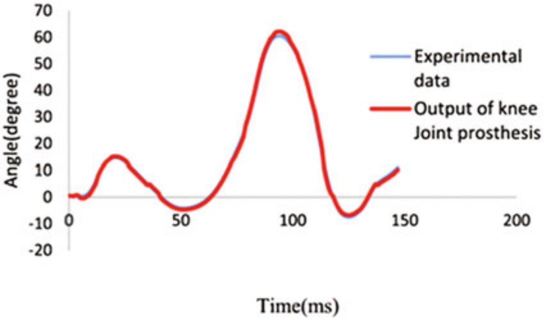


Fig. 6 Comparison of the output of knee joint with experimental data in fast gait

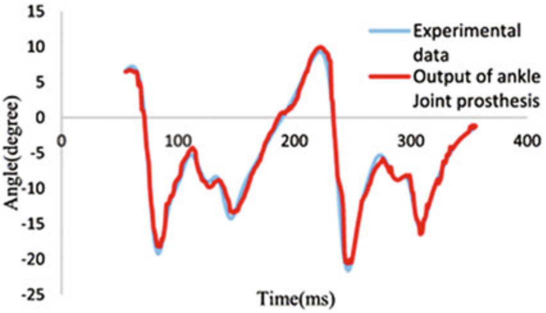


Fig. 7 Comparison of the output of ankle joint with experimental data in normal gait

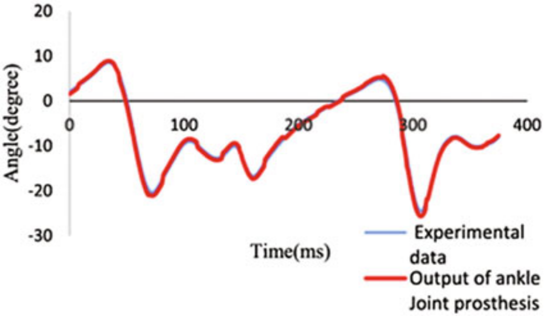


Fig. 8 Comparison of the output of ankle joint with experimental data in slow gait

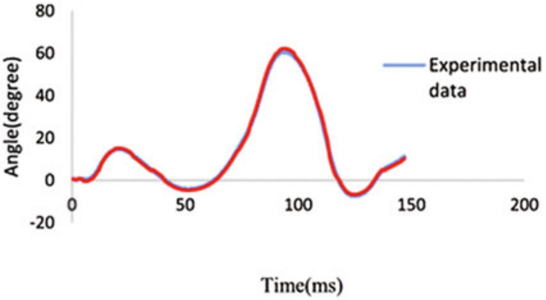


Fig. 9 Comparison of the output of ankle joint with experimental data in fast gait

4 Conclusion

In this study, the human motion was first simulated and then it was controlled by the prosthesis. Received data from the rehabilitation center using PLC to the microcontroller transferred. In the processing section, microcontroller was calculates the information received and applies to the ankle and knee actuators. Using sensors embedded in different parts of the prosthesis, the prosthesis detects the actual position of the prosthesis and controls the movement of the prosthesis.

References

1. Wang RC, Jin DW (2009) Research and development of bionic and intelligent limb prosthesis. Chin Med Devices Inf 15(1):3–5
2. Kim JH, Oh JH (2001) Development of an above knee prosthesis using MR damper. In: Proceedings of IEEE international conference on robotics & automation. IEEE, Seoul, Korea, pp 3686–3691
3. Donelan JM, Kram R, Kuo AD (2002) Mechanical work for step-to-step transitions is a major determinant of the metabolic cost of human walking. J Exp Biol 205:3717–3727
4. Grabowski A, Farley CT, Kram R (2005) Independent metabolic costs of supporting body weight and accelerating body mass during walking. J Appl Physiol 98:579–583
5. Kuo AD, Donelan J. M, Ruina A (2005) Energetic consequences of walking like an inverted pendulum: step-to-step transitions. Exerc Sport Sci Rev 33:88–97. (https://doi.org/10.1097/00003677-200504000-00006)
6. Dollar AM, Herr H (2008) Lower extremity exoskeletons and active orthoses: challenges and state-of-the-art. IEEE Trans Robot 24:144–158
7. Chakraborty J, Patil k (1994) A new modular six-bar linkage trans-femoral prosthesis for walking and squatting. Prosthet Orthot Int 18:98–108
8. Jun K, Chengdong W, Fei W, Shiguang W (2011) The research of the four-bar bionic active knee. Adv Mater Res 308–310:1988–1991
9. Jin D, Zhang R, Dimo H, Wang R, Zhang J (2003) J Rehabil Res Dev 40(1):39–48
10. Sancisi N, Caminati R, Parenti-Castelli V (2009) "Optimal four-bar linkage for the stability and the motion of the human knee prostheses", presented at the atti del xix congresso dell'associazione italiana di mechanics teorica e applicata, Ancona, pp 1–10
11. Al-Araidah O, Batayneh W, Darabseh T, Banihani S (2011) Conceptual design of a single DOF human-like eight-bar leg mechanism. Jordan J Mech Ind Eng 5(4):285–289
12. Xie H, Wang S, Li F (2014) Knee joint optimization design of intelligent bionic leg based on genetic algorithm. Int J BIO Autom 18(3):195–206
13. Hannes PS, Sliman JB (2015) Biomimetic approaches to bionic touch through a peripheral nerve interface. Int J Neuropsychologia 79:344–353, ISSN: 0028-3932
14. Diego F, Lukas M, Brittany P, Zhen X, Carlo M (2017) A case study of a force-myography controlled bionic hand mitigating limb position effect. J Bionic Eng 14:692–705
15. Javad Mowafaghian research center of intelligent neuro-rehabilitation technologies, Sharif University of technology, Tehran, Iran (1396)

Part III
Electronic Engineering

Implement Deep SARSA in Grid World with Changing Obstacles and Testing Against New Environment

Mohammad Hasan Olyaei, Hasan Jalali, Ali Olyaei and Amin Noori

Abstract In this paper, the Deep SARSA method is used to find the path of the robot on 5 × 5 environment with the presence of moving obstacles. This problem is known as Grid world with changing obstacles (GWCO). In GWCO problem, obstacles move on specific paths. Due to a permanent change in the location of obstacles, this can be considered as a dynamic problem. In dynamic problem, the environment is constantly changing. In this paper, we first refer to the applications of Deep learning and Reinforcement learning (RL), then to the details of Grid World and GWCO. Then it is discussed about the Deep SARSA algorithm and the results show that the agent could well find the optimal path and receive the highest reward. After learning the agent, we change the environment and add a new obstacle in the agent's path. The results show that the agent has been able to quickly propose a new path. Simulation of this paper is done with the Python software and Tensorflow.

Keywords Grid world · Deep reinforcement learning · Deep learning RL · Python · Tensorflow

M. H. Olyaei (✉) · H. Jalali · A. Noori
Faculty of Electrical Engineering, Sadjad University of Technology,
Mashhad, Iran
e-mail: mh.olyaei123@sadjad.ac.ir

H. Jalali
e-mail: h.jalali144@sadjad.ac.ir

A. Noori
e-mail: amin.noori@sadjad.ac.ir

A. Olyaei
Department of Computer Engineering, Faculty of Engineering,
Ferdowsi University of Mashhad, Mashhad, Iran
e-mail: ali.olyaei@mail.um.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_20

1 Introduction

In this paper we use Deep Reinforcement Learning (Deep RL), before we use Deep RL, we should talk about it. Deep RL is a new subject in the world that wants to revolutionize the field of artificial intelligence (AI) and show a step towards building self-governing system with a higher level understanding of visual world. Currently deep learning is enabling reinforcement learning to recognize unmanageable problem such as learning to play video games directly from pixels [1]. There are many articles talk about playing games with deep reinforcement learning for example researchers apply deep reinforcement learning methods to seven Atari 2600 games from arcade learning environment, with no using adjustment of architecture or learning algorithm [2]. Reinforcement learning is useful when we need an agent and handler to carry out many tasks, an agent must increase a control policy for taking actions in an environment, the goal of some projects and articles to learn a policy to have an agent successfully play the special games and Reinforcement Learning Agents Providing Advice in Complex Video Games. For example playing the FPS games (see Fig. 1) and the flappy bird game (see Fig. 2) [3–5].

We can use deep reinforcement learning to solve the perspective-taking task. Perspective taking is the ability that allows us to take the point of view of another agent. This capability is not special to humans, its common to humans. This is a necessary ability for agents to achieve good social interactions, including cooperation and competition [6]. Some articles talk about obstacle avoidance through deep reinforcement learning, Obstacle avoidance is a basic requirement for autonomous or self-governing robots which operate in, and communicate with, the real world [7]. Deep learning reinforcement in image captioning with embedding reward, Image captioning, the task of automatically recognize the contact of an image and

Fig. 1 Flappy bird game

Fig. 2 The FPS game

describing it with natural language, has fascinated increasingly interests in the computer vision. It can be important because it aims at endowing machines with one of the core human intelligence to find out the huge amount of visual information and to explain it in natural language [8]. The most important point of deep reinforcement learning is deep reinforcement learning in robotics, Reinforcement learning determine to robotics a framework and set of tools that can design complex and hard-to-engineer behaviors. Reinforcement learning (RL) helps a robot to autonomously discover an excellent behavior through trial-and-error interactions with its environment [9].

2 Grid World with Changing Obstacles (GWCO)

Grid World (GW) is one of the issues that can be solved today by various algorithms. The GW environment contains a number of agents, obstacles, starting point, and goal point. The number of squares in this environment varies and is selected in each case. The starting point is where the agent starts moving from this point and must learn to reach the goal point without dealing with obstacles. In Fig. 3, the sample environment of the GW problem is observed with two agents [10]. GW issues can be solved by methods such as Reinforcement Learning (RL).

The GWCO problem is similar to the GW problem, except that the obstacles are moving in a certain direction and the agent must identify the course of the obstacles and choose a path to avoid these obstacles. GWCO issues cannot be solved by the RL method because the obstacles are constantly moving, and the number of states and actions is very high. Deep RL method is used for problems with a large number of state-action (or, in other words, the Q-table has a lot of dimensions). In Sect. 3 we will introduce Deep RL. The GWCO problem defined in this paper includes three moving obstacles that move on a specific path. In Figs. 4 and 5, the location

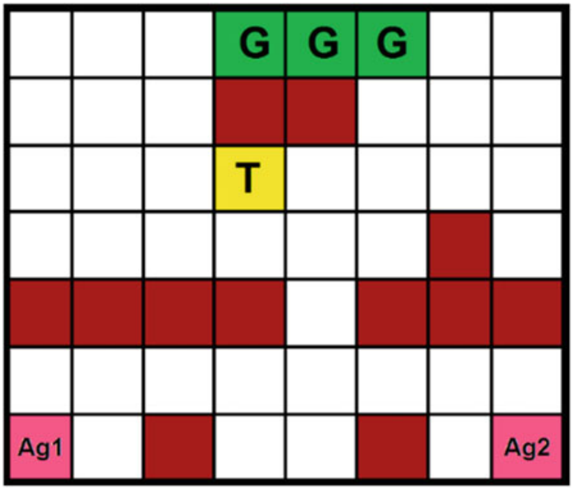


Fig. 3 Sample of GW problem [10]

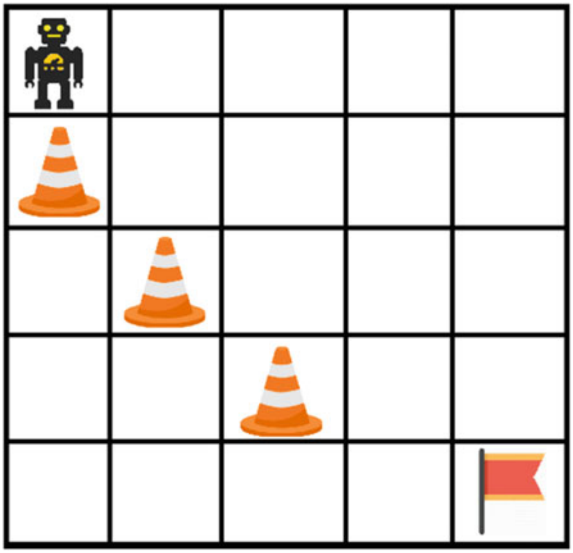


Fig. 4 Details of environment

of the obstacles, the agent, the starting point and the target have been shown, as well as in Table 1, their coordinates have been identified.

Figure 4 shows that the robot's location (agent) is located on (1, 1). Three obstacles in the places (2, 1), (3, 2) and (4, 3) and the target point are located in (5, 5).

In Fig. 5, the direction of movement of obstacles is indicated. At any moment, the location of the obstacles changes to one square. This is a dynamic issue because of the permanent change in the location of the obstacles and the deep learning is the best way to solve these types of problems.

Fig. 5 Direction of obstacles movement

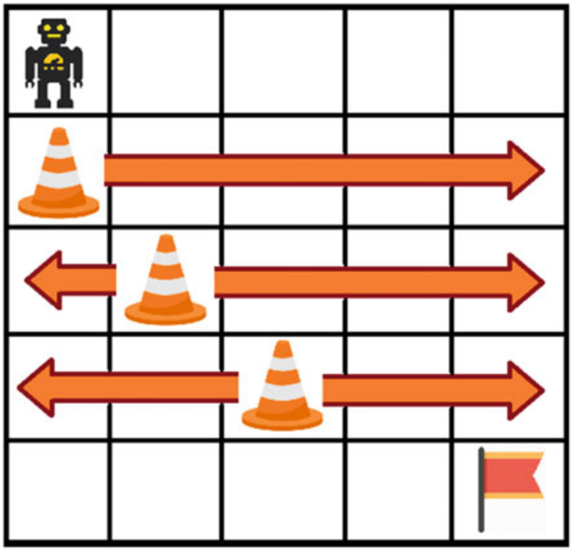


Table 1 The details of environment

Name	X coordinate	Y coordinate
(1) Robot	(2) 1	(3) 1
(4) Goal point	(5) 5	(6) 5
(7) Obstacles	(8) 1	(9) 2
	(10) 2	(11) 3
	(12) 3	(13) 4

3 Deep Reinforcement Learning

3.1 Deep Learning

Deep learning is part of the machine learning, which today is widely used in various sciences [11]. Its applications can be used to distinguish and categorize objects in images [12] and audio processing [13]. The advantage of deep learning over the neural network is the number of layers and its specific architecture, and images can be considered directly as inputs for the network. Deep learning can extract and categorize the features of an image. Features that are extracted from deep learning can be low, medium, and high level, and this depends on the choice of our network architecture [14].

3.2 Reinforcement Learning

Today, Reinforcement learning is used to solve many issues. One of the newest learning topics for reinforcement learning is the multi-agent reinforcement learning [15]. In reinforcement learning, the agent must learn to choose the best action in each state and earn the highest reward in each episode. This choice is made according to the policy of the problem. After selecting any action, the agent goes to the next state and receives the corresponding reward. Given this reward, the Q-table is updated. With a high number of state-action, the Q-table will be larger and more memory will be needed. For this reason, it is not possible to use reinforcement learning on these issues. The best way to solve such problems is to combine deep learning with reinforcement learning known as Deep Reinforcement Learning (DRL).

3.3 Sarsa Deep Reinforcement Learning (SDRL)

In deep reinforcement learning, there is no need to store a Q-table with very large dimensions, and the value of each state-action is calculated and estimated by the deep network. The deep network input can be an image with dimensions L * W * D, where L is the image length, W is the image width and D represents the depth of the image. The size of the images for this paper is 50 * 50 * 1, with a length and width of the image equal to 50 and depth is equal to 1. The network input is also a state vector. In this paper, the size vector of states is 1 * 15 and is shown in Fig. 6.

The number of deep network outputs is equal to the number of actions that the agent can choose. In this paper, the agent can do five types of moves: move up, down, left, right, and stay constant. So the number of deep network outputs is 5. Each output represents the value of an action in the considered state. The policy chosen in this paper is ε-greedy. With this policy, the action that is best, selected and after a_t action, the agent is passed to a_{t+1} state. Again for the new state-action,

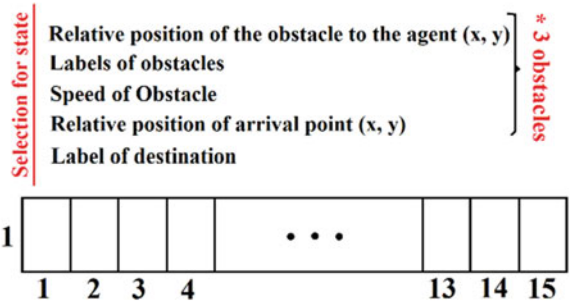


Fig. 6 States in this paper

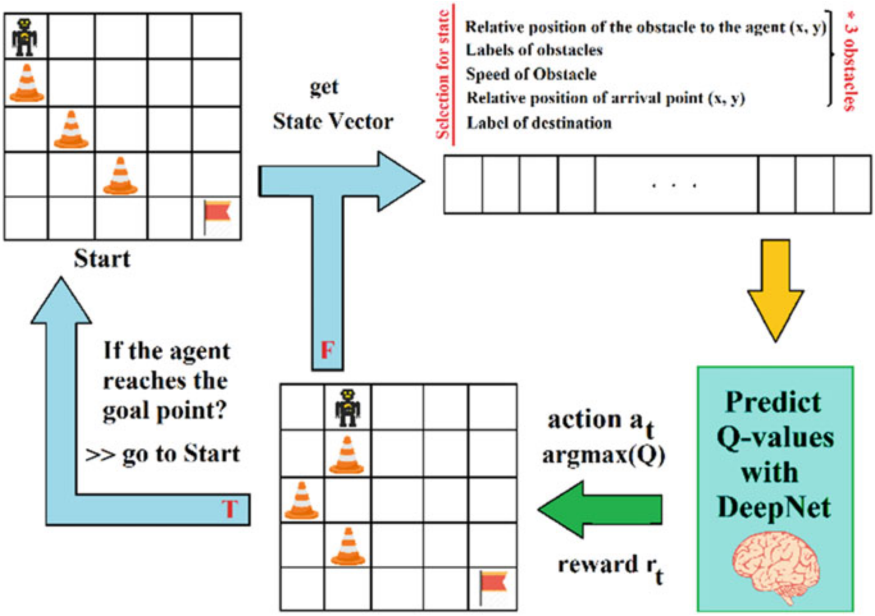


Fig. 7 The algorithm used in this paper

$Q_{(s,\ a)}$ is estimated. This process is performed as long as the agent can receive the highest amount of reward received during each episode. The algorithm performed in this paper is shown in Fig. 7.

The deep network used to estimate Q-values consists of three layers. Layers are Fully Connected. The first layer and the second layer each have 30 neurons and the output layer also has 5 neurons. There are 5 output neurons, due to 5 actions in the problem. In Fig. 8, the deep network architecture is shown. Activation functions in the first and second layers are selected ReLU and the final layer is selected as Linear.

4 Simulation Results

The simulation of this problem was done using the Python software. To define the deep network model, Tensorflow and Keras library are used. In this paper, the learning parameters are set out in Table 2.

According to Table 2, we set the learning rate to 0.01, 0.001 and 0.0001, and according to the results, we will examine which learning step is better for this problem. We also consider the reward function in accordance with Eq. (1).

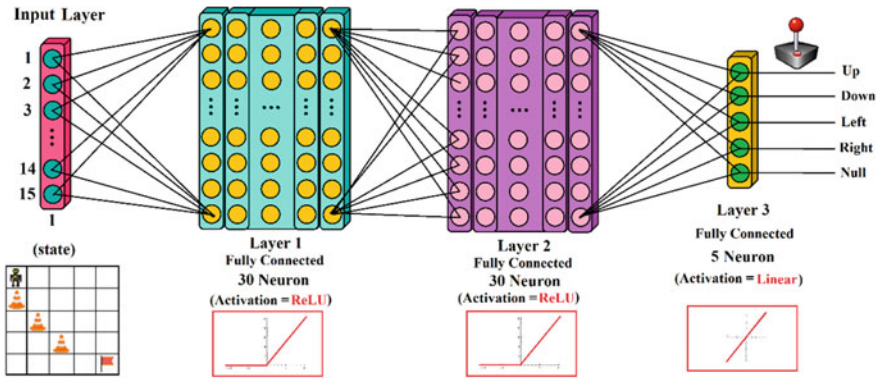


Fig. 8 Deep network architecture

Table 2 Parameters of learning

Name	Value
α—Learning rate	0.01, 0.001, 0.0001
γ—Discount factor	0.99
ε—Epsilon	1 * (0.99^Episode number)

$$\text{Reward} = \begin{cases} \text{R(obstacle)} = -1 \\ \text{R(non)} = 0 \\ \text{R(goal)} = 3 \end{cases} \tag{1}$$

In Eq. (1), it is specified that if the agent reaches the target point, the reward is +3 and if it hits the obstacles, then the reward is −1 and otherwise it receives 0 reward. Each episode also ends when the agent reaches the target point. The number of episodes is also 1000. First we will learn for 1000 episodes. Figure 9 shows the result of the reward received for 1000 episodes. Learning rate is set to 0.001.

Figure 9 shows that the robot has received a negative score of up to 300 episodes and the total reward received has reached about −400, and since the 300th episode, he has been able to learn not to encounter obstacles and increase rewards. Repeat the simulation for 1000 episodes and α = 0.01. In Fig. 10, we see the result with α = 0.01. It can be seen that with a large learning rate, the agent cannot learn and increase the reward.

Simulation video for this section (that α = 0.001) can be downloaded from www.goo.gl/CdBMxa.

In Fig. 11 we also see the result of simulation with α = 0.0001. This reduces the learning speed, and the agent in the 1000th episode cannot receive an acceptable reward.

Fig. 9 Total reward received with α = 0.001

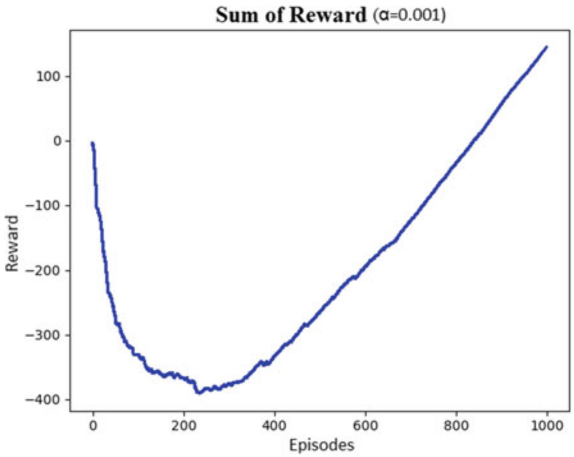


Fig. 10 Total reward received with α = 0.01

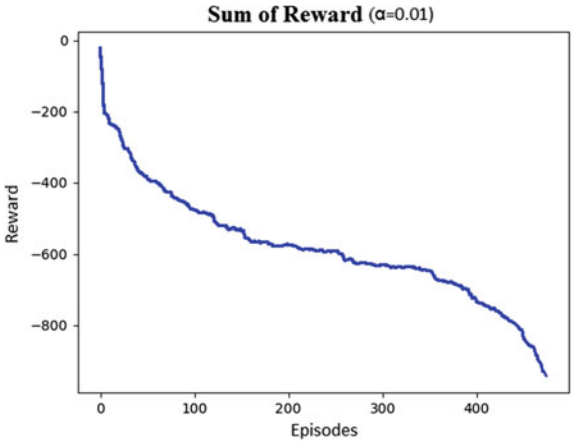
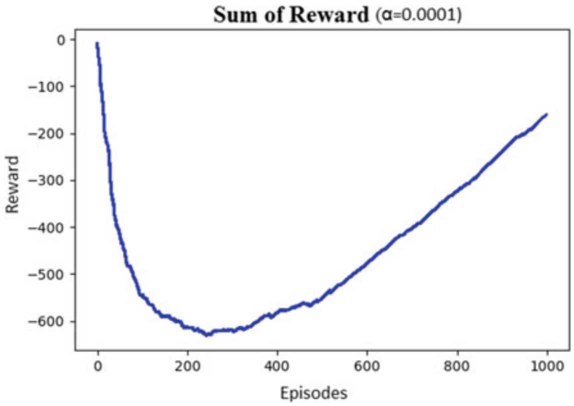


Fig. 11 Total reward received with α = 0.0001



5 Adding Uncertainty and Changing the Environment After Learning

After complete learning, we change the environment and create an obstacle in the final path that the agent moves after complete learning. We expect the agent to propose a new path by dealing with this new obstacle. Figure 12 shows a new obstacle location.

At this point, the agent will deal with the new obstacle several times and receive reward=-1, after the last 25 episodes he learns to change his path. The simulation results in Figs. 13 and 14 show that the agent proposes two new paths.

Simulation video for this section (Chang the environment after learning) can be downloaded from www.goo.gl/ZaerDT and www.goo.gl/EY9K87.

As shown in Figs. 15 and 16, the agent has been able to find a path after 30 episodes that will increase its rewards. The final reward after 175 episodes in the first path is 78 and in the second path is 115, this means that on the second path there are less obstacles.

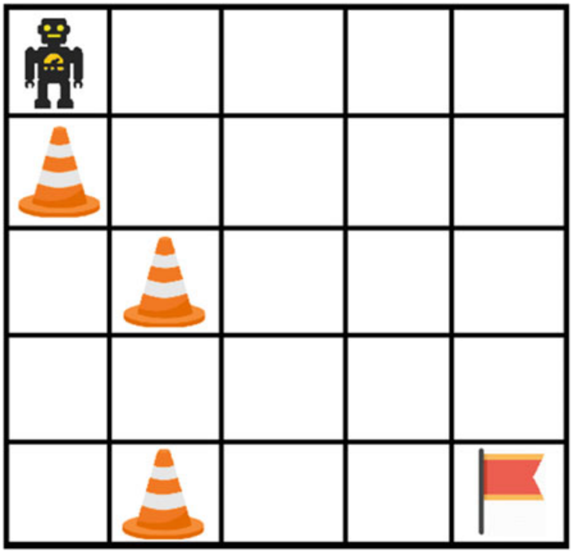


Fig. 12 Changed environment

Fig. 13 New path 1

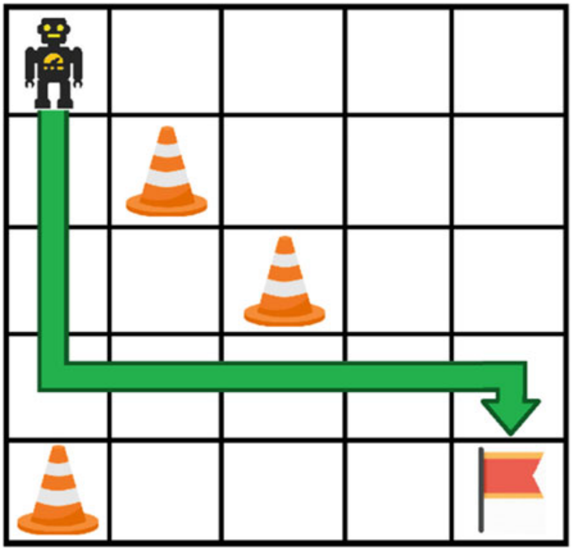


Fig. 14 New path 2

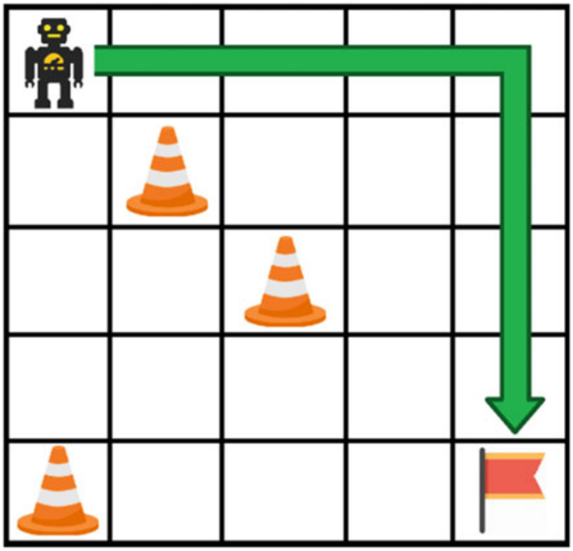


Fig. 15 Sum of reward in new path 1

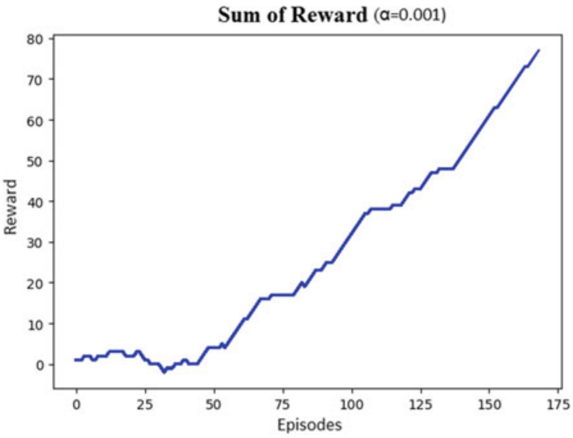
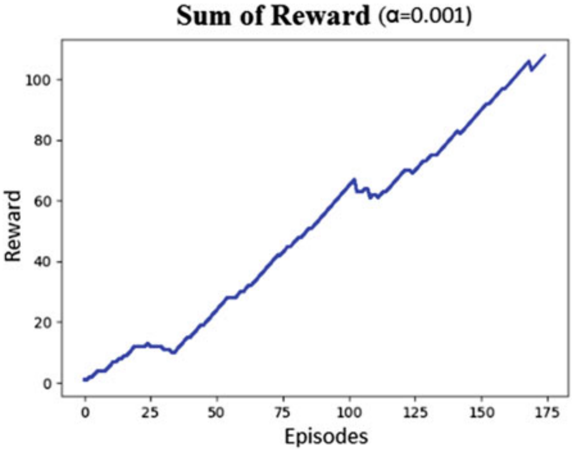


Fig. 16 Sum of reward in new path 2



6 Conclusion

In this paper, we first talked about Deep Learning, Reinforcement Learning, and Deep Reinforcement Learning. Then the GWCO problem and its conditions are explained. One of the results is that, for issues that have a large number of state-action (such as the GWCO problem), the Deep RL method is a very good method and yields an acceptable result. One of the other results is that the learning rate (α) for Deep RL issues should not be very large and very small. By comparing simulation results, it can be concluded that in order to solve the GWCO problem with the Deep RL method, it is better to consider learning rate to $\alpha = 0.001$. Finally, after changing the environment and adding a new obstacle, the results indicate that the agent, using the Deep RL method, has been able to adapt itself to the new environment and suggest new paths.

Acknowledgements We thank Danny George Juha and his friends, to provide the necessary conditions for study and research.

References

1. Arulkumaran, K, Deisenroth MP, Brundage M, Bharath AA (2017) A brief survey of deep reinforcement learning. ArXiv: 1708.05866
2. Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M (2013) Playing Atari with deep reinforcement learning. ArXiv: 1312.5602
3. Lample G, Chaplot DS (2017) Playing FPS games with deep reinforcement learning. In: AAAI, pp 2140–2146
4. Chen K (2015) Deep reinforcement learning for flappy bird. https://www.cs229.stanford.edu/proj2015/362_report.pdf
5. Taylor ME, Carboni N, Fachantidis A, Vlahavas I, Torrey L (2014) Reinforcement learning agents providing advice in complex video games. Connection Sci 26(1):45–63
6. Aqeel L (2017) Using deep reinforcement learning to solve perspective-taking task
7. Xie L, Wang S, Markham A, Trigoni N (2017) Towards monocular vision based obstacle avoidance through deep reinforcement learning. ArXiv: 1706.09829
8. Ren Z, Wang X, Zhang, N, Lv X, Li LJ (2017) Deep reinforcement learning-based image captioning with embedding reward. ArXiv: 1704.03899
9. Kober J, Bagnell JA, Peters J (2013) Reinforcement learning in robotics: a survey. Int J Robot Res 32(11):1238–1274
10. Olyaei Torqabeh MH, Noori A (2017) Reinforcement learning application of multi-agent path planning. Shiraz, Iran. https://www.civilica.com/EnPaper-ECCONF02-ECCONF02_002.html
11. Bengio Y (2009) Learning deep architectures for AI. Foundations and trends®. Mach Learn 2 (1):1–127
12. Airola R, Hager K (2017) Image classification, deep learning and convolutional neural networks: a comparative study of machine learning frameworks
13. Yu D, Deng L (2011) Deep learning and its applications to signal and information processing [exploratory dsp]. IEEE Sig Process Mag 28(1):145–154
14. Le QV (2013) Building high-level features using large scale unsupervised learning. In: 2013 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, pp 8595–8598
15. Busoniu L, Babuška R, De Schutter B (2010). Multi-agent reinforcement learning: an overview. Innovations in multi-agent systems and applications-1, vol 310, pp 183–221.

A New 1 GS/s Sampling Rate and 400 μV Resolution with Reliable Power Consumption Dynamic Latched Type Comparator

Sina Mahdavi, Maryam Poreh, Shadi Ataei, Mahsa Jafarzadeh and Faeze Noruzpur

Abstract A new high-speed dynamic latched type comparator with reliable resolution is presented in this paper. The proposed paper presents a 1 GS/s sampling rate in presence of 8 mV input offset, and it can detect the very low voltage differences such as ±200 μV at the output nodes, reliably. The power consumption and delay time of the proposed circuit are 750 μw and 257 ps with the power supply of 1.8 V, respectively. Furthermore, the proposed structure is the suitable candidate for high-speed SAR ADC, as well. Simulation results of the suggested circuit are performed using the BSIM3 model of a 0.18 μm CMOS process with the power supply of 1.8 V at all process corners along with the different temperatures in the region −50 to +50 °C, reliably.

Keywords High-resolution · High-speed · Comparator · ADC
SAR ADC

S. Mahdavi (✉)
Young Researchers and Elite Club, Tabriz Branch, Islamic Azad University,
Tabriz, Iran
e-mail: mahdavi9099@yahoo.com

M. Poreh · S. Ataei · M. Jafarzadeh · F. Noruzpur
Department of Microelectronics Engineering, Urmia Graduate Institute,
Urmia, Iran
e-mail: m.m.poreh@urumi.ac.ir

S. Ataei
e-mail: m.s.ataei@urumi.ac.ir

M. Jafarzadeh
e-mail: m.m.jafarzadeh@urumi.ac.ir

F. Noruzpur
e-mail: m.f.noruzpur@urumi.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_21

1 Introduction

Mostly, in electronic circuits, the operational amplifier is designed to be used with negative feedback. Meanwhile, it can be used as the comparator in open loop configuration, too. Due to that, the comparator is designed for open loop configuration without any feedback, especially. Comparators are mostly used in analog-to-digital converter (ADCs) [1–26]. Clearly, in the conversion process, first the input signal is sampled, then the sampled signal is applied to a number of comparators to determine the digital equivalent of the analog value. It is well-known that each comparator compares two analog input or reference signal and produces a binary signal based on the comparison. Dynamic regenerative comparators are used to signify the addition of clocks to the input of the circuit design. Also, in regenerative comparators the positive feedback like as a latch to compare the signals [1, 2, 6, 12, 14, 15, 23, 24, 26–28]. The feedback aids in providing higher speed in the circuit. Obviously, there are three main stages in a comparator; the first stage is the pre-amplifier stage. In this stage, the input signal which is fed to the comparator is being amplified. The second stage is a positive feedback stage. This is mainly used to identify the input signal which is high or low. The final stage is the decision-making stage and an output buffer stage [1–5, 13, 15, 22, 27, 29–33]. Generally, in single-stage or multi-stage comparator design, some parameters must be considered, carefully, for example, in single-stage dynamic comparators a preamplifier circuit is directly connected to a cross-coupled latch circuit, these comparators provide higher speed and lower power consumption compared to the static comparators [1–3, 8, 12, 30]. However, they suffer from kickback noise which is due to capacitive paths from output nodes to input nodes.

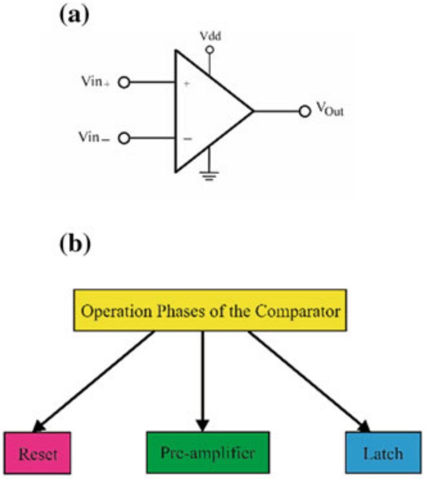


Fig. 1 **a** The simple model of the comparator and **b** its operation phases

On the other hand, in two-stage dynamic comparators, the kickback noise problem by weakening the capacitive path from the output nodes to the input nodes is solved but, the main drawback of the pre-amplifier based comparator is its comparatively large offset voltage [2, 3, 10, 12, 22, 24, 27, 29, 30, 32, 34].

In this paper, a new high-speed dynamic latched type comparator with reliable resolution is presented in this paper. The proposed paper presents a 1 GS/s sampling rate in presence of 8 mV input offset, and it can detect the very low voltage differences even ±200 μV at the output nodes, reliably.

The proposed paper is organized as follows. The general concept of the comparator is presented in Sect. 2. Section 3 describes the proposed comparator. Simulation results of the paper are supported in Sect. 4, and finally, Sect. 5 concludes the paper.

2 General Concept

Usually, each comparator has three operation phases; reset, pre-amplifier and latch, simply. In the reset mode the outputs of the comparator are connected together, in the pre-amplifier phase the input signal is amplified until the desired level and finally, the function of the latch stage is to produce full range signal as well [1, 2, 6, 19, 30]. Also, it is well-known that the comparator compares two instant analog voltages and generates digital output "1" or "0", in order to represent the polarity of the input difference. The general symbol of the comparator and its operation phases are indicated in Fig. 1a, b, respectively [1, 2, 6, 7, 23, 24, 26, 28, 30, 33–36].

In designing of the comparator circuit, some parameters are so significant which must be considered, carefully such as gain, resolution, etc. It is notable that, since the comparator is not ideal, it has a finite gain Av. The gain describes an input range where the output is "digitally" unknown, therefore, it defines how steep the transition between the digital levels, the relation of the gain, difference input voltage and power supply voltage is shown in (1), simply [1, 2, 13, 17, 24, 26, 34, 37].

$$\Delta Vin = \frac{Vdd}{Av} \tag{1}$$

On the other hand, the resolution is the minimum input voltage difference which is detectable through a comparator. Meanwhile, input referred offset and noise are limiting factors of the resolution. In an ADC the minimum required resolution is represented as LSB. For example, in 8-bit ADC with the continuous-time comparison, assume that Vdd = 1.8 V, Full Scale Range (FSR) = 0.9 V, and when the resolution of 1/2 LSB is needed, in this condition the required gain of the comparator is achieved as (2) [1, 2, 24, 34, 38].

$$Av = \frac{1.8}{(1/2)LSB} = \frac{1.8}{0.5 * (0.9/256)} \cong 1011 \quad (2)$$

Another effective parameter in dynamic latched comparator circuit is kickback noise it should be considered, the disturbance of the voltage at the input nodes of the comparator, due to the large variation of the voltage at internal nodes is called kickback noise [1, 2, 5, 8, 22, 24, 26, 30, 31, 33–35].

3 The Proposed Comparator

The proposed comparator is indicated in Fig. 2. As it is clear that in Fig. 2, the suggested comparator is collected of one pre-amplifier stage and latch stage, noticeably. The first stage is a voltage amplification stage and the second one is the regenerative latch stage, that uses two inverters to avoid static current and that achieves rail-to-rail digital outputs, out+, and out−, furthermore, they guarantee that

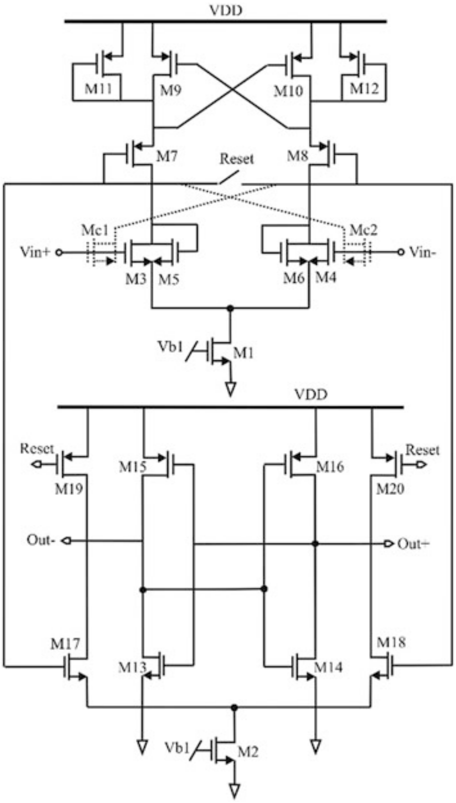


Fig. 2 The proposed comparator structure

the output signal will produce logic levels for the succeeding digital circuit [1, 2, 5, 6, 12–15, 17, 18, 21, 24, 32]. It is important that, in a latch type comparator, other than thermal noise, the main source of errors are comparator offset, clock feedthrough, and kick-back noise. Commonly, in a SAR ADC, the comparator offset, normally constant, can be tolerated, since it results in an offset in the ADC I/O transfer curve that can easily be tolerated. However, the other two errors must be minimized [1, 13, 18, 19, 21–25, 28, 33, 34, 36–39]. As discussed above, clock feedthrough and kick-back noise are generated at the edge of the comparator clock and the edge of the comparator input signal, respectively. A basic solution is to use a preamplifier that reduces the input- referred effects of these errors. However, this noticeably increases the comparator power consumption. Normally, the pre-amplifier is used to amplify the input signal and block the kickback noise. Although the output signal of the pre-amplifier is larger than the input signal, it is not large enough to drive the digital circuitry, due to that a latch stage is required to create the full range signal at the output nodes of the comparator basically. In the proposed structure, the input signals (Vin+, Vin−) are applied to the gate terminals of the M3–M4. Also, M9–M10 and M11–M12 are utilized to play the function of positive and negative feedback, respectively. Meanwhile, the dashed transistors Mc1–Mc2 are applied to decrease the effect of the kick-back noise [2], and, M5–M8 are utilized to increase the reliability of the circuit, generally. Then the amplified signals are employed to the gate terminals of the M17 and M18 correspondingly to generate full range signals at the output nodes (out+, out−) as well. Meanwhile, M19–M20 and reset switch are applied to increase the Dc gain and short the outputs of the preamplifier stage, consistently. Besides, M1 and M2 are the current sources of the circuit, and the transistors of the M13–M16 play function of the latch stage to produce full range signal, reliably.

4 Simulation Results

The simulation results of the proposed comparator are presented in this section. Figure 3 indicates the comparison results with the input differences of +200 and −200 µV at 1 GS/s, in presence of 8 mV input offset, consistently. Also, Fig. 4 illustrates the comparison results of the proposed comparator, when it is utilized in the SAR ADC, in that condition, as the mentioned Fig shows that, the proposed comparator is capable detect the small input difference voltage (output voltage of DAC) even 150 µV, as well. Meanwhile, Fig. 5 shows the delay time in proposed comparator versus input differential voltage for different supply voltages, merely. It is notable that, as Fig. 5 indicates, in the Worst-Case state, when Vdd = 0.8 V the delay time increases to 348 ps, also, when Vdd = 1.8 V, the delay time is decreased to 257 ps as well. The power consumption of the proposed circuit is 750 µW with

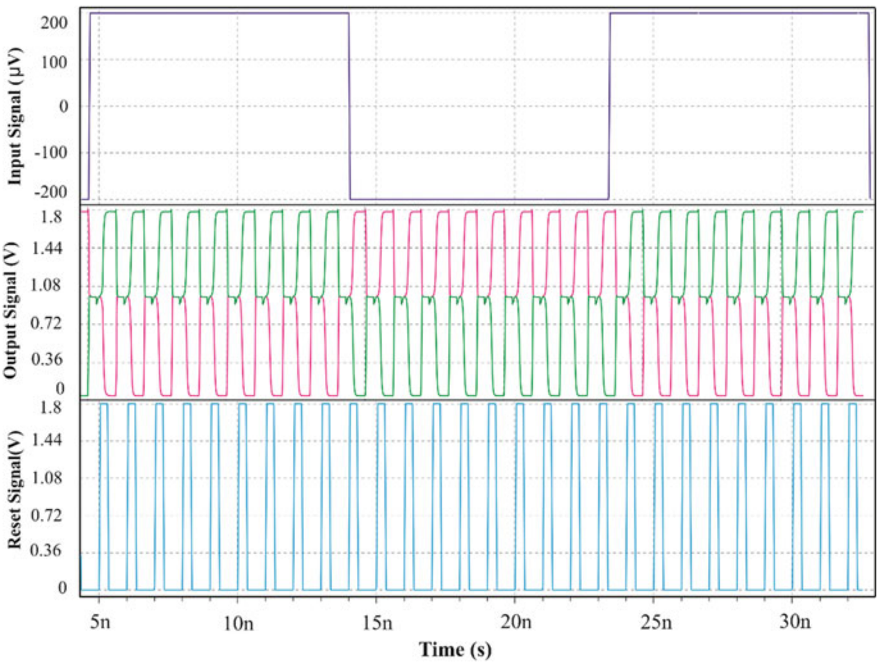


Fig. 3 The comparison results of the comparator with the input differences of +200 and −200 μV at 1 GS/s, in presence of 8 mV input offset

the power supply of 1.8 V. Finally, Table 1 summarizes the performance of the proposed dynamic comparator. Simulations are performed for all corner conditions using the BSIM3 model of a 0.18 μm CMOS process with the power supply of 1.8 V at all process corners along with the different temperatures in the region −50 to +50 °C, reliably.

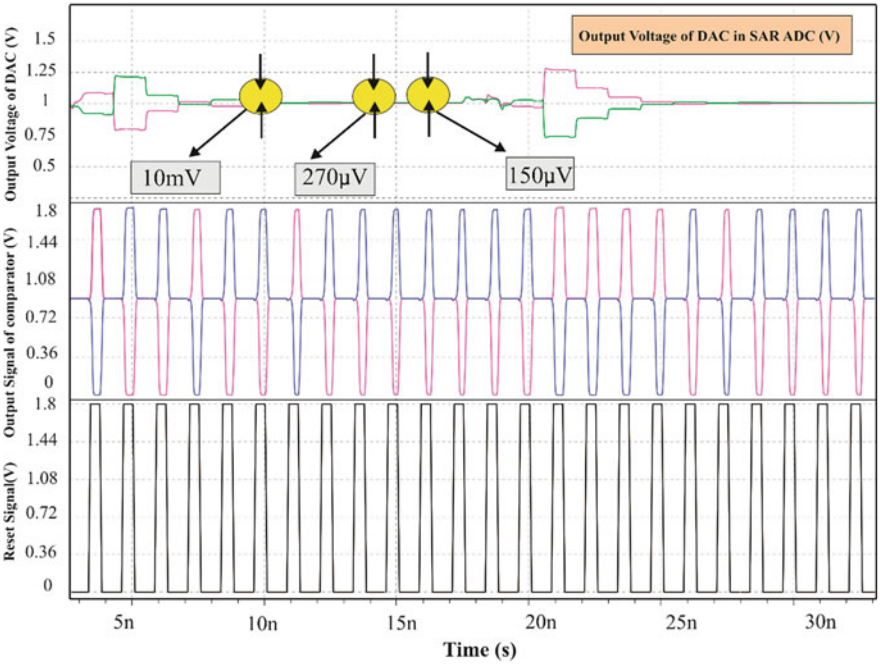


Fig. 4 The comparison results of the comparator with applying the Worst-Case input (output voltage of DAC) when comparator is used in SAR ADC at 1 GS/s, in presence of 8 mV input offset

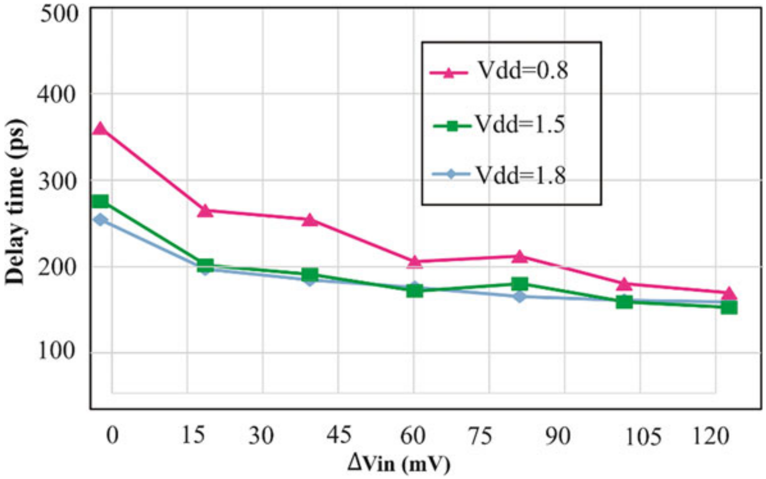


Fig. 5 Delay time of the proposed comparator versus supply voltage and input voltage difference

Table 1 Specifications obtained for designed comparator

Specifications	Value
Technology (nm)	180
Supply voltage (V)	1.8
Resolution (μV)	400
Clock frequency (GS/s)	1
Power consumption (μW)	750
Offset voltage (mV)	8
Power supply noise (mV)	100
Propagation delay time (ps)	257

5 Conclusion

A new high-speed dynamic latched type comparator with reliable resolution is presented in this paper. The proposed paper presents a 1 GS/s speed in presence of 8 mV input offset, and it can detect the very low voltage differences even ±200 μV at the output nodes, reliably. The power consumption and delay time of the paper are 750 μw and 257 ps with the power supply of 1.8 V, respectively. Furthermore, the proposed structure is the suitable candidate in high-speed SAR ADC, as well. Simulation results of the proposed structure are simulated using the HSPICE BSIM3 model of a standard 0.18 μm CMOS process.

References

1. Mahdavi S (2017) A 12 bit 76 MS/s SAR ADC with a capacitor merged technique in 0.18 μm CMOS technology. J Elect Comput Eng Innov (JECEI) 5(2):4
2. Figueiredo PM, Vital JC (2006) Kickback noise reduction techniques for CMOS latched comparators. IEEE Trans Circ Syst II Express Briefs 53(7):541–545
3. Kazeminia S, Mahdavi S (2016) A 800 MS/s 150 μV input-referred offset single-stage latched comparator. In: 2016 MIXDES—23rd international conference mixed design of integrated circuits and systems, pp 119–123
4. Jackuline Moni D, Jisha P (2012) High-speed and low-power dynamic latch comparator. In: 2012 ICDCS, pp 259–263
5. Akbari M, Nejad, MM, Mirbozorgi SA (2013) A new rail-to-rail ultra low voltage high speed comparator. In: 2013 ICEE, pp 1–6
6. Rabiei A, Najafizadeh A, Khalafi A, Ahmadi SM (2015) A new ultra low power high speed dynamic comparator. In: 2015 23rd Iranian conference on electrical engineering, pp 1266–1270
7. Majumder A, Das M, Nath B, Mondal AJ, Bhattacharyya BK (2016) Design of low noise high speed novel dynamic analog comparator in 65 nm technology. In: 2016 26th international conference radio elektronika (Radioelektronika), pp 115–120
8. Sharifi Gharabaghlo N, Moradi Khaneshan T (2017) High resolution CMOS voltage comparator for high speed SAR ADCs. In: 2017 ICEE, pp 511–514
9. Nanda S, Panda AS, Moganti GLK (2015) A novel design of a high speed hysteresis-based comparator in 90-nm CMOS technology. In: 2015 ICIP, pp 388–391

10. Singh A, Marwah A, Akashe S (2015) Design of novel low power dynamic latch comparator using multi-fin technology. In: 2015 ICCN, pp 107–110
11. Lahariya A, Gupta A (2015) Design of low power and high speed dynamic latch comparator using 180 nm technology. In: 2015 ISPCC, pp 129–134
12. Kapadia DN, Gandhi PP (2013) Implementation of CMOS charge sharing dynamic latch comparator in 130 and 90 nm technologies. In: 2013 IEEE conference on information and communication technologies, pp 16–20
13. Solis CJ, Ducoudray GO (2010) High resolution low power 0.6 µm CMOS 40 MHz dynamic latch comparator. In: 2010 53rd IEEE international midwest symposium on circuits and systems, pp 1045–1048
14. Li Y, Zeng T, Chen D (2013) A high resolution and high accuracy R-2R DAC based on ordered element matching. In: 2013 IEEE international symposium on circuits and systems (ISCAS2013), pp 1974–1977
15. Taherzadeh-Sani M, Lotfi R, Nabki F (2014) A 10-bit 110 kS/s 1.16 µW SA-ADC with a hybrid differential/single-ended DAC in 180-nm CMOS for multichannel biomedical applications. IEEE Trans Circ Syst II Express Briefs 61(8):584–588
16. Liu C, Huang M, Hsuan Tu Y (2016) A 12 bit 100 MS/s SAR-assisted digital-slope ADC. IEEE J Solid-State Circ 51(12):2941–2950
17. Aghaie S, Mueller J, Wunderlich R, Heinen S (2014) Design of a low-power calibratable charge-redistribution SAR ADC. In: 2014 10th conference on Ph.D. research in microelectronics and electronics (PRIME), pp 1–4
18. Mahdavi S, Hadidi K, A 14 bit 17 MS/s 80 dB SNDR low power SAR ADC with energy-efficient switching procedure (unpublished)
19. Lai W, Huang J, Hsieh C (2014) A 10-bit 20 MS/s successive approximation register analog-to-digital converter using single-sided DAC switching method for control application. In: 2014 CACS international automatic control conference (CACS 2014), pp 29–33
20. Chung Y, Wu M, Li H (2015) A 12-bit 8.47-fJ/conversion-step capacitor-swapping SAR ADC in 110-nm CMOS. IEEE Trans Circ Syst I Regul Pap 62(1):10–18
21. Cho Y, Jeon Y, Nam J, Kwon J (2010) A 9-bit 80 MS/s successive approximation register analog-to-digital converter with a capacitor reduction technique. IEEE Trans Circuits Syst II Express Briefs 57(7):502–506
22. Mahdavi S, Ghadimi E (2017) A new 13-bit 100 MS/s full differential successive approximation register analog to digital converter (SAR ADC) using a novel compound R-2R/C structure. In: 4th international conference on knowledge-based engineering and innovation (KBEI-2017), pp 0237–0242
23. Babayan-Mashhadi S, Lotfi R (2014) Analysis and design of a low-voltage low-power double-tail comparator. IEEE Trans Very Large Scale Integr VLSI Syst 22(2):343–352
24. Mahdavi S, Jafarzadeh M, Poreh M, Ataei S (2017) An ultra high-resolution low propagation delay time and low power with 1.25 GS/s CMOS dynamic latched comparator for high-speed SAR ADCs in 180 nm technology. In: 4th international conference on knowledge-based engineering and innovation (KBEI-2017), pp 0260–0265
25. Mahdavi S, Poreh M, Alizadeh L, Moradkhani B, Ebrahimi R (2017) A 1.25 GS/s 12 bit and 2.27 mW digital to analog converter (DAC) with 70.22 SNDR based on new hybrid R-C procedure in 180 nm CMOS. Int J Microelectron Comput Sci Poland (in press)
26. Madhab Dhal L, Pradhan A (2012–2013) Study and analysis of different types of comparators. A thesis submitted to the National Institute of Technology for the degree of Bachelor of Technology in Electronics and Communication Engineering, Rourkela, 2012–2013
27. Khorami A, Dastjerdi MB, Ahmadi AF (2016) A low-power high-speed comparator for analog to digital converters. In: 2016 ISCAS, pp 2010–2013
28. Lee H, Park S, Lim C, Kim C (2015) A 100-nW 9.1-ENOB 20-kS/s SAR ADC for portable pulse oximeter. IEEE Trans Circ Syst II Express Briefs 62(4):357–361

29. Haenzsche S, Höppner S, Ellguth G, Schüffny R (2014) A 12-b 4-MS/s SAR ADC with configurable redundancy in 28-nm CMOS technology. IEEE Trans Circ Syst II Express Briefs 61(11):835–839
30. Ahmadi M, Namgoong W (2015) Comparator power minimization analysis for SAR ADC using multiple comparators. IEEE Trans Circ Syst I 62(10):2369–2379
31. Sujatha K, Narayana Bhagirath T, Garje KK (2015) Design and simulation of high speed comparator for LVDS receiver application. In: 2015 annual IEEE India conference (INDICON), pp 1–6
32. Aakash S, Anisha A, Jaswanth Das G, Abhiram T, Anita JP (2017) Design of a low power, high speed double tail comparator. In: 2017 ICCPCT, pp 1–5
33. Kazeminia S, Mahdavi S, Gholamnejad R (2016) Bulk controlled offset cancellation mechanism for single-stage latched comparator. In: 2016 MIXDES—23rd international conference mixed design of integrated, pp 174–178
34. Diplom O (2013) A study of SAR ADC and implementation of 10-bit asynchronous design. A thesis Submitted to the University of Texas at Austin for the degree of Master of Science in Electrical Engineering
35. Mousazadeh M (2006) A 10 bit 100 MS/s ADC based on new pipeline of successive approximation ADC. A thesis submitted to the Urmia University for the degree of Master of Science in Electrical Engineering
36. Yoshioka M, Ishikaw K, Takayama T, Tsukamoto S (2010) A 10-b 50-MS/s 820-μW SAR ADC with on-chip digital calibration. IEEE Trans Biomed Circ Syst 4(6):410–416
37. Ha H, Lee S, Kim B, Park H, Sim J (2014) A 0.5-V, 1.47-μW 40-kS/s 13-bit SAR ADC with capacitor error compensation. IEEE Trans Circ Syst II Express Briefs 61:840–844
38. Lim S, Kim J, Yoon K, Lee S (2013) A 12-b asynchronous SAR type ADC for biosignal detection. J Semicond Technol Sci 13(2):108–113
39. Khosrov DS (2010) A new offset cancelled latch comparator for high-speed, low-power ADCs. In: IEEE Asia Pacific conference on circuits and systems, APCCAS 2010, pp 13–16

Improved Ring-Based Photonic Crystal Raman Amplifier Using Optofluidic Materials

Amire Seyedfaraji

Abstract Ring-based photonic crystal (PhC) structure for Raman amplifier is investigating in this article. Then using optofluidic materials in the holes on the sides of the signal path, pump and signal group velocity reducing that cause Raman gain increase. In order to achieve bigger Raman gain, we use two-ring structure. The time evolution and propagation of picosecond signal pulses and dispersion inside the device are analyzed and Raman gain, Raman bandwidth and bit rate are studied in one-ring and two-ring structures. Maxwell equations are solved by finite difference time domain (FDTD) method and considering the optical nonlinear parameters of two photon absorption, free carrier absorption, Kerr effect and self-phase modulation in PhC structure. From a structure with a length of 100 μm, Raman gain of 19.01 dB and bit rate of 0.6493×10^{12} pulse/sec are achieved.

Keywords Raman amplifier · Ring-based · Photonic crystal · Optofluidic materials · Bit rate · Maxwell equation

1 Introduction

Recently, we observe a significant development in use of silicon (Si) based fast modulators, photo-detectors, optical amplifiers and sources. Silicon photonics is turning out as low-cost optoelectronic solutions for a variety of applications from telecommunications and interconnects to optical sensing and biomedical applications. Stimulated Raman scattering has been used as a successful approach for optical amplification and lasing in Si [1].

In Raman spontaneous scattering, thermal vibration of lattice at frequency ω_v, produce a sinusoidal modulation in optical susceptibility. This frequency is 15.6 THz in Si. With a collision of input pump field (ω_l), with optical susceptibility

A. Seyedfaraji (✉)
Faculty of Engineering and Technology, Alzahra University,
P.O. Box 1993893973, Tehran, Iran
e-mail: sfaraji@alzahra.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_22

(ω_v), some polarizations create at the sum frequency $(\omega_v + \omega_l)$, and at the difference frequency $(\omega_l - \omega_v)$. The radiation produced by these two polarization components, respectively called anti-Stokes and Stokes waves. In stimulated Raman scattering, atomic vibration can be excited by simultaneous propagation of pump and Stokes field, which amplifies the Stokes field [2].

In recent years, Raman gain in Si waveguides has been extensively studied [3–5]. For improve the efficiency of Raman amplifiers, Si nano waveguides [6], SiGe waveguides [7], PhC [8, 9], and hybrid PhC (HPhC) waveguides [10] and slow-light grating waveguides [11] have been used. Some rules have been developed to design Raman amplifiers, using analytical and semi-analytical methods [12] and geometric waveguides parameters have been optimized to improve Raman amplifiers performance [13].

On the other hand, in slow-light regime, where light moves slower through the material, light-matter interaction time is increased and nonlinear effects will be intensified. As well as reduction of group velocity reduces the pump power or physical length required to appear nonlinear effects. For this reason, to reduce the size and increase the intensity of nonlinear effects, slow light is employed [14, 15].

PhC waveguides, provide strong mode confinement and low-group velocities through structural Bragg reflections. Such properties account for enhancing nonlinear optical phenomena such as Raman scattering and therefore larger Raman gain will be obtained with smaller input pump power [16].

To make efficient use of the pump power, we have already presented a new configuration of ring-based Raman amplifier [17], PhC ring-based Raman amplifier and HPhC ring-based Raman amplifier [18]. The resonance effect enhances the effective pump power and thus can achieve the same level of Raman gain at a much lower input pump power.

In this paper, using optofluidic materials [19, 20] in the holes on the sides of the signal path, we reduce the group velocity for signal and pump wavelength and increase the matter and light interaction time which consequently results in stronger nonlinear effects. In this way, by decreasing group velocity in PhC ring-based structure, input pump power can be reduced and higher Raman gain can be achieved.

This paper is organized as follows. In Chap. 2 Using Maxwell equations, we model Raman amplification in PhC waveguides considering nonlinear effects of two photonic absorption (TPA), free carrier absorption (FCA), Kerr effect and self-phase modulation (SPM) effect. Chap. 3 deals with Raman amplification simulation results in PhC ring-based structure and improved PhC ring-based structure. We summarize the results of the paper in Chap. 4.

2 Modeling Theory

As a consequence of Kerr nonlinear effect, two photon absorption effect and other nonlinear effects, the refractive index of a material is dependent on the optical power. When a strong light pulse passes through a medium, it will induce a phase shift $\Delta\varphi$ due to optical nonlinearities. When ultra short pulses are used, the intensity rapidly varies in time which results in phase change. The time derivative of the phase change yields a frequency shift $\Delta\omega$ across the pulse defined by (1)

$$\Delta\omega(x, y, t) = -\frac{d(\Delta\varphi(x, y, t))}{dt} \tag{1}$$

That $\Delta\varphi$ is given by (2) [21].

$$\Delta\varphi(x, y, t) = \frac{2\pi L_{\text{int}}}{\lambda_P}(\Delta n_{kerr}(x, y, t) + \Delta n_{FC}(x, y, t)) \tag{2}$$

L_{int} is the interaction length. Δn_{kerr} is Kerr-induced refractive index change and Δn_{FC} is free-carrier induced refractive index change. λ_p is center wavelength of the pump pulse along the waveguide.

As a result of falling and rising edges of $\Delta\varphi$, positive and negative changes create in $\Delta\omega$. In addition, passing the pump pulse from anywhere on the waveguide, causes carrier density to be increased, which consequently results in larger reflective index. So, central wavelength will have blue shift. Center wavelength of the pump pulse is given by (3) in each point [21].

$$\lambda_P(x, y, t) = \frac{\lambda_0}{1 - (L_{\text{int}}/c).(d\Delta n(x, y, t)/dt)} \tag{3}$$

In this equation, Δn is the sum of the changes of reflective index arises from nonlinear Kerr effect and FCA effect. λ_0 is initial wavelength of entrance pump pulse, and c is light velocity.

By increasing pump power, the losses due to TPA and FCA are intensified. So these two phenomena should be considered in simulation [22]. Raman effect and TPA are modeled with 3rd order nonlinear optical susceptibility [23, 24]. TPA causes some carrier density changes that affect the refractive index and gain coefficient and therefore real and imaginary parts of the first order optical susceptibility are changed [24]. SPM effect is modeled with real part of 3rd order nonlinear optical susceptibility. Considering these effects, the electric polarization for the pump and signal are given by (4) and (5).

$$\begin{aligned}\mathbf{P_S} &= \mathbf{P}(\omega_S)\\ &= \chi(\omega_S)\mathbf{E_S} + \varepsilon_0\chi^{(3)}_{\text{Im}}(\omega_S)\mathbf{E_P}.\mathbf{E^*_P}.\mathbf{E_S} + \varepsilon_0\chi^{(3)}_{\text{Im}}(\omega_S)\mathbf{E_S}.\mathbf{E^*_S}.\mathbf{E_S} + \varepsilon_0\chi^{(f)}_P(\omega_S, N)\mathbf{E_S}\end{aligned} \tag{4}$$

$$\begin{aligned}\mathbf{P_P} = \mathbf{P}(\omega_P) = \chi(\omega_P)\mathbf{E_P} + \varepsilon_0\chi^{(3)}_{\text{Im}}(\omega_P)\mathbf{E_S}.\mathbf{E^*_S}.\mathbf{E_P} + \varepsilon_0\chi^{(3)}_{\text{Im}}(\omega_P)\mathbf{E_P}.\mathbf{E^*_P}.\mathbf{E_P}\\ + \varepsilon_0\chi^{(f)}_P(\omega_P, N)\mathbf{E_P} + \varepsilon_0\chi^{(3)}_{\text{Re}}(\omega_P)\mathbf{E_P}.\mathbf{E^*_P}.\mathbf{E_P}\end{aligned} \tag{5}$$

where $\mathbf{E_s}$ is signal electric field, $\mathbf{E_p}$ is pump electric field, ε_0 is permittivity of free space. N is carrier density. ω_S and ω_P are signal and pump frequency, respectively. χ is the first order optical susceptibility and $\chi^{(3)}_{\mathrm{Im}}$ and $\chi^{(3)}_{\mathrm{Re}}$ are imaginary part and real part of 3rd order nonlinear optical susceptibility, respectively. $\chi^{(f)}$ is the optical susceptibility that models the FCA defined as (6) [24].

$$\chi_v^{(f)} = 2n_0\left(n_{f\upsilon} + ic\frac{\alpha_{f\upsilon}}{2\omega_\upsilon}\right) \tag{6}$$

$$n_{f\upsilon}(\omega_\upsilon, N) = -\frac{q^2N}{2\varepsilon_0 n_0\omega_\upsilon^2}\left(\frac{1}{m_{ce}} + \frac{1}{m_{ch}}\right) \tag{7}$$

$$\alpha_{f\upsilon}(\omega_\upsilon, N) = \frac{q^3N}{\varepsilon_0 cn_0\omega_\upsilon^2}\left(\frac{1}{\mu_e m_{ce}^2} + \frac{1}{\mu_h m_{ch}^2}\right) \tag{8}$$

where the index υ denotes S or P for signal or pump, respectively. n_0 is linear refractive index, $n_{f\upsilon}$ is free carrier index change or rate of change in the refractive index due to the carrier density changes and $\alpha_{f\upsilon}$ is free carrier absorption or rate of change in the absorption coefficient due to the carrier density changes. m_{ce}, m_{ch}, μ_e and μ_h are effective mass of electron, effective mass of hole, electron mobility and hole mobility, respectively [24].

Carrier density change caused by TPA is modeled as (9)

$$\frac{dN}{dt} = -\frac{N}{\tau} + \frac{\beta_2(\omega_S)I_S^2}{2\hbar\omega_S} + \frac{\beta_2(\omega_P)I_P^2}{2\hbar\omega_P} \tag{9}$$

$$I_S = \frac{\varepsilon_0 cn}{2}|E_S|^2 \quad I_P = \frac{\varepsilon_0 cn}{2}|E_P|^2 \tag{10}$$

$$\beta_2(\omega) = \frac{3\omega\chi^{(3)}_{\mathrm{Im}}}{2\varepsilon_0 c^2 n_0^2} \tag{11}$$

where β_2 is TPA coefficient and τ is the carrier lifetime.

The Kerr effect causes the linear refractive index of a material to be linearly dependent on the optical intensity, according to (12)

$$n = n_0 + n_2 I_P \tag{12}$$

where n_2 is Kerr coefficient [21].

We obtain the relative permittivity from electric polarization and substitute it into Maxwell's equations and solve them using the finite difference time domain (FDTD) method. A perfectly matched layer boundary condition has been used.

3 Results and Discussion

Schematic structure of proposed PhC ring-based Raman amplifier (SR) and improved PhC ring-based Raman amplifier using optofluidic materials (SR-a) are shown in Fig. 1a, b, respectively.

These micro rings have been made in Si hexagonal 2D PhC slab with air holes. Structure period is *a* and holes radii are *r*. Structural parameters should be selected

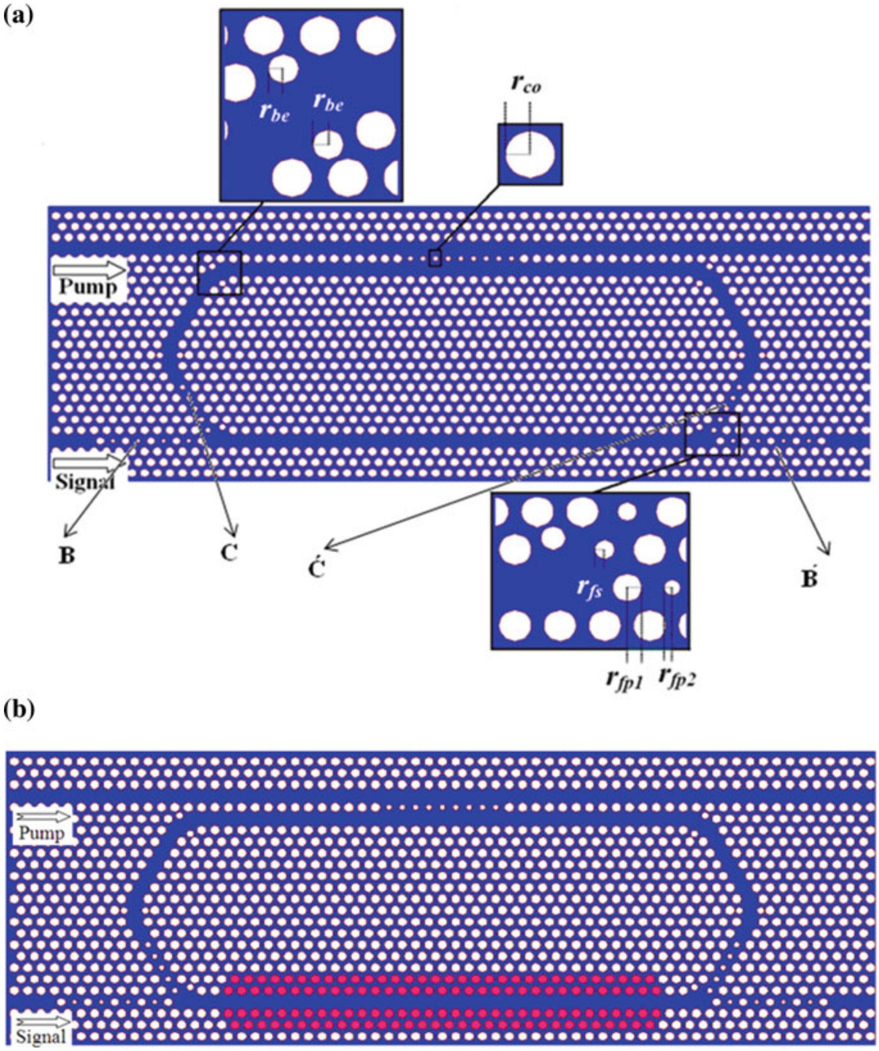


Fig. 1 Structure of **a** PhC ring-based Raman amplifier (SR), and **b** improved PhC ring-based Raman amplifier using optofluidic materials (SR_a)

such that pump wavelength (TM, 1.55 μm) and signal wavelength (TE, 1.686 μm) propagate inside the PhC waveguide.

These structures have two separate entrances for pump and signal [18]. Upper waveguide is the entrance of pump. The perimeter of the ring is optimized for the resonance of pump wavelength. The geometric parameters have been selected such that the structures have critical coupling for pump wavelength. Lower waveguide is the entrance of signal which passes through the lower side of the ring.

Inside the waveguide before the ring and after that indicated by B and B′, several air holes (nano defects) with specified radii and distances are designed such that pump wavelength cannot pass. We call these two parts as pump filter. Thus pump wavelength can resonate inside the ring causing its intensity to be increased. On the other hand, inside the two sides of the ring that indicated by C and C′, air holes (nano defects) with specified radii and distances that are called signal filter prevent the signal entry. So, the signal only passes the direct path without any perturbation inside the ring.

In Fig. 1b using optofluidic materials in the holes on the sides of the signal path, the advantage of lower group velocity can be used more efficiently. The refractive index of optofluidic material is $n_{of.}$

Geometric properties of these structures are presented in Table 1.

Transmission spectra of pump and signal filter, corresponding to B, B′ and C, C′, are shown, respectively in Figs. 2 and 3. In these figures, part (a) corresponds to around pump wavelength (TM, 1.55 μm) and part (b) is relevant to the signal wavelength (TE, 1.686 μm). As can be seen, pump filter does not pass pump wavelength and signal filter does not pass signal wavelength, whereas pump wavelength and signal wavelength pass respectively through the signal filter and pump filter with very little loss.

There are different methods to improve the quality of passing through the bend in PhC structures [19, 20, 25, and 26]. Due to the [25] to design the bends, the air hole at the inner corner is made smaller, and one air hole is added at the outer corner of the bend. The two air holes have the radii of r_{be}= 0.75 × r, and they are moved 0.3 × a oppositely along the symmetric axis of the bend.

Table 1 Physical parameters of PhC ring-based Raman amplifiers

	SR	SR_a
a (nm)	460	460
$\boldsymbol{r}$ (nm)	165.6	165.6
$\boldsymbol{r_{fs}}$(nm)	89.7	89.7
$\boldsymbol{r_{fp1}}$(nm)	147.2	147.2
$\boldsymbol{r_{fp2}}$(nm)	69	69
$\boldsymbol{r_{co}}$(nm)	308.2	308.2
$\boldsymbol{r_{be}}$(nm)	124.2	124.2
$\boldsymbol{n_{of}}$	–	1.5

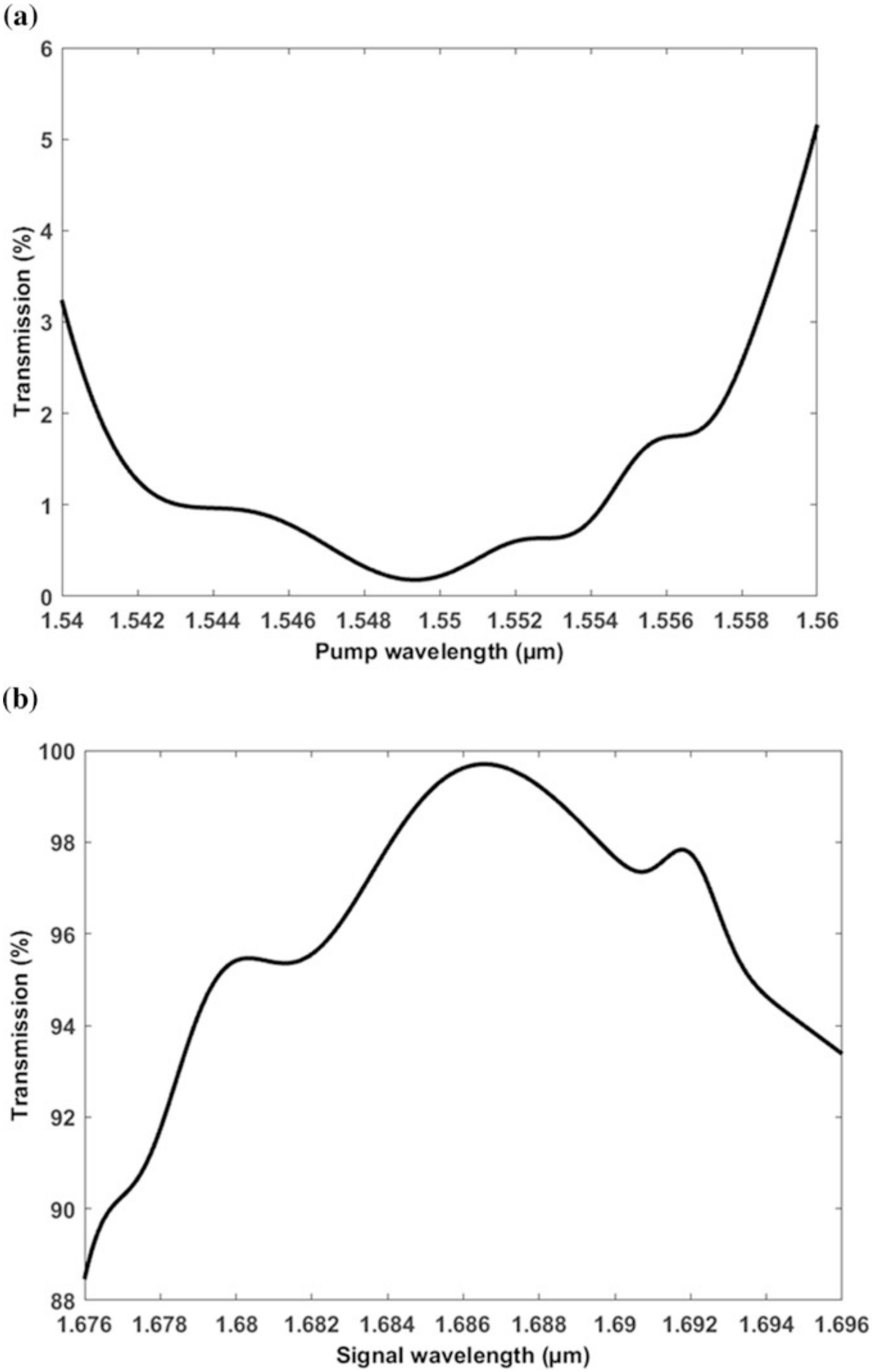


Fig. 2 Transmission spectra of pump filter around **a** pump wavelength and **b** signal wavelength

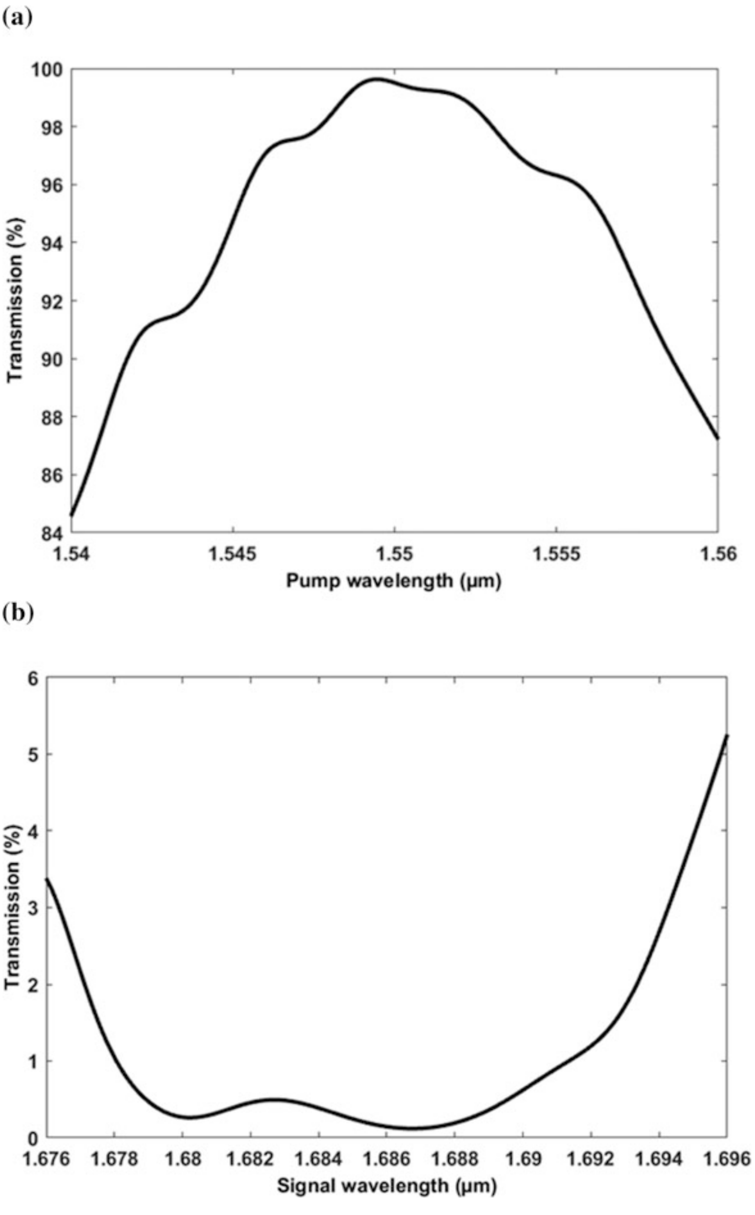


Fig. 3 Transmission spectra of signal filter around **a** pump wavelength and **b** signal wavelength

The bend transmission spectra around pump wavelength (TM, 1.55 μm) of PhC ring-based structure is shown in Fig. 4. As can be seen, pump wavelength pass through the bend with little loss.

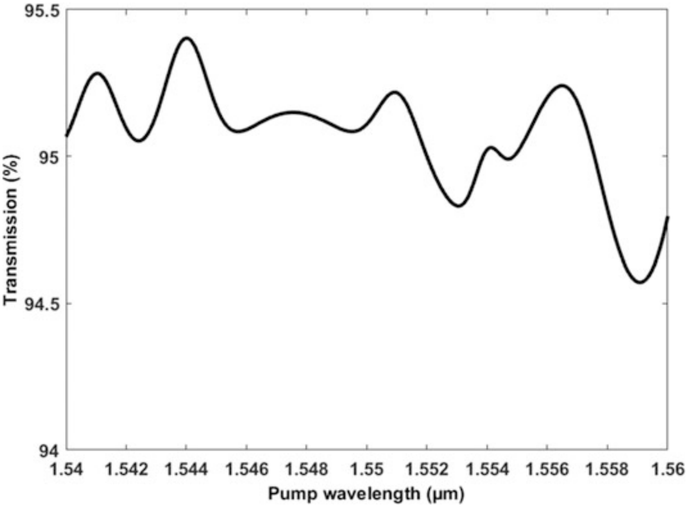


Fig. 4 The bend transmission spectra around pump wavelength (TM, 1.55 μm) of PhC ring-based structure

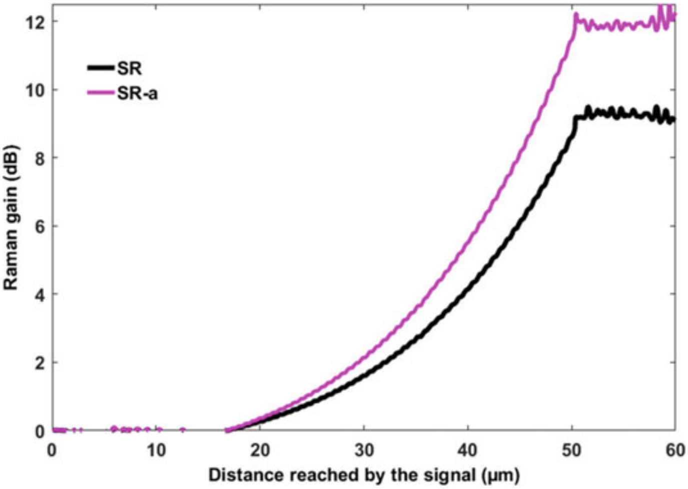


Fig. 5 Raman gain along signal path for 2 ring-based PhC structures (SR and SR-a) for pump power of 0.3 W

Figure 5 shows Raman gain along signal path for these 2 structures (SR and SR-a). Length of amplification region in these structures is about 35 μm and pump power is 0.3 W. Each of these curves has three parts. The first and the last parts correspond to regions where the signal passes through them but the pump is prevented to enter. Thus, no additional Raman gain is created. There is small amount of noise caused by structural scattering. The middle part is relevant to the region where both signal and pump are simultaneously present which results in Raman gain.

By filling the holes on the two sides of the signal path by optofluidic materials (SR-a), pump and signal group velocity are reduced causing a higher Raman gain compared with SR structure. So that, Raman gain in the output of SR-a structure is almost 3 dB greater than obtained Raman gain of SR structure.

To achieve greater Raman gain we can add an extra ring along the first one, as shown in Fig. 6. Here we assume that these two rings have no coupling with each other.

Figure 7 shows Raman gain along signal path for two-ring Raman amplifiers (SR and SR_a). Pump power is 0.3 W. As shown in this figure, for a specified

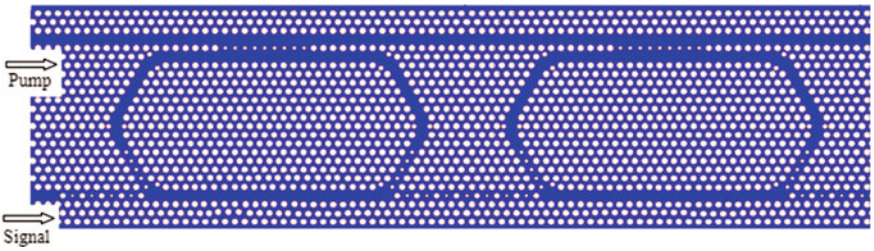


Fig. 6 Two ring-based PhC Raman amplifier structure

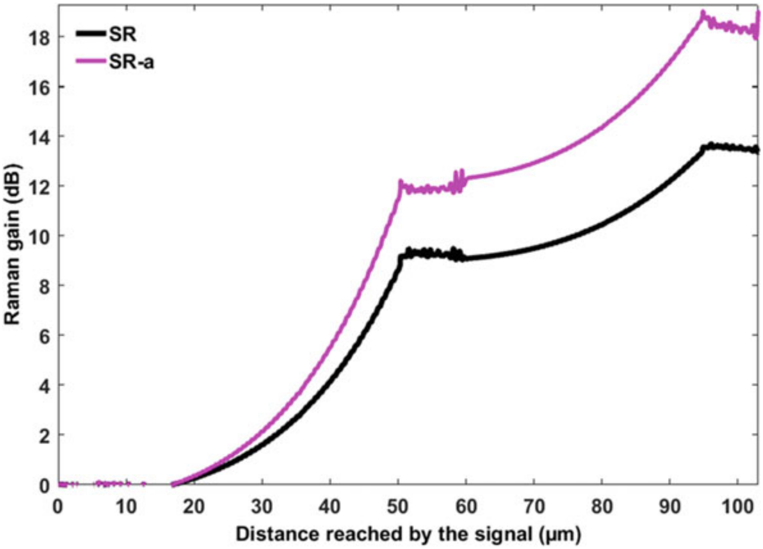


Fig. 7 Raman gain along signal path for two-ring Raman amplifiers (SR and SR-a). Pump power is 0.3 W

amount of input pump power, Raman gain has almost doubled compared to the single ring structures. As well as using optofluidic materials in SR-a structure has caused Raman gain to be about 6 dB greater than Raman gain in SR structure for input pump power of 0.3 W.

High Raman gain is one of the important parameters in choosing the structure of Raman amplifier. Another important parameter for evaluation of Raman amplifier's performance is bit rate of input pulses which corresponds to the minimum time distance between successive pulses. As the input pulse passes through the waveguide and both the initial and final parts (pump filters) creates dispersion and therefore signal pulse is broadened in time domain. Thus the minimum distance between successive pulses increases and bit rate is reduced.

Figures 8 and 9 show the time evolutions of output signal for, respectively, single-ring and 2-ring PhC-based Raman amplifier structures (SR and SR-a) for propagation of 3 successive pulses.

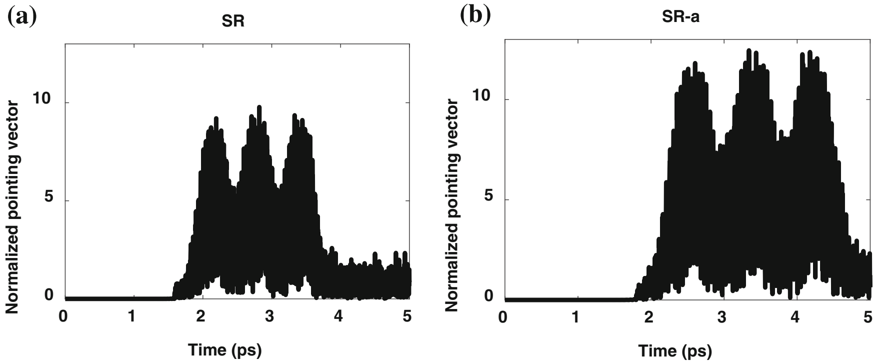


Fig. 8 Single-ring PhC-based amplifiers output after sending three successive signal pulses. Pump power is 0.3 W

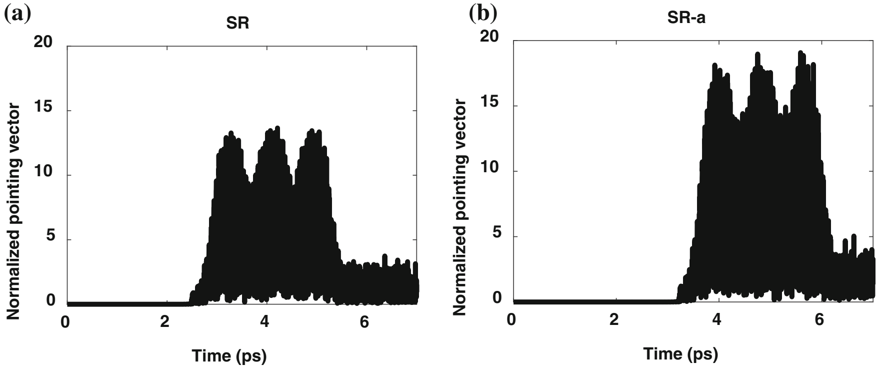


Fig. 9 2-ring PhC-based amplifiers output after sending three successive signal pulses. Pump power is 0.3 W

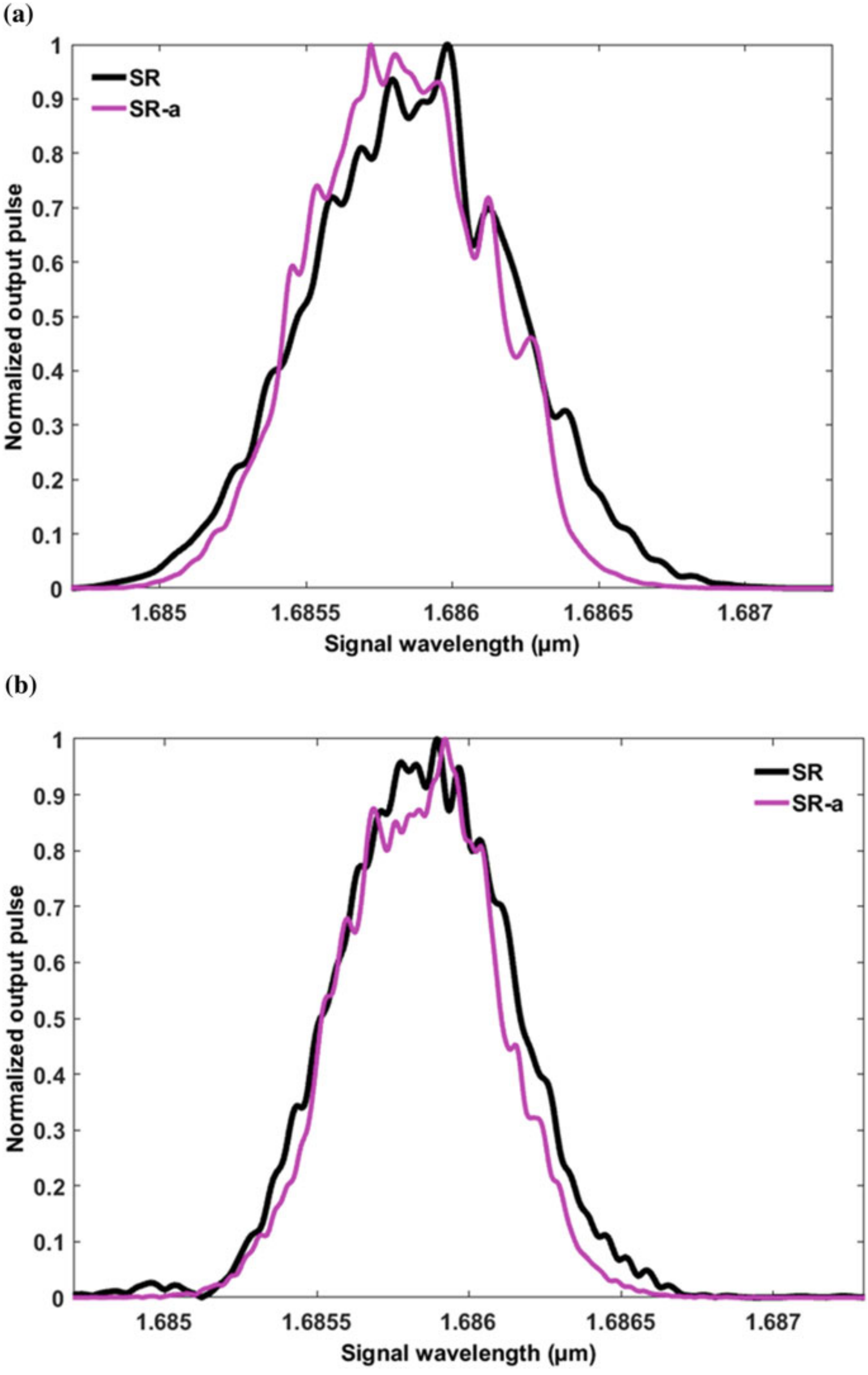


Fig. 10 Fourier transform of signal output pulses for **a** single-ring and **b** 2-ring PhC-based Raman amplifiers. Pump power is 0.3 W

Table 2 Results of Raman gain, bit rate and FWHM study in SR and SR-a Raman amplifiers

	SR (single-ring)	SR-a (single-ring)	SR (2-ring)	SR-a (2-ring)
Raman Gain(dB)	9.192	12.19	13.56	19.01
Bit rate ($\times 10^{12}$ s)	0.759	0.716	0.7077	0.6493
FWHM (nm)	0.79	0.77	0.69	0.64

The pump power is 0.3 W

Comparing the results of Figs. 8 and 9, we find that the short length path traveled by the signal in single-ring based structures provides smaller dispersion in the output. However, in two-ring Raman amplifiers, longer signal path gives larger Raman gain. On the other hand, using optofluidic materials in SR_a structure increases Raman gain significantly, but increases dispersion a little. Therefore, SR_a has a proper bit rate besides having great Raman gain.

For better understanding of the dispersion effect on the output signal pulses, Fourier transform of output pulses are shown in Fig. 10a, b for single-ring and 2-ring PhC-based Raman amplifiers, respectively. The results exhibit that dispersion reduces the pulse width in frequency domain. In single-ring structures the dispersion's full width at half maximum (FWHM) is larger than two-ring structure.

For better comparison, the obtained values of Raman gain, bit rate and FWHM of single-ring and two-ring SR and SR_a structures are summarized in Table 2. As can be seen, single-ring SR structure exhibits the lowest Raman gain and the largest bit rate (the lowest dispersion). In two-ring Raman amplifiers, longer signal path gives larger Raman gain and greater dispersion. The greatest Raman gain and smallest bit rate is seen in two-ring SR-a structure.

4 Conclusion

PhC ring-based Raman amplifier (SR) and improved PhC ring-based Raman amplifier (SR_a) have been compared in this paper. Using optofluidic materials in the holes on the sides of the signal path in the improved PhC structure (SR-a) has caused enhancement in Raman gain significantly, with same input pump power. So that, the obtained Raman gain from SR-a structure is about 2.99 dB greater than the Raman gain in SR structure, for input pump power of 0.3 W. To achieve greater Raman gain we presented the two-ring Raman amplifier structures. Longer signal path in two-ring Raman amplifier structures gives larger Raman gain. Obtained Raman gain from 2-ring SR structure and 2-ring SR-a structure are 13.56 and 19.01 dB, respectively, for input pump power of 0.3 W.

On the other hand, passing the input pulse through the waveguide, creates dispersion and therefore signal pulse is broadened in time domain. Thus, the minimum distance between successive pulses increases and bit rate is reduced.

Optofluidic materials in SR-a structure increases the dispersion and consequently deteriorates the bit rate.

But the decline of bit rate against increasing the Raman gain is not so considerable. Thus, using improved 2-ring PhC-based Raman amplifier (2-ring SR-a) is suggesting for amplifying signal.

References

1. Rong H et al (2007) Monolithic integrated ring resonator Raman silicon laser and amplifier. Proc. SPIE 6485:1–8
2. Jalali B, Raghunathan V, Shori R (2006) Prospects of silicon Mid-IR raman lasers. IEEE J Sel Top Quantum Electron 12:1618–1627
3. Claps R et al (2004) Influence of nonlinear absorption on Raman amplification in silicon waveguides. Opt Express 12:2774–2780
4. Liu A, Rong H, Paniccia M (2004) Net optical gain in a low loss silicon-on-insulator waveguide by stimulated Raman scattering. Opt Express 12:4261–4268
5. Rukhlenko ID, Premaratne M (2010) Spectral compression and group delay of optical pulses in silicon Raman amplifiers. Opt Lett 35:3138–3140
6. Kroeger F et al (2010) Saturation of the Raman amplification by self-phase modulation in silicon nanowaveguides. Appl Phys Lett 96:241102-1–241102-3
7. Claps R et al (2005) Raman amplification and lasing in SiGewaveguides. Opt Express 13:2459–2466
8. Seidfaraji A, Ahmadi V (2012) Enhanced Raman amplification by photonic crystal based waveguide structure. ICTON 1–4
9. Seyedfaraji A, Ahmadi V (2013) Improvement of Raman amplifier bandwidth by means of slow light in photonic crystal based waveguide structure. Opt Quant Electron 45:1237–1248
10. Seyedfaraji A, Ahmadi V (2010) Enhanced Raman amplification by hybrid photonic crystals. ICTON 1–4
11. Yi-Hua H, Iwamoto S, Arakawa Y (2013) Design of slow-light grating waveguides for silicon Raman amplifier. CLEO-PR 1–2
12. Krause M, Renner H, Brinkmeyer E (2010) Silicon Raman amplifiers with ring-resonator-enhanced pump power. IEEE J Sel Top Quant 16:216–225
13. Rukhlenko ID et al (2010) Optimization of Raman amplification in silicon waveguide with finite facet reflectivities. IEEE J Sel Top Quant 16:226–233
14. Monat C et al (2010) Slow light enhanced nonlinear optics in silicon photonic crystal waveguides. IEEE J Sel Top Quantum Electron 16:344–356
15. Corcoran B (2010) Optical signal processing on a silicon chip at 640 Gb/s using slow-light. Opt Express 18:7770–7781
16. McMillan JF et al (2006) Enhanced stimulated Raman scattering in slow-light photonic crystal waveguides. Opt Lett 31:1235–1237
17. Seyedfaraji A, Ahmadi V (2013) New design of ring-based Raman amplifier using optofluidic materials. Opt Eng 59(9):097103-1–097103-6
18. Seyedfaraji A, Ahmadi V (2016) Enhanced Raman amplification by conventional and hybrid photonic crystal based ring structure. Optical Quantum Electronic 48(190):1–13
19. Bakhshi S, Moravvej-Farshi MK, Ebnali-Heidari M (2011) Proposal for enhancing the transmission efficiency of photonic crystal 60° waveguide bends by means of optofluidic infiltration. Appl Opt 50:4048–4053
20. Bakhshi S, Moravvej-Farshi MK, Ebnali-Heidari M (2012) Design of an ultracompact low-power all-optical modulator by means of dispersion engineered slow light regime in a photonic crystal Mach-Zehnder interferometer. Appl Opt 51:2687–2692

21. Dekker R et al (2007) Ultrafast nonlinear all-optical processes in silicon-on-insulator waveguides. J Phys D Appl Phys 40:R249–R271
22. Keyvaninia S et al (2008) Gain variation of Raman amplifier in silicon micro-ring coupled resonator optical waveguides. Proc SPIE 6998:699818-1–699818-8
23. Kippenberg, TJA (2004) Nonlinear optics in ultra-high-Q whispering-gallery optical microcavities. PhD thesis, California Institute of Technology
24. Lin Q, Painter OJ, Agrawal GP (2007) Nonlinear optical phenomena in silicon waveguides: modeling and applications. Opt Express 15:16604–16644
25. Zheng W et al (2009) Integration of photonic crystal polarization beam splitter and waveguide bend. Opt Express 17:8657–8668
26. Xing FF et al (2005) Optimization of bandwidth in 60-photonic crystal waveguide bends. Opt Commun 248:179–184

Considering Factors Affecting the Prediction of Time Series by Improving Sine-Cosine Algorithm for Selecting the Best Samples in Neural Network Multiple Training Model

Hamid Rahimi

Abstract As stock exchange is complex and there is a high volume of information to process, no good prediction results are obtained using a simple system. Therefore, researchers have presented combined model to propose a system with lower sophistication and higher accuracy. System only uses information of one index for predicting in most prediction models but a two-level system of multi-layer perceptron neural network is proposed in this model and several indices are used to prediction and sine-cosine algorithm is also used to select the best samples after neural network training in order to train the neural network better and consequently gain better results. Results show that the proposed model is able to perform with lower prediction error compared with other models.

Keywords Sine-cosine evolutionary algorithm · Multi-layer perceptron neural network · Prediction · Time series

1 Introduction

Prediction is an inevitable part of human's life. Desire to live better and in ideal situation leads human to control these events better by looking ahead and forecasting events in advance and preparing the context to improve the future events as much as possible and/or reducing the expenses by estimating harmful events. Time series prediction is critical among other issues. Predicting future values of time series is applicable in fields such as economy, business, sales and inventory control, weather and temperature forecast, signal processing and control [1].

The requirement to determine an official model with a hypothetical probability distribution for data is a difficulty in time series analysis [2]. On the other hand,

H. Rahimi (✉)
Department of Computer, Mashhad Science and Research Branch,
Islamic Azad University, Mashhad, Iran
e-mail: Rahimihmd@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_23

initial conditions and start point are considered the most important parameters in time series responses if dynamic system is unstable; this means system reaches different responses with the least difference in start position. This position is called "sensitive to initial conditions" and expresses chaotic behavior of such dynamic systems [3–5].

Also, time series probability theory deals mostly with static time series and consequently dynamic time series analysis should transform these dynamic systems into static ones [6]. Therefore, common classic methods such as Box-Jenkins may be unsuitable and have weak points and achieve unsatisfactory responses when system has dynamic variable behavior in time. Other ways based on non-linear methods including soft computing have recently used; just like genetic and neural network algorithms and fuzzy logic which are used in analyzing and predicting complex time series and have clear and visible advantages compared with classic statistical methods [7, 8]. General advantage of soft computation methods is that unlike classic regression there is no need to determine the structure of initial model [9].

As stock exchange is sophisticated and there is a high volume of information to process, a two-level system of multi-layer perceptron neural networks is proposed in this study in which sine-cosine algorithm are used to select the best samples for neural network training so that neural network is trained better and consequently results are improved. Data analyzed in this study include stock price index, trade volume, and return rate of companies in Tehran Stock Exchange collected from March 24, 2012 to June 19, 2016 [10].

2 Sine-Cosine Optimization Algorithm

2.1 Sine-Cosine Algorithm

In sine-cosine optimization algorithm, waves search the search space to find the overall optimum point of the issue with sine and cosine behavior and waves always move toward a wave showing the best optimum point for that moment. This algorithm is used in many papers including characteristics selection [11, 12], improving multi-layer perceptron neural network training [13], and in several functional papers in contexts of handwriting recognition [14], and economic and thermal optimization in plants [15].

Sine-cosine algorithm is not a method based on elitism population and it is based on a trajectory which has shown high extraction power and accurate convergence and this leads to an accurate optimum point even in high dimensions of functions. It has both discovery and extraction in optimization and achieves an overall optimum for the problem. The best response always shows the destination for search waves in this algorithm. Therefore, search waves are not distracted from the main problem and also the swinging behavior in this algorithm allows it to search the search space

```
Start: generating search waves
Performing the following steps until it reaches the final iteration:
    Assessment: each wave is assessed by problem function.
    Update: the best obtained response is updated.
    Update: $r_1$, $r_2$, $r_3$, and $r_4$ are updated.
        $r_1 = a - t\frac{a}{T}$
        a is a constant variable, t is current iteration and T is the final iteration
    Update: new coordinates of waves are updated.
        $x_i^{t+1} = \begin{cases} x_i^t + r_1 \times \sin(r_2) \times |r_3 p_i^t - x_i^t| & r_4 < 0.5 \\ x_i^t + r_1 \times \cos(r_2) \times |r_3 p_i^t - x_i^t| & r_4 \geq 0.5 \end{cases}$
End of algorithm final iteration and return of the best obtained response as the optimum of the problem.
```

Fig. 1 Pseudo-code of sine-cosine algorithm [16]

around the optimum point of the problem well and has good accuracy for obtaining the optimum.

Steps of this algorithm are as follows [16]:

This algorithm of variables in the operator is as follows: (Fig. 1)

x_i^{t+1} new coordinate of search wave; x_i^t previous coordinate of search wave; p_i^t coordinate of the best obtained response which is considered as destination; and if r_1 is less than 1, movement toward destination is started and if it is larger than 1, the movement will be far from the destination. R_2 is used for modeling swinging movement and varies between 0° and 306° and r_3 is a weight for the destination and if it is considered larger than 1, the next step will be a larger movement toward the destination and if it is less than 1, movement will have a smaller step and r_4 is a random variable between 0 and 1 to switch between sine and cosine movement.

2.2 *Proposed Algorithm*

Proposed model includes a two-layer structure. First stage is consisted of multi-layer perceptron neural networks as predictor which each one is trained to predict a certain indicator of different data. In other words, each basic categorical prediction network predicts a certain indicator independent from other systems. There is a synthesizer multi-layer perceptron neural network in the second stage trained using best sample selection mechanism and it finally provides the model output (Fig. 2).

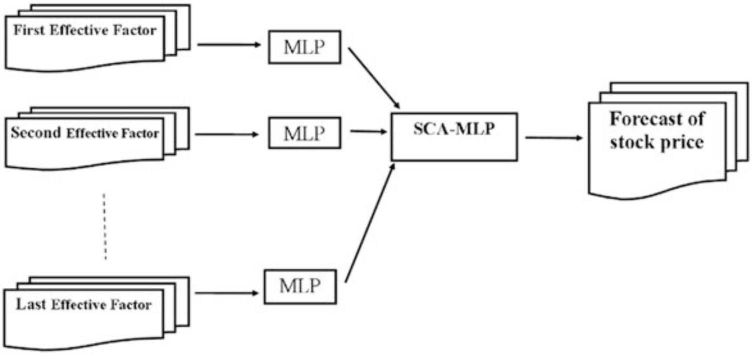


Fig. 2 Proposed model

General importance and characteristic of this model include using different features and data to predict final behavior of given data. The role of synthesizer layer in this structure is prediction through information and changing data and other indicators of basic prediction networks to predict more accurately and efficiently.

First layer of proposed model includes a window of data as input and data after the window is considered as the output of each network and each basic network is trained separately.

In other words, input of each neural network at first stage of training is data and a certain indicator at time window of t and before it and its output is the prediction of that indicator at t + 1. Therefore, first stage includes neural network systems which each one predict a certain indicator.

At the second layer, neural network is trained by best samples selected using sine-cosine optimization method and when it is going to be used data predicted in first layer are fed into the second stage to forecast final financial indicator.

Neural network of second layer is trained by best samples. Sine-cosine optimization method select samples using fitness obtained from neural network and these selected samples are transferred to neural networks to train and they produce their own output neural network and fitness of selected samples is determined by the difference between the outputs of neural network and real outputs under fitness function (Fig. 3).

3 Results

Data of stock price index, trading volume and Tehran's stock exchange return is used in this paper collected from March 24, 2012 to June 19, 2016.

Next, three important data sets of Tehran's stock exchange used as substantial data for final forecast of Tehran's stock price index are examined:

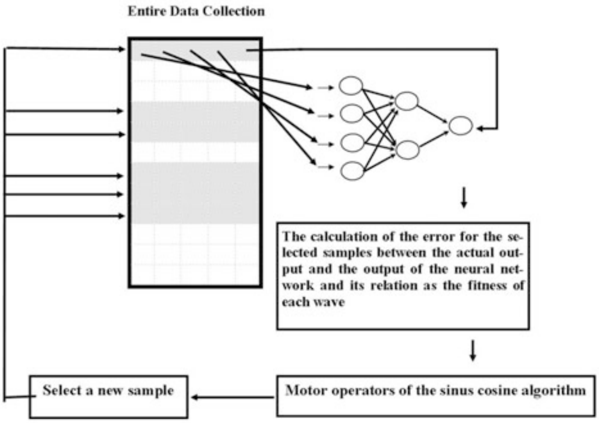

Entire Data Collection
The calculation of the error for the selected samples between the actual output and the output of the neural network and its relation as the fitness of each wave
Select a new sample
Motor operators of the sinus cosine algorithm

Fig. 3 Structure of second stage of the proposed model

(1) Price index of Tehran’s tock:

Tehran’s stock exchange attempted to calculate and distribute its price index under the title of “Topix”. This index indicates price changes in whole market.

Tehran’s stock price index includes the share of all companies accepted in the exchange and if the symbol of a company is closed or not traded for a while, the last traded price is considered. As the formula shows, number of shares distributed by companies is a weighting criterion for the given index. Price index of the exchange is an indicator which shows the general level of stock price in companies accepted in the exchange. Stock price is the base of decision making for market agents in tradable assets market. Investors of this market look at the latest information about price situation and analysis of future price variation predictions as price analysis is performed more easily for all but there are other affecting factors such as trading volume and interest rate return.

(2) Trading volume:

Trading volume is the number of purchase and sales performed in a time range such as one working day. Trading volume can be used as another tool for determining what events happen in the market, especially events of trends, as even credible and reliable trends may sometimes fail [17]. Investment liquidity is one of the main criteria in short-term investment. Undoubtedly, number of trading each share and number of traded share of a company can be an indication of share liquidity.

Share change is performed through trading in Tehran’s exchange. Share price of the company will change if trade volume leads share price of a company towards increase or reduction directionally.

Trade volume is important for Tehran's stock exchange, brokers, and tax officers as wage and all benefits of these organizations is determined based on tariffs of trade volume and they all benefit from trade volume rather than increasing and decreasing share price. Buyers and sellers can also obtain effective results from trade volume. They can review offer situation and demand on each share through trade volume and use these tools in their decision making. Therefore, trade volume can also be used as an effective factor in prediction. McMillan suggests trade volume for predicting share range which shows trade volume can contain good information as price and they also show how price and volume are related [17].

(3) Return rate:

Return rate (interest) is the interest or income obtained from investment which is expressed in percent. Return rate is one the effective factors in assessing share price changes [18].

In researches performed on Tehran's stock exchange, mostly the relation of internal information of companies and their financial status with company's share price is studied and the market itself and its internal indicators are less considered as factors affecting company's share price.

Also, effects of releasing company's information are studied in most researches while the purpose of this study is examining the release of market information and trader's usage of market.

Prediction error measurement criteria will be used to measure the conformity of one prediction with a pattern of time series data. Prediction error is e = y − ŷ if y and ŷ show real and predicted value of variable in time t. Therefore, prediction measurement criteria for a time period and n predicted values are as follows:

$$MSE = \frac{\sum_{i=1}^{n} (e_i)^2}{n}$$

where y is real value and $\widehat{y}$ is the value predicted by the model. In fact, the closer model estimation is to the reality, the less error exists in the prediction. So, mean square error (MSE) is used which is considered an acceptable criteria by researchers.

Results obtained from this paper are resulted from three different prediction model with multi-layer perceptron neural network:

First model: direct prediction of exchange price index without considering factors affecting it.

Data of exchange price index of Tehran's stock exchange are considered from March 24, 2012 to June 19, 2016 which 1565 cases were considered. 1173 cases are used for network training and 392 cases are considered for network test. Numerical value of vertical graph is in scale of 1000 and shows price index number and horizontal graph is based on day.

Windows with the length of 50 data for network input are considered to predict these sets using perceptron neural network and next data is adjusted for network output which this window move as much as one data for all training data until neural network is trained and then test data are fed similarly (window with the length of 50 for the input) into the network for prediction.

Selecting the best architecture for neural network is performed by true trial and error and adjusting neurons of hidden layer of neural network and network test during training and supervising obtained error for test and training data is used to avoid over-training of the network. The best numeric result obtained from errors of different architectures of neural networks is presented in Tables 1 and 2 (Fig. 4).

Second model: predicting price index in two levels using affective factors and without sample selection.

In this model, price index prediction is performed by considering affective factors, trading volume, and return rate so that neural network training at first level is designed for time series of trading volume and return rate and at second level, all samples for neural network training to predict price index are considered.

Selecting the best neural network is performed through trial and error and by adjusting neurons of hidden layer in neural network and network test during training and error monitoring of test and training data is used to train networks and prevent over-training of the network. Table 2 shows the best numerical result (Fig. 5; Tables 3, 4, 5 and 6).

Third model: predicting price index in two-level form considering affecting factors and sample selection.

In this model, price index prediction is performed by considering factors affecting trading volume and return rate so that at first level, neural network training is designed for time series of trading volume and return rate and samples selected

Table 1 Comparing results obtained from first model per different neuron

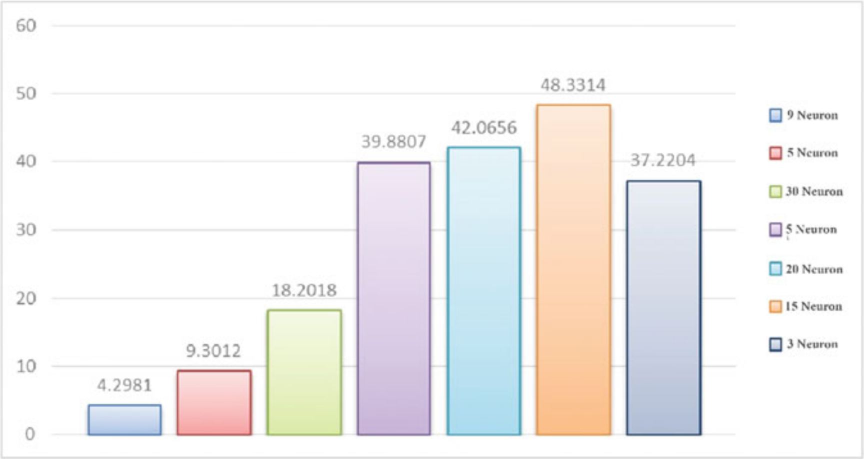


Table 2 Prediction results obtained from first model for best model

First model	MSE
Best network architecture with 9 neurons in one hidden layer	4.2981

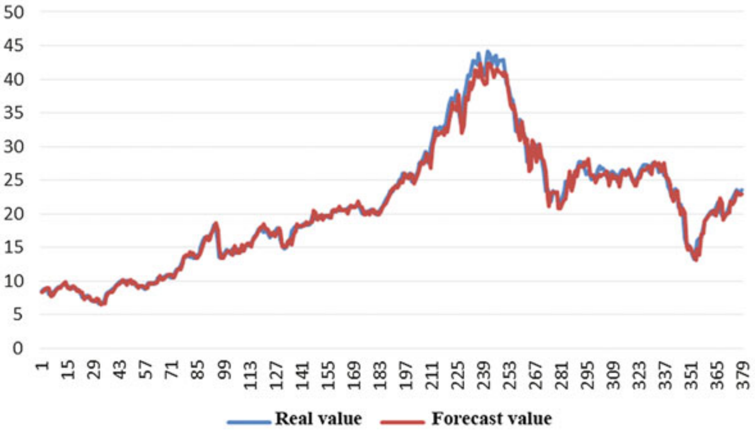


Fig. 4 Comparing time series predicted by neural network having 9 neurons in a hidden layer in first model and its real value

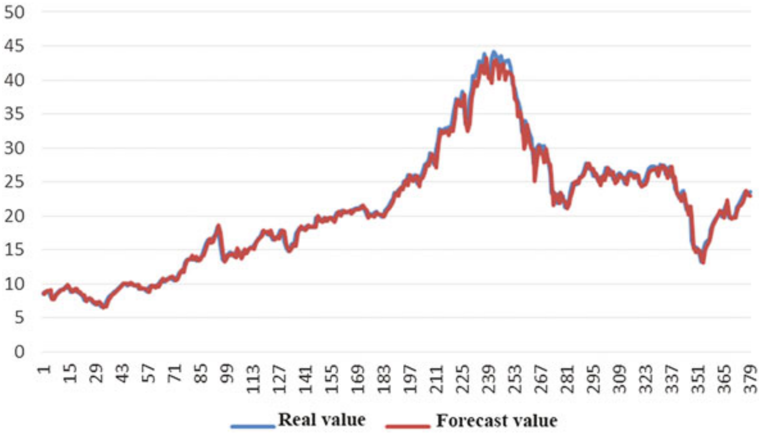


Fig. 5 Comparing graphs of time series predicted with neural network and its real value

by sine-cosine algorithm are used for training neural network at second level to predict price index.

In this algorithm, each sine-cosine wave is defined as an array including a number of samples subscript to select samples:

Table 3 Comparing results obtained from trading volume in second model per different number of neurons

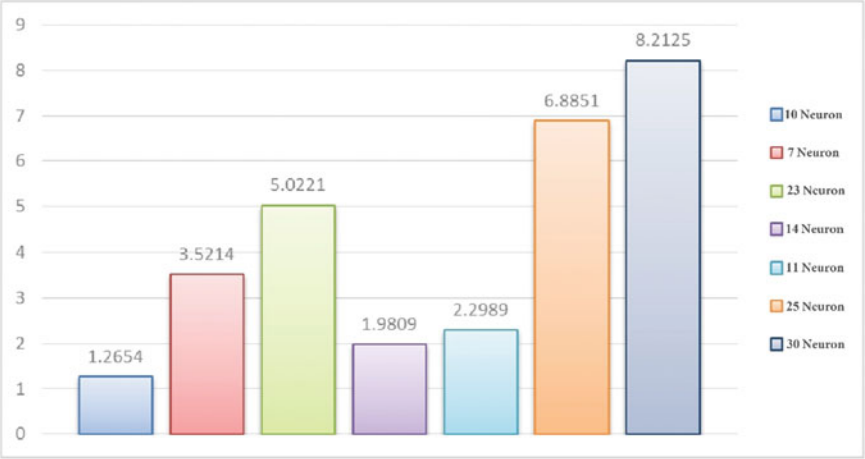


Table 4 Comparing results obtained from return rate factor in second model per different number of neurons

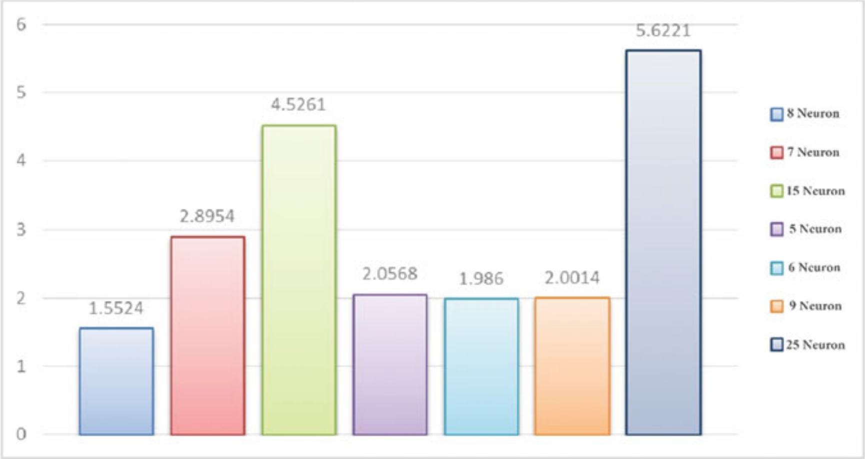


Table 5 Comparing results of prediction obtained from price index factor in second model per different number of neuron

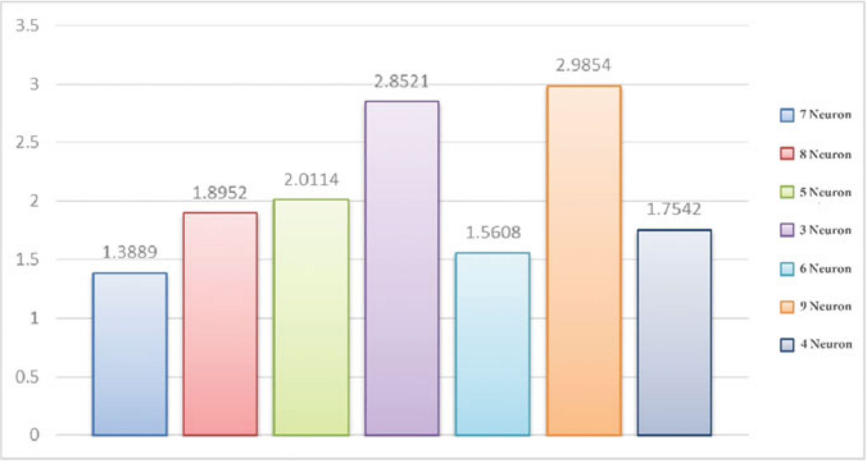


Table 6 Results of prediction obtained from second model

Second model	MSE
Best network architecture with 10 neurons in one hidden layer for trading volume factor (first level)	1.2654
Best network architecture with 8 neurons in one hidden layer for return rate factor (first level)	1.5524
Best network architecture with 7 neurons in one hidden layer for price index factor (second level)	1.3889

$$\mathrm{Wave} = [\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_\mathrm{n}]$$

X_1 is the location of first sample participating in network training and X_n is the last sample.

Components of each wave can include the numbers 1 and/or 0. Number 1 means that sample number corresponding the component is selected and number 0 means lack of selection. Fitness function of the problems is as follows in which error between network output and real output shows the fitness of samples selected for network training: (Fig. 6; Tables 7 and 8)

$$\mathbf{MIN\ Fitness} = \frac{1}{\boldsymbol{n}} \sum_{i=1}^{\boldsymbol{n}} (\boldsymbol{Y}_{\boldsymbol{net}} - \boldsymbol{Y}_{\boldsymbol{act}})^2$$

Comparing results of three stimulated models:

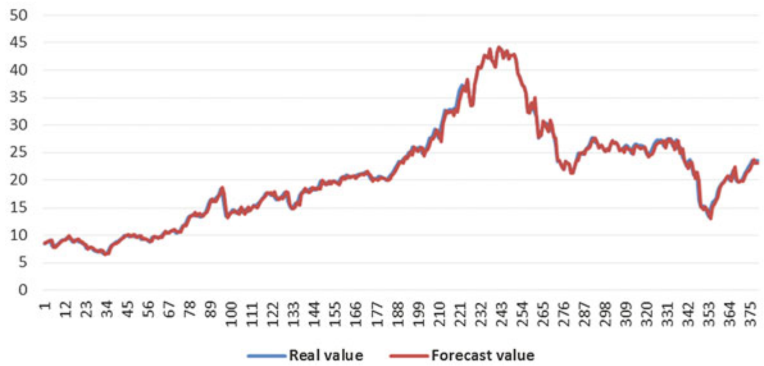


Fig. 6 Graphical comparing of times series predicted by neural network and its real value

Table 7 Comparing prediction results of price index factor in third model per different number of neurons

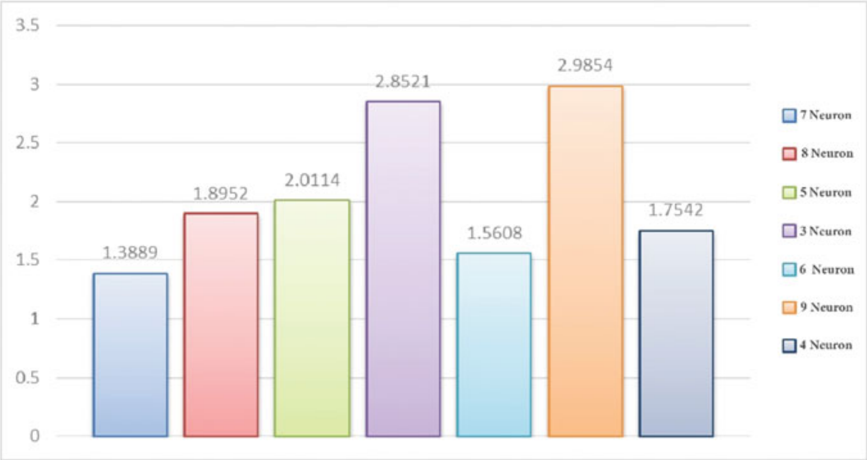


Table 8 Prediction results obtained from the third model

Third model	MSE
Best network architecture with 10 neurons in one hidden layer for trading volume factor (first level)	1.2654
Best network architecture with 8 neurons in one hidden layer for return rate factor (first level)	1.5524
Best network architecture with 7 neurons in one hidden layer for price index factor (second level)	0.2719

Results of three stimulated models used for predicting stock price index of a company in Tehran's stock exchange show that the proposed model (third model) has predicted with lower error compared with other models (Table 9).

Tables 9 and 10 shows errors of three applied model in this study. As it can be seen, results of proposed model are better than two other models.

Next, results obtained from proposed model are compared with results of previous works including prediction of Tehran's stock exchange index by combining main component analysis, support vector regression and particle aggregative movement [1] which its data are similar to the time range selected for the proposed model (2012–2017). Results of comparing the proposed model with the model of article [19] is presented in Table 11.

Table 11 shows errors obtained in article [19] and from the proposed method. As it can be seen, results of proposed method is better than methods of main component analysis, support vector regression, and particle aggregative movement and this improvement is about 45%.

Table 9 Comparing results obtained from three models

Models	MSE
First model	4.2981
Second model	1.3889
Third model	0.2719

Table 10 Graphs comparing prediction results obtained from three models

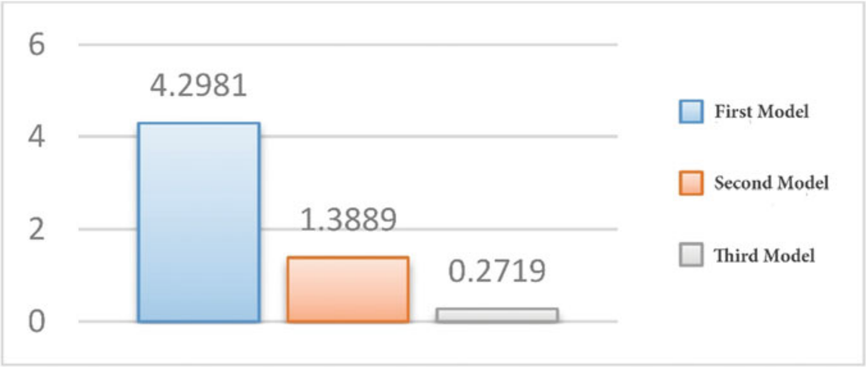


Table 11 Comparing prediction results obtained from the proposed model and other works

Methods	MSE
First model: support vector regression and particle aggregative movement [19]	0.753
Second model: analysis method of main components, support vector regression, and particle aggregative movement [19]	0.486
Third model (proposed model)	0.271

4 Conclusion

A two-level system of multi-layer perceptron neural networks is proposed in this study as stock exchange is complex and there is high volume of information to process. Sine-cosine algorithm is used to select better samples for neural network training so that neural network is trained better and results will be improved.

Studied data of this paper include exchange price index, trading volume, and return rate of companies in Tehran's stock exchange collected from March 24, 2012 to June 19, 2016. At first level of proposed model, time series of trading volume and return rate is predicted and exchange price index using the best samples is predicted at second level.

Three models were used to compare results; first model: direct prediction of exchange price index without considering factors affecting it; second model: predicting price index in two levels using affective factors and without sample selection; and third model: predicting price index in two-level form considering affecting factors and sample selection. In first model, trial and error and adjusting neurons of hidden layer of neural network were used to select the best architecture of the neural network and network test during training and error monitoring for test and training data was used to train networks and prevent over-training of the network. In the second model, price index prediction was performed using factors affecting trading volume and return rate so that neural network training at first level was designed for time series of trading volume and return rate and at second level, all samples of neural network training were considered for price index prediction. In third model, factors affecting trading volume and return rate were considered to predict price index so that neural network training at first level was designed for time series of trading volume and return rate and samples selected by sine-cosine algorithm were used at second level to train neural network for predicting price index.

Numerical results of company's data from Tehran's stock exchange including three data sets of exchange price index, trading volume, and return rate are presented using three models: first model: direct prediction of exchange price index without considering factors affecting it; second model: predicting price index in two levels using affective factors and without sample selection; and third model: predicting price index in two-level form considering affecting factors and sample selection. Proposed method is able to present better prediction results with lower error compared with other methods.

References

1. Wang LX (1997) A course in fuzzy systems and fuzzy control. Inc publisher, Bernard Goodwin, pp 151–179
2. Hansen JV, McDonald JB, Nelson RD (1999) Time series prediction with genetic-algorithm designed neural networks: an empirical comparison with modern statistical models. Comput Intell 15(3)
3. Brock WA, Hsieh DA, LeBaron B (1991) Nonlinear dynamics, chaos and instability. MIT Press, Cambridge, MA
4. Peters E (1994) Fractal market analysis: applying chaos theory to investment and economics. Wiley, New York
5. Lai KK, Yu L, Wang S, Wei H (2006) A novel nonlinear neural network ensemble model for financial time series forecasting. Int Conference Comput Sci (1)790–793
6. Chatfield C (2002) The analysis of time series introduction. Ferdusi University publisher pp 1–111
7. Rojas I, Pomares H (2004) Soft-computing techniques for time series forecasting. ESANN'2004 proceedings—european symposium on artificial neural networks bruges (Belgium), 28–30 April 2004, d-side public. ISBN 2-930307-04-8, pp 93–102
8. Maddala GS (1996) Introduction to econometrics. Prentice-Hall, Englewood Cliffs, NJ
9. Zheng Z (1996) Automated mathematical modeling for financial time series prediction using fuzzy logic, dynamical system theory and fractal theory. In: Proceedings IEEE international conference on computational intelligence for financial engineering (CIFEr), New York, pp 120–126
10. http://new.tse.ir/
11. Sindhu R, Ngadiran R, Yacob YM (2017) Sine–cosine algorithm for feature selection with elitism strategy and new updating mechanism. Neural Computing Appli
12. Hafez AI, Zawbaa HM, Emary E, Hassanien AE (2016) Sine cosine optimization algorithm for feature selection. International symposium on innovations in intelligent systems and applications (INISTA)
13. Sahlol AT, Ewees AA, Hemdan AM, Hassanien AE (2106) Training feed forward neural networks using Sine-cosine algorithm to improve the prediction of liver enzymes on fish farmed on nano-selenite Sign In or Purchase. In: 12th international computer engineering conference (ICENCO)
14. Elfattah MA, Abuelenin S Hassanien AE, Pan J-S (2016) Handwritten arabic manuscript image Binarization using sine cosine optimization algorithm. In: International conference on genetic and evolutionary computing
15. Turgut OE (2016) Thermal and economical optimization of a shell and tube evaporator using hybrid backtracking search—sine–cosine algorithm. Arabian J Sci Eng
16. Mirjalili S (2016) SCA: a sine cosine algorithm for solving optimization problems. Knowledge-Based Systems
17. Grinblatt M, Moskowitz TJ (2004) Predicting stock price movements from past returns: the role of consistency and tax-loss selling, Journal of Financial Economics 71:541–579
18. McMillan DG (2007) Non-linear forecasting of stock returns: does volume help? Int J Forecasting 23(1):115–126
19. Raee R, Nik Ahd A, Habibi M (2017) Predicting price index of Tehran's stock exchange by combining methods of main component analysis, support vector regression, and particle aggregative movement. Financial Management Strategy 1–23:4.4

Advantages of Using Cloud Computing in Software Architecture

Alireza Mohseni and Mehrpooya Ahmadalinejad

Abstract In recent years, various software architectures have emerged. Service oriented architecture and model driven architecture are two examples of this architecture. Service-oriented architecture (SOA) has disadvantages, such as: the need for further development, precise design, and the creation of a service infrastructure. According to Exforsys research, SOA is used in some cases, such as single-user applications that do not have distributed properties, a uniform connection of the asynchronous type is required, but there is no need for interconnection between the components; programs that are short-lived and in A short run is run, programs whose function is more dependent on the GUI. Model-based architecture also has some disadvantages, such as: ignoring some of the aspects and features of the system, including quality requirements, because of the model's model, the inadequacy of the small software projects and the complexity of the software to The reason for the lack of a comprehensive definition of requirements. On the other hand, cloud computing has many benefits that access is based on the demand and the widespread use of networks and shared resources of these benefits. In this article, we are trying to provide a newer and more advanced architecture with the use of cloud computing and the development of software architecture. In cloud computing, the user designs his own cloud for his cloud and receives cloud services from the cloud without even coding a line or designing his own software at a great cost. Cloud architecture is very effective in designing small and large software, and even the user can add objects as services to the reference cloud. Finally, other benefits of the new architecture are expressed.

Keywords Software architecture · Cloud computing · Model driven architectureword · Service-Oriented architecture

A. Mohseni (✉)
Graduated from Islamic Azad University, Qaem Shahr Branch, Qaem Shahr, Iran
e-mail: mohsenia94@yahoo.com

M. Ahmadalinejad
Graduated from Islamic Azad University, Qazvin Branch, Qazvin, Iran
e-mail: mehrpooya@qiau.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_24

1 Introduction

Since the introduction of the software so far, different architectures have been designed and implemented. The above architectures on the one hand, derive from the features and nature of hardware in their time, and on the other, represent the type and attitude of users' expectations. Keep in mind that the software is a dynamic body and should at any time adapt itself to the huge needs and expectations of new users. Because the application is a human extract for the purpose of actualizing on the hardware platform over time. Obviously, from the past, the spectrum of human demands has changed and will change, and hardware will undergo a major change. In this regard, it is necessary that the software, with full respect for the principle of flexibility, accepts all developments from the past and can perform its mission at any time. Accordingly, from the past, different architectures have been proposed for the design and implementation of the software. Each architecture has its own unique features and attributes, and software that relying on each of the above architectures will inherit its properties from the architecture it employs.

Software architectures and their advantages and disadvantages include: MainFrame architecture, File Server architecture, Client/server architecture, Two-Tier Architecture, Three-Tier Multilayer Tier Architecture.

1.1 Software Architectures and Their Advantages and Disadvantages

1.1.1 MainFrame Architecture

Features of this architecture include: In the 1960s–1970s, the architecture was considered to be serious, The host computer is responsible for all processes, Terminal users are able to communicate with the host system, Terminals are not smart and are limited to a keyboard and display, Pressing the keyboard keys is the only thing that will mean the connection, between the users (terminals) and the original system, The data and logic of the program are stored on the same system (Host).

Advantages of this architecture are: Security in this type of architecture is very high. Due to the focus of data and logic, centralized management and its implementation will be easy.

Disadvantages of this architecture are: The cost of providing, renting and supporting these types of systems is very high. The program (logic), along with the corresponding data, is deployed in a location and from the same processing environment.

File Server Architecture

Features of this architecture include: Rotation relative to the MainFrame architecture. Multiple servers or servers are used centrally. Various resources are shared, such as a printer or storage space (hard drive). The server will send the required files to the users and will be stored by shared resources. The work requested by the user (logic and data) will be executed on the user's system. The logic of the program will run on the user's system (client). The data will be deployed on the client.

Advantages of this architecture are: To use this architecture, you will not need to spend a lot of money. In terms of utilizing resources, it has a good flexibility.

Disadvantages of this architecture are: All logic of the program will be implemented on the client. Existence of specific constraints such as memory size or type of client processor makes it difficult to use the program. Improving the performance of the program or implementing the desired improvements is always one of the serious challenges.

Client Server Architecture

Features of this architecture include: In the above architecture, service providers and clients with different characteristics are used. The principle of division of labor follows, and the server will process heavy operations with high processing and client-style operations. Two different parts of a program work together to carry out the operation. The client will display a request for assistance in the processing of the operation by sending the request and the server, responding to the request. Platform and server and client server operating systems can be different. The operation that an application executes between the server and the client is divided.

Advantages of this architecture are: Proper use of existing hardware potentials in accordance with the principle of division of operations. Optimizing the use and utilization of shared resources. Optimize the ability of users to perform different activities.

Disadvantages of this architecture are: The lack of facilities to encapsulate policy software policies. Reduced program performance while simultaneously increasing the number of users. Improving the performance of the program or implementing the desired improvements is always one of the serious challenges.

Client Server Architecture: Two Tier

Features of this architecture include: Similar to the Client Server model in the previous section. The model uses a server and a client on the network. The above model will consist of three sections that will be located in two layers of the server and the client. UI sections, process management, and database management are three parts of the above model. Program logic is distributed between two physical locations.

Advantages of this architecture are: he most appropriate method for distributed processing in a network with up to 100 users. Ease of implementation. Direct attribution of the user interface with data sources.

Disadvantages of this architecture are: The lack of possibilities for encapsulating policy software policies. Reduced program performance while simultaneously increasing the number of users. The lack of flexibility and flexibility of transferring an application from a server to another server without making any major changes.

1.1.2 Multi Tier Architecture, Three Tier

Features of this architecture include: The above model was introduced in 1990. In this model, another middle tier is used between the client (user interface) and the database server. The middle layer includes a set of tools for accessing system resources, regardless of the type of platform format. The middle layer will be responsible for managing the processes. The middle layer is responsible for decomposing or combining the results from data sources such as databases. UI sections, process management, and database management are three parts of the above model. The middle layer can be divided into two or more parts with distinct functions (multi-tier). The business logic layer can be placed on multiple servers. The above model is an ideal option for software implementation on the Internet.

Advantages of this architecture are: Increase efficiency, flexibility, reusability and support power. Improved performance as the number of users increases. Hide existing complexities according to the nature of distributed processes from the perspective of users. Provide the necessary resources for programmers to design and implement software with a similar approach. Provide the necessary resources for programmers to follow the same methods for accessing data. Use design and programming patterns in faster implementation and development.

1.1.3 Service Oriented Architecture

Features of this architecture include: Use of technology-independent standards and agreement to provide software components under the service template. Introduces a specific and agreed method for defining and communicating between software components. Individual software components can be used to make other software. Reinforcing the approach of assembling predefined components for software development instead of developing and implementing them. It can connect to external software applications, like its internal ones.

Service-oriented architecture enables you to quickly change your systems. Easy integration with both domestic and foreign partners. reuse. Support for products with a short life span. Improve return on capital. Direct Implication of Professional Processes to Information Technology. Gradual development and implementation. Easy flexibility and easy change from one service provider to another. Identifying services helps the organization better understand core processes. Reduce the cost of

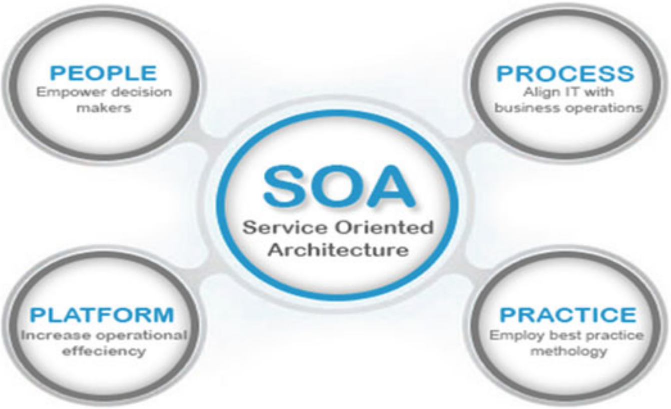

PEOPLE
Empower decision makers
PROCESS
Align IT with business operations
SOA
Service Oriented Architecture
PLATFORM
Increase operational effeciency
PRACTICE
Employ best practice methology

Fig. 1 Service oriented architecture [1]

developing and maintaining systems. The information architecture for the business domain is visible (visible). Applying standards will ensure interoperability. Service-oriented architecture is platform independent Non-dependence of systems and architecture to physical location. The explicit definition of the responsibility for each service results in the accountability of most parts of the organization.

Disadvantages of this architecture are: Service-oriented architecture is not always the best choice for developing software systems, because SOA application requires more development time, accurate design and creation of service infrastructure, and this itself requires a high cost. Unicode programs that do not have distributed properties. Programs that require asynchronous one-way communication, but loose coupling between components is not much needed. Programs that will run for a short time and in a short time. Programs whose performance is more dependent on the graphical user interface (Fig. 1).

Model driven Architecture

Features of this architecture include: In this architecture, models have the task of guiding and leading the flow of understanding, analysis, design, construction, deployment, operation, maintenance and evolution. The purpose of software developers is to produce high-quality code, which in this approach can produce many automation and automation software development steps and products, leading to higher product and code quality.

Advantages of this architecture are: More time can be devoted to analyzing and designing models. The time required to code is reduced due to the availability of the code auto-generate tool. The quality of the developed system improves.

Disadvantages of this architecture are: The driven model architecture has devastating effects on system development. First, there is no guarantee that more

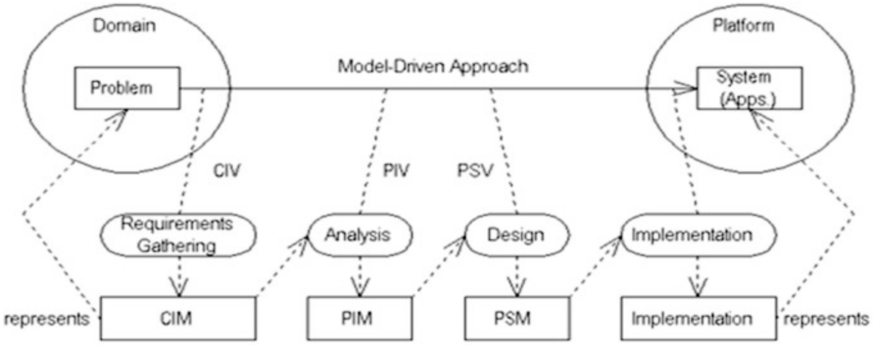


Fig. 2 Model driven approach [2]

abstraction results in better software, and in fact the results of programming psychology in particular indicate that abstraction can have a negative effect, because thinking is abstract and difficult, and people's desire for objective examples More abstract conceptualization. Second, model driven engineering involves affiliated activities that have positive and negative effects. For example, code generation from a model at first glance seems to have a positive effect on it, but a lot of effort to develop models that enable code generation and the need for manual changes to these codes seems to have a negative effect on Thirdly, there are several varieties of engineered modeling driven, and so developers will be in trouble choosing the right kind of business for their business [2] (Fig. 2).

1.2 Software Life Cycle

The stages of the software life cycle are: UR Step—User Requirements. SR Step—Software Requirements. AD stage—Architectural Design. DD stage—Detailed Design. TR stage—Transfer of the software. Step OM—Operations and Maintenance [3] (Fig. 3).

1.3 What Is Cloud Computing?

Cloud computing is a model for easy access and network sharing to a set of configurable computing resources (such as networks, servers, storage, applications, and services) that can be quickly deployed with the least effort or effort or service provider interference, or Free (abandoned). This cloud-based model supports availability and incorporates five basic features, three shapes and four forms of preparation.

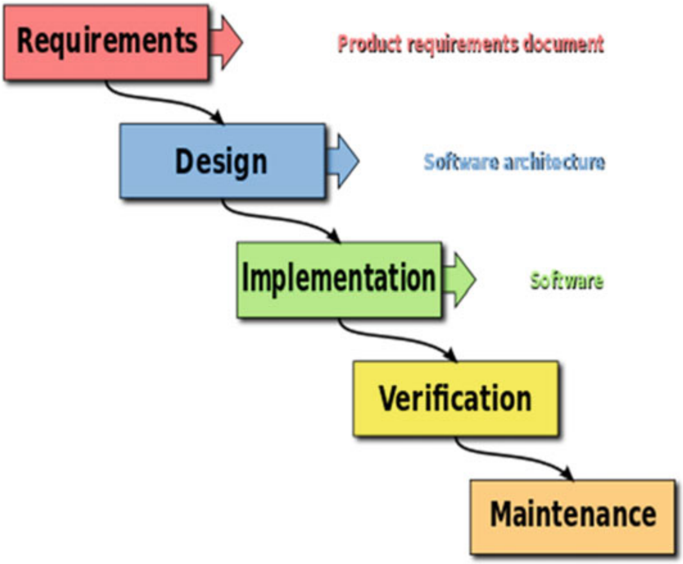


Fig. 3 Software life cycle [3]

Cloud computing features:

On dimand: The client can obtain one-way computing features such as server and network storage as needed from any provider automatically and without human intervention. Network-wide access. Facilities on the network are available and can be achieved by standardized mechanisms, supporting mechanisms that are used for heterogeneous platforms for weak and strong clients (such as mobile phones, laptops, and PDAs). Uniform Synthesis to the Sources of Resource: The compiler resources are provided to serve the needs of all customers through the use of the multimodal model, carried out by various physical or virtual resources that are dynamically retrieved and retrieved according to the client's request. The customer usually does not have control or knowledge about the exact location of the provided resources, but it may be able to determine the location at higher abstract levels (such as country, province, or data center). For example, resources include storage space, processing power, memory, network bandwidth and virtual machines. Fast Flexibility (In-Place): Flexibility can be achieved quickly and flexibly to be expanded quickly (scale-free) or on-site to quickly reach a smaller scale. From the customer's point of view, the facilities available to access are often unlimited and can be purchased at any time and at any time. Measurement Services: Cloud-based systems control and optimize resources by using the ability to measure at the level of abstraction appropriate to the type of service (such as storage space, processing power, bandwidth, and active users' number). The use of resources can be monitored, controlled and reported transparently for both the customer and the provider.

Providing service in cloud computing:
Cloud Computing as a Service (SaaS). What is provided to the client is a cloud-based provider that runs on the cloud infrastructure and is available on various client devices through an interface to a weak client such as a web browser (such as webmail). The client does not manage or control the cloud infrastructure, network, servers, operating systems, underlying storage space, or even software, with the exception of the limited configurations of the program at the user level. Cloud Platform as Service (PaaS). The client may place his application on the cloud infrastructure. This program is made using the programming languages and tools supported by the provider (such as Java, Python, and .NET). The client does not manage or control the cloud infrastructure, network, servers, or storage space, but it is based on an application and possibly a host of host configurations.

Cloud Infrastructure as Service (IaaS). The option provided to the customer is the processing power, storage space, networks and other computing base resources, so that the client can put and run its own software, which can include operating systems and applications. The client does not manage or control the underlying cloud infrastructure, but it controls on operating systems, storage, applications, and possibly the selection of networking components (such as firewalls, load balancing).
Preparation forms:
Private cloud. The cloud infrastructure only works for an organization and may be managed by the organization itself or another company, it can also fit inside or outside the organization. Cloud community. Cloud infrastructure is shared between multiple organizations and supports a specific group that shares common tasks (such as missions, security needs, policies, and legal considerations). This cloud can be managed by these organizations or another company, it can also be embedded inside or outside the organization. Public cloud. Cloud infrastructure is available for the public or for a large number of customers, and is the owner of the organization that sells these cloud services. Hybrid cloud. Cloud infrastructure is a mix of two or more clouds (private, group or public), each of which holds its own unique features, but are connected by standardized or exclusive technology that transmits data and applications [4] (Fig. 4).

1.4 Use Cloud Computing in Software Architecture

Considering the features of cloud computing, we use step-by-step to use these features in cloud computing: Step One: Create a cloud of requirements for the user. This cloud contains all the documentation and other requirements extraction tools based on the type of entities. Cloud requirements have all the entities and attributes and their relationship. By creating a cloud of requirements for a model, the user extracts his or her own entities along with the attributes and relationships between

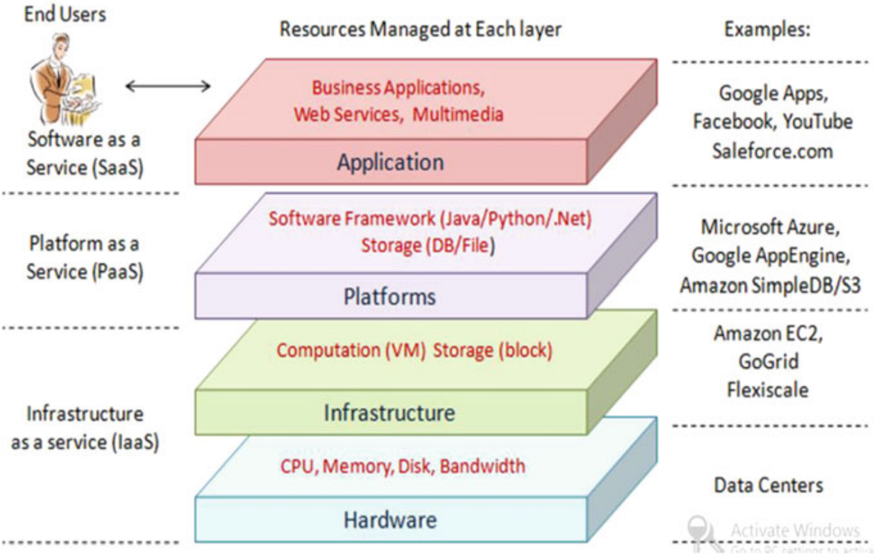


Fig. 4 Cloud computing layer architecture [4]

them, and, using the software environment that the cloud requirements, it creates its own cloud requirements. And in the end it’s possible to share all the entities inside this cloud for other users.

Step Two: In this phase, the cloud design is made up of a variety of models and services to prepare the product. In other words, using cloud requirements facilities, the user passes requirements from cloud requirements to cloud design, and the cloud design includes models that models user software to the user’s choice.

Step Three: At this point, by creating a cloud-based implementation, the user uses the various features of this cloud to code the cloud-based design.

Step Four: At this stage, the cloud testing of deployment and delivery of the system will be performed after having passed the appropriate test and is approved for publication, sale or any distribution for the final work environment.

Step Five: In this step, by creating a cloud maintenance user after the final acceptance of the software, it corrects it in order to correct the revealed errors during the previous steps or because of new needs that occur. This action is called “maintenance”. This architecture is shown in the following figure: (Fig. 5).

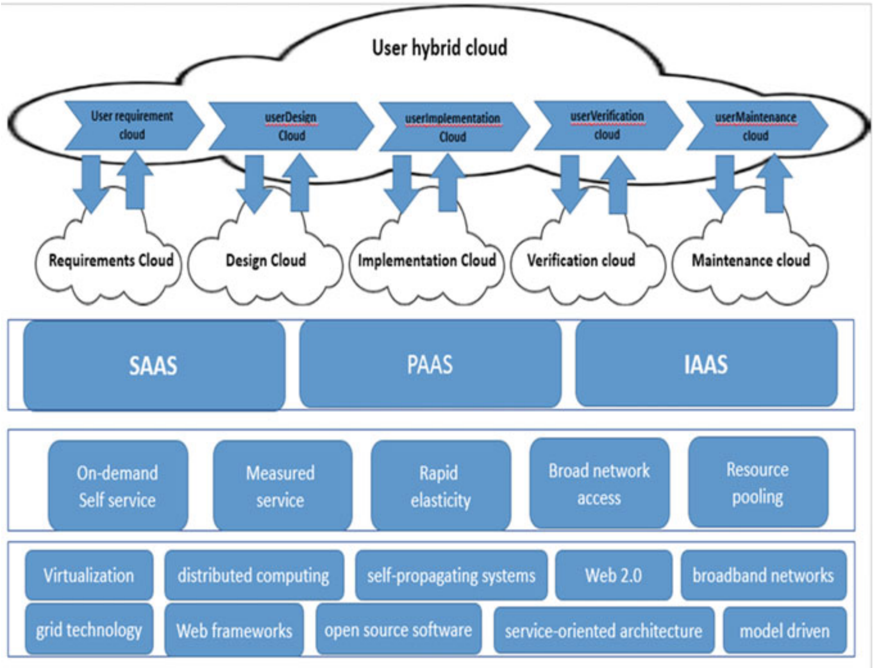


Fig. 5 Software architecture by cloud computing(cloud driven architecture)

1.5 Future Research and Conclusions

Due to the shape of the cloud architecture, the user can only produce the software with a system and even with low knowledge. Access to all types of architectures and software and hardware features in the cloud architecture provides a fast, agile design of software, from the smallest to the largest software without increasing complexity. The cloud architecture features a drastic reduction in the design and training costs of employees and programmers and their managers.

References

1. http://www.maktreesolution.com/content/service-oriented-architecture-soa-web-services
2. Alhir SS (2003) Understanding the model driven architecture (MDA). Method and Tools 11(3):17–24
3. https://en.wikipedia.org/wiki/Waterfall_model
4. Zhang Q, Cheng L, Boutaba R (2010) Cloud computing: state-of-the-art and research challenges. J Internet Serv Appl 1(1):7–18

Designing and Implementation a Simple Algorithm Considering the Maximum Audio Frequency of Persian Vocabulary in Order to Robot Speech Control Based on Arduino

Ata Jahangir Moshayedi, Abolfazl Moradian Agda and Morteza Arabzadeh

Abstract Based on the definition, speech process is known as the audio signals Conversion process as an input for checking systems by computer algorithms. The importance of this field is for many applications such as aerospace. Automatic translation, providing news texts from lectures, home intelligent automation, Computer games. Serving the blinds and low—power people, collecting and organizing different information resources like books, websites and also facilitate and expedite in educational services. This study has done with the purpose of introducing and testing a simple algorithm for Persian vocabulary and implementation of a robot structure based on the Arduino. The important point of this study is using audio signals in Persian language which has a history less than two decades. Although some attempts such as Nevisa and speech to text conversion software has been made, but based on researcher's investigation, there was not found a controllable product based on Persian vocabulary. Creating relevant response with the highest audio frequency based on three Persian terms, "right", "left" and "straight" is the purpose which has considered in this article. In this article, Sound is received by the microphone and through Matlab software by getting and processing maximum received signal frequency, the higher frequency of receiving sound will be determined and transferred to the Arduino board existence on robot through serial communication which leads to a movement reaction based on the definition of this number by a robot. The way of sampling is totally free considering the kind of microphone and distance, and results of this study in the review of three words, "left", "right" and "straight" is an indicator of the efficiency of the sug-

A. J. Moshayedi (✉)
Department of Electrical Engineering, Doalt abad Branch,
Islamic Azad University, Isfahan, Iran
e-mail: moshayedi@gmail.com

A. M. Agda · M. Arabzadeh
Department of Medical Engineering, Ragheb Esfahani, Isfahan, Iran

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_25

gested algorithm with the success rate of 73 and 24% failure at the first repetition and with the success rate of 86 and 14% failure at the second repetition.

Keywords Speech control · Sound processing · Audio processing
Speech processing · Arduino · Audio signal · Hearing system

1 Introduction

Speech is the first meant of communication between humans and also is one of the most effective methods in this regard. In the Classical definition, Speech processing is the process of converting audio signals into word sequences by computer algorithms which takes place with the purpose of development of methods and systems for using this signal as an input single [1]. This tends as one of the artificial intelligence fields is simulating speech issues in human, included detecting and understanding speech, speech production and improving speech quality. Ayat [2], and also it investigates for increasing people's quality of life in different levels of society and making a quick and easy human communication with machines around with no needs for hardware connections [3]. The initial step for the topic started in 1950 [4] which, considering available basic facilities in computers, it was able to identify and recognize limited words. One of the most Significant scientific activities in this field has made efforts in1971–1976 named SUR program, as one of the largest program in the field of speech recognition. Faster processor entry in 90s leads to use this signal increased to 80% [5]. With the internet development, even using these systems got increasing growth and leads to enter voice commands in windows and MAC operating systems [6]. Nowadays, different activities have done in this regard, which among them we can mention to the different uses, such as aerospace, automatic translation, providing news texts from lectures, home intelligent automation, Computer games, Serving the blinds and low—power people, Collection and organizing different information resources like books, websites and also facilitate and expedite in educational services [4, 7–12]. Even some efforts such as designing and developing a voice control wheelchair and voice control robotic arm have done

1.1 Speech Recognition Techniques

The Purpose of speech recognition for a device, is hearing, realizing. And running a command in order to analysis, extraction and identification of information related to the identity of the owner of the voice. The voice recognition system can be included 4 steps: speech analysis, feature extraction, modeling, testing. For distinction these audio patterns, up to now, various methods have presented containing main components analysis, linear separation analysis (LDA) 2 independent components analysis, linear prediction coding, and Cepstral analysis, frequency Cepstral (MFFC), which has shown briefly with the, way of execution and features in Fig. 1 [13–15].

As it has shown in Fig. 1, each mentioned methods have different specifications used by researchers [16]. This investigation has done with the purpose of introduction and the experiment of a simple Persian word algorithm and a robot structure implementation based on Arduino which, according to a researchers' investigation, it has an antecedent of less than two decades and hardly Persian words with the few articles. The purpose of this study, is designing a robot using an

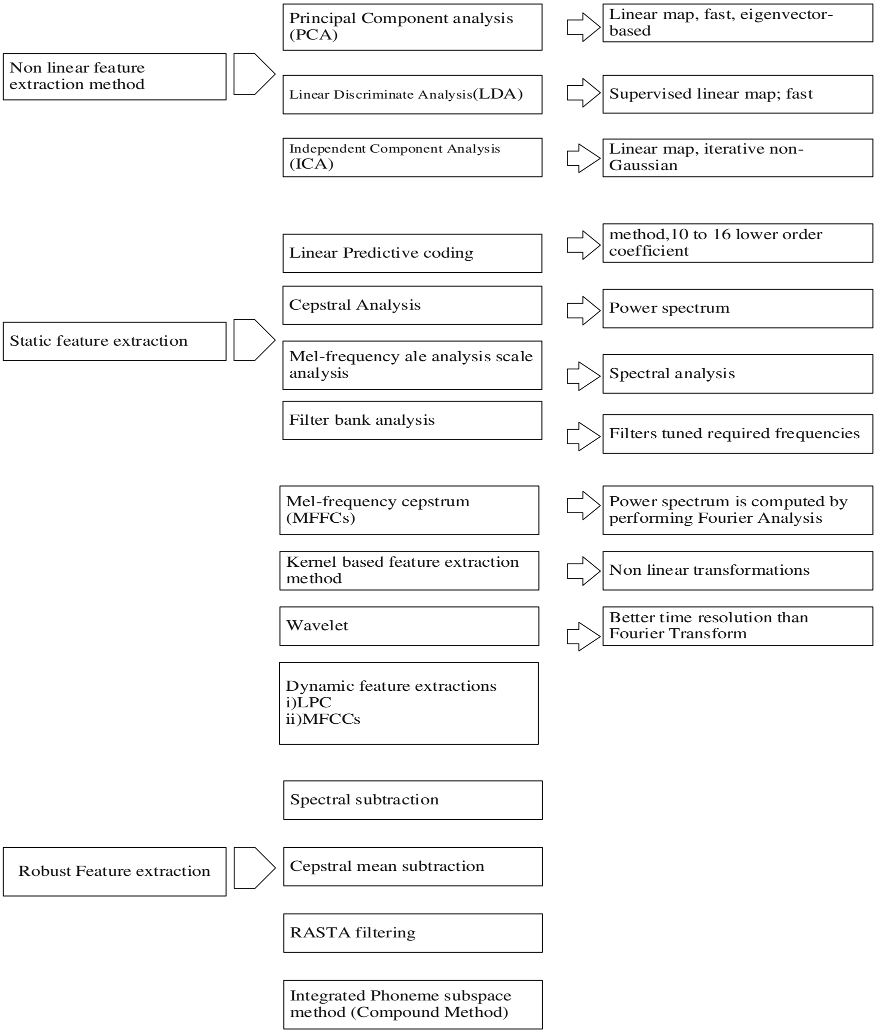

Non linear feature extraction method
Principal Component analysis (PCA)
Linear map, fast, eigenvector-based
Linear Discriminate Analysis(LDA)
Supervised linear map; fast
Independent Component Analysis (ICA)
Linear map, iterative non-Gaussian
Linear Predictive coding
method,10 to 16 lower order coefficient
Cepstral Analysis
Power spectrum
Static feature extraction
Mel-frequency ale analysis scale analysis
Spectral analysis
Filter bank analysis
Filters tuned required frequencies
Mel-frequency cepstrum (MFFCs)
Power spectrum is computed by performing Fourier Analysis
Kernel based feature extraction method
Non linear transformations
Wavelet
Better time resolution than Fourier Transform
Dynamic feature extractions
i)LPC
ii)MFCCs
Spectral subtraction
Robust Feature extraction
Cepstral mean subtraction
RASTA filtering
Integrated Phoneme subspace method (Compound Method)

Fig. 1 The speech recognition methods

Arduino board In order to steer acoustic line through the Persian words and also using maximum speech frequency of simple algorithm. In this article, first using hardware and its components has studied, then the speech signal process algorithm has introduced. In the next section, after a description of the method of sampling, the suggested algorithm will test and at the end, the article will be finished by study of conclusion.

2 System Designing and Implementation Suggested Algorithm

With attention to mentioned subjects, various algorithms have used for speech control. This study has done with the purpose of introducing and implementation of a simple algorithm in order to robot navigation based on Arduino which all components and way of relating between this, system has shown in Fig. 2.

As it has shown in Fig. 2, the system components consist of two parts. (A) Signal process: consists of recording speech command and issuing of movement control by computer. (B) Consists of robot movement control. The first part is responsible for speech signal conversion by taking record word samples into processed data, and the second part is responsible for navigation which has explained in the following.

2.1 The Components of Signal Processing Part

As it has been explained, the signal processing part has responsibility of analysis and sending a proper command to control the robot. The components of this part have shown in Fig. 3 which is described separately.

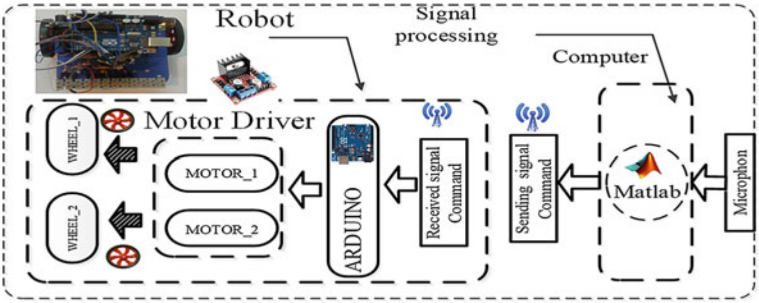


Fig. 2 The Robot block Diagram

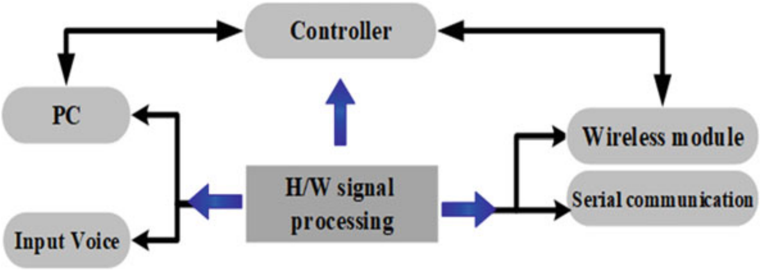


Fig. 3 The Signal processing segments

2.1.1 Sound Input

In order to sound input considering kind of sampling, different microphones have used which has explained in algorithm part.

2.1.2 Computer(PC)

In the Matlab software, a personal computer or laptop as a main processor (CPU) and algorithm implemented in tests and experiments (Lenovo 7500 with the processor Intel core i7 2.2 C-HZ and RAM memory 8 GB with Matlab 2017 software has used.

2.1.3 Wireless Connection Module

For sending data to robot, sending module NRF 2401 has used which has explained in the following [17] (Fig. 4a).

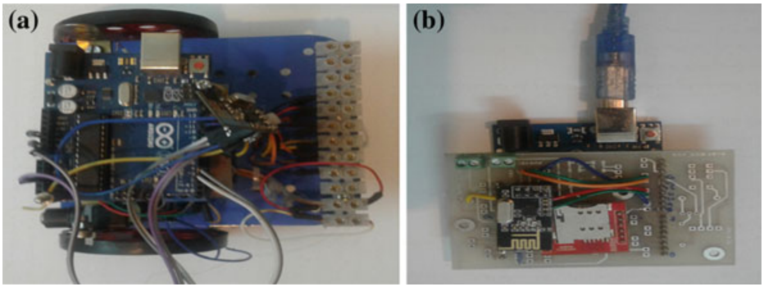


Fig. 4 **a** The robot with the receive NRF Module. **b** The Send NRF Module

2.1.4 Controller Number One

In order to sending data from computer to robot, has used UNO board has used [18].

2.2 *Robot Components*

Robot hardware components have shown in Fig. 5. As it mentioned before, this part is responsible for receiving the final signal and robot movement.

Figure 5 shows the robot structure and way of relating between hardware components in robot section. This robot consists of Arduino board UNO. Serial port connection on the board, 2 DC Motors, L298 motor driver, and wireless connection to receive command signals which both parts are managed by UNO controller. Also, robot frame is included a plate and 2 wheels (with 3.5 cm radius) that will be explained in the following.

2.2.1 Controller Number Two

In order to receive final signal from receiving signal process part, robot part needs a controller for receiving data from process part and the wireless signal is responsible for robot steering and creating proportion command for the wheels. In this section with attention to the robot nerd, UNO controller with serial port connection has used. The controller features have selected according to Table 1 [18] and other robot components have connected to it by the wires.

2.2.2 DC Motor

In order to robot structure, movement, two 5 V DC motors with gearbox at the left and right side has placed under the robot body, has used.

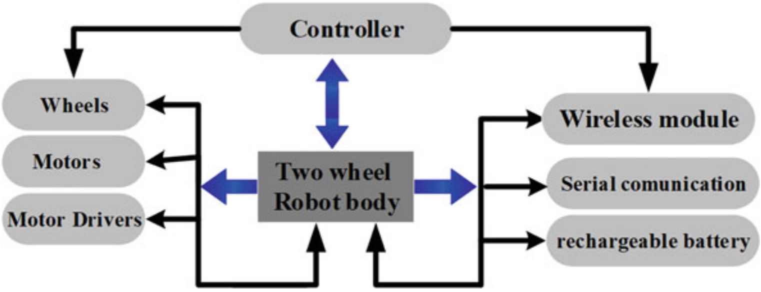


Fig. 5 The robot component diagrams

Table 1 The Uno specification

1	Microcontroller	ATmega328P
2	Operating voltage	5 V
3	Input voltage (recommended)	7–12 V
4	Input voltage (limit)	6–20 V
5	Digital I/O pins	14
6	PWM digital I/O pins	6
7	Analog input pins	6
8	DC current per I/O pin	20 mA
9	DC current for 3.3 V pin	50 mA
10	Flash memory	32 KB
11	SRAM	2 KB
12	EEPROM	1 KB
13	Clock speed	16 MHz

2.2.3 Motor Driver

Considering the needs of DC motors to driver, chip L298 included two DC drive, has used. This driver has a high thermal protection and low saturation voltage. It also has low noise which is responsible for moving the robot wheels with the capability of turning left and right. The driver specifications are shown in Table 2.

2.2.4 Wheel

In this robot, two wheels with 3.5 cm radius have used which is connected to the sides of robot body and with attention to the motor command, they move to “left”, “right” and “straight”.

Table 2 The L 298 dual full bridge driver specification

1	Driver power supply	+5 V to +46 V
2	Driver Io	2A
3	Logic power output Vss	+5 to +7 V (internal supply +5 V)
4	Logic current	0 to 36 mA
5	Controlling level	Low −0.3 to 1.5 V, high: 2.3 V to Vss
6	Enable signal level	Low −0.3 to 1.5 V, high: 2.3 V to Vss
7	Max power	25 W (Temperature 75 cesus)
9	Dimension	60 mm * 54 mm

2.2.5 Battery

With respect to robot movement independently from the power supply in this design, 3 rechargeable 3. 8 V battery with 2150 ma current, 8.2 used as the power supply [19, 20].

2.2.6 Wireless Connection Module:

Considering the possibility of existing distance between the computer and robot, the NRF module, model NRF 24L01 in order to wireless information transmission between two Points (for example computer and microcontroller) has used. The effective distance through this module is 1000 m and information are transferred through radio wave with 2.4 GH frequency [17]. Table 3 shows the features of module NRF.

3 The Robot Direction Detection Algorithm

With attention to the main purpose, this study needs to suggestion and implementation of algorithms in order to issue a control command based on Persian words, which, considering existing changes received from different samples, we can't use the created frequency with them, directly. In this case, motion algorithm with the following stages has considered. First stage: receiving a sound input signal. Second stage: sound processing and extracting features from sound signal. Third stage: Sending maximum number to Arduino in order to determine reaction. Fourth stage: A comparison between receiving amount and distinguish the Operation in Arduino board. Fifth stage: Sending a message to motors for movement (On and Off) which has shown in Fig. 5.

3.1 Sound Processing Algorithm

As it explained before, in this study, Matlab software 2017 has used for receiving and analyzing speech data which mentioned software moreover calculation, has the accessibility to the board and serial port.

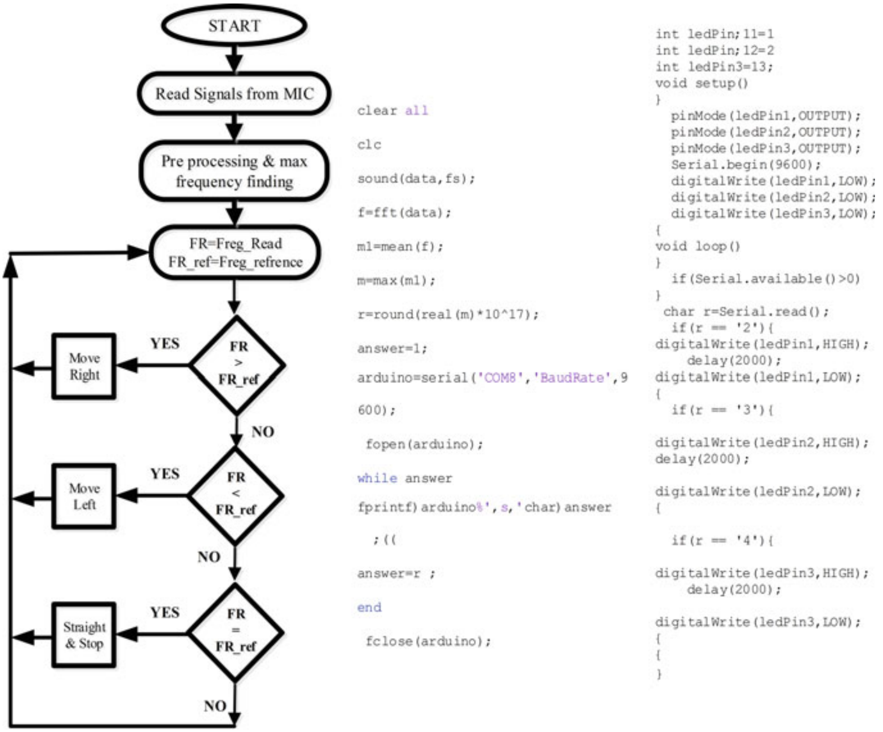
Table 3 The NRF send and received module specification

Specification NRF24L01	
Supply range	1.9–3.6 V
Frequency	2.4 GHZ
Maximum power for data transmission	100 mW equal to 20 dbm
Succeeded range	1000 m

As it has shown in Fig. 6, in this design. First, speech sample receives, Then, the maximum input signal in Matlab is obtained and is also calculated which consists of extracting sound spectrogram, using filters in the way of finding the maximum amount of signal and multiply in signal number, and again finding maximum signal on the generated spectrogram frames [6].

In the next stage, the obtained information in Matlab software, by serial port connection, transfers to the Arduino board. In this system, reactions are based on audio original frequencies and their conformity with a reference number is determined and as sampling, three reactions are considered.

First motion: turning right if the tested audio frequency is less than reference is less than reference number. Second motion: Turning left: if the tested audio frequency is more than reference number. Third motion: straight movement: if the tested audio frequency is equal to the reference number.



```
clear all

clc

sound(data,fs);

f=fft(data);

m1=mean(f);

m=max(m1);

r=round(real(m)*10^17);

answer=1;

arduino=serial('COM8','BaudRate',9

600);

 fopen(arduino);

while answer

fprintf)arduino%',s,'char)answer

  ;((

answer=r ;

end

 fclose(arduino);
```

```
int ledPin;11=1
int ledPin;12=2
int ledPin3=13;
void setup()
}
  pinMode(ledPin1,OUTPUT);
  pinMode(ledPin2,OUTPUT);
  pinMode(ledPin3,OUTPUT);
  Serial.begin(9600);
  digitalWrite(ledPin1,LOW);
  digitalWrite(ledPin2,LOW);
  digitalWrite(ledPin3,LOW);
{
void loop()
}
  if(Serial.available()>0)
}
 char r=Serial.read();
  if(r == '2'){
digitalWrite(ledPin1,HIGH);
    delay(2000);
digitalWrite(ledPin1,LOW);
{
  if(r == '3'){

digitalWrite(ledPin2,HIGH);
delay(2000);

digitalWrite(ledPin2,LOW);
{

  if(r == '4'){

digitalWrite(ledPin3,HIGH);
    delay(2000);

digitalWrite(ledPin3,LOW);
{
{
}
```

Fig. 6 The algorithm with the Matlab and Arduinos Program

4 Algorithm Operation Study

In order to study of mentioned algorithm operation, first sampling of audio signals without any distance limitation, kind of microphone and frequency has done. Then, by Matlab software, it has analyzed and final number has extracted. At the end, considering the output number, some command is sent to a robot which is explained in the following.

4.1 Audio Signal Sampling and Receiving Sound from Microphone

Considering the main purpose of this study in robot control, we need to have audio data which among 26 different samples, in order to study of efficiency for suggested algorithm, without any limitation in microphone distance and kind of microphone, by the mobile phone of each case has done. The purpose of this kind of sampling is detecting the voice of person, independent from sound environmental filtering and way of speech in sampling. Used samples in this study are included six samples of woman voice 35% and 20 samples of man voice 65% in different ages which has shown in Fig. 7.

As it explained, for testing algorithm, among samples, 17 men and 9 women have selected. At the time of sampling, Use has the choice of using the words “right”, “left” and “straight”. The specifications of receiving data along with age of people have shown in Tables 4 and 5.

Obtained results of the audio frequency analysis in order to distinguish motion, has shown in Fig. 8.

As it has shown in Table 6 and Figs. 8 and 9. Different wave form in signal analysis for the word “right”, “left” and “straight” give the capability of using maximum audio frequency and refereeing the reference number to user.

According to the presented algorithm and as it has shown in above Fig. 9 for “Straight” command, number 19 as the maximum amount, for “right” command,

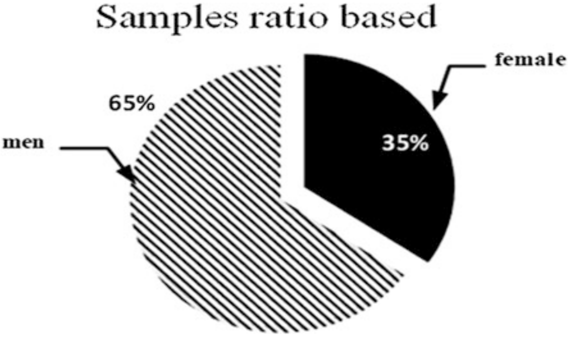


Fig. 7 The ration of used data based on genders sample

Table 4 The used sound specification for men

No		1	2	3	4	5	6	7	8	9	10
Gender		Men									
Age		59	27	25	26	31	37	25	37	49	31
Sound channel number		2									
Sample rate		4800									
Total samples	Straight	106,560	81,720	68,088	34,560	68,088	69,120	56,376	56,376	95,040	85,560
	Right	198,720	65,592	80,568	48,960	68,088	118,080	56,376	80,568	51,840	93,432
	Left	132,480	70,968	61,752	34,560	61,752	175,680	56,376	73,464	54,720	68,088
Time duration	Straight	2.22	1.703	1.419	0.72	1.419	1.44	1.175	1.175	1.98	1.783
	Right	4.14	1.367	1.679	1.02	1.419	2.46	1.175	1.679	1.08	1.947
	Left	2.76	1.479	1.287	0.72	1.287	3.66	1.175	1.531	1.14	1.419
Bits per sample		16									
No		11	12	13	14	15	16	17	18	19	20
Gender		Men									Women
Age		20	26	27	37	18	32	36	22	66	30
Sound channel number		2							1	1	2
Sample rate		48,000						800	48,000		
Total samples	Straight	73,464	68,088	93,432	48,960	69,120	80,568	3865	94,224	176,112	65,592
	Right	73,464	68,088	80,568	48,960	31,680	85,560	3568	96,240	153,600	65,592
	Left	56,376	61,752	80,568	40,320	28,800	68,088	3865	136,176	166,896	76,344
Time duration	Straight	1.5305	1.4185	1.9465	1.02	1.44	1.6785	0.48313	1.963	3.669	1.3665
	Right	1.531	1.419	1.679	1.02	0.66	1.783	0.446	2.005	3.2	1.367
	Left	1.1745	1.2865	1.6785	0.84	0.6	1.4185	0.48313	2.837	3.477	1.5905
Bits per sample		16						8			

Table 5 The used sound specification for woman

No		21	22	23	24	25	26
Gender		Female					
Age		62	11	45	4.5	65	56
Sound channel Number		2	2	2	2	1	2
Sample rate		48,000			8000	48,000	
Total Samples	Straight	77,760	41,976	36,984	6871	172,032	73,464
	Right	92,160	51,000	36,984	3761	167,952	80,568
	Left	120,960	41,976	41,976	3865	186,384	61,752
Time duration	Straight	1.62	0.875	0.771	0.8589	3.584	1.531
	Right	1.92	1.063	0.771	0.4701	3.499	1.679
	Left	2.52	0.875	0.875	0.4831	3.883	1.287
Bits per sample		16				8	16

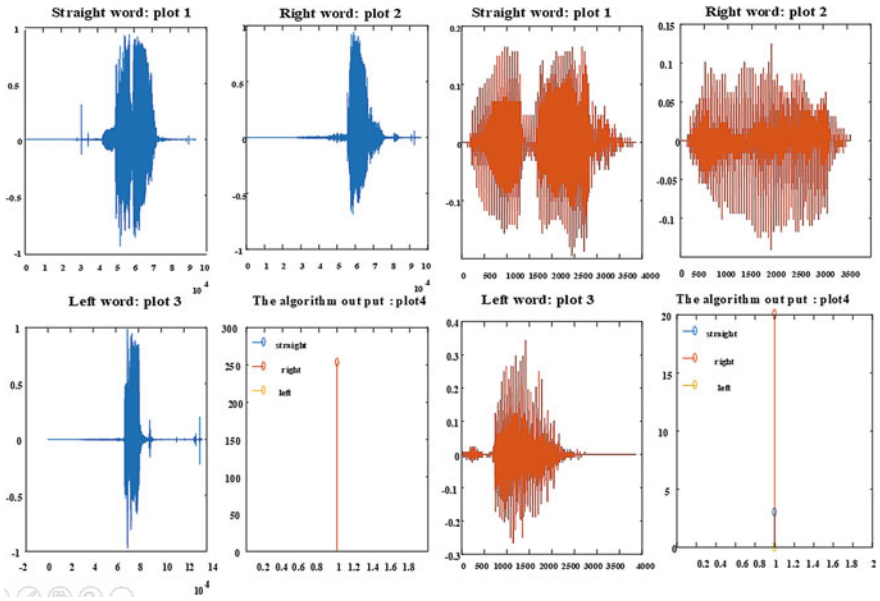


Fig. 8 The algorithm result for the men with the age 37 and woman in 56

the numbers of more than 19 and for "left" command, number "0" have considered which above amounts has imported into Arduino and the robot operation truth has tested. Also. In some samples, for some of terms, user had to repeat the sample, which obtained results of this repetition has shown in Fig. 10.

Table 6 The Algorithm, output for the three word of straight, left and right

No	Gender	Age	The algorithm o/p		
			Straight	Right	Left
1	Woman	11	4	33	0
2		4.5	3	57	0
3		65	12	73	0
4		45	4	312	0
5		20	11	226	0
6		31	7	354	0
7		62	2	53	0
8		30	12	43	0
9		56	11	162	0
10	Men	37	2	237	0
11		25	1	53	0
12		37	6	34	0
13		31	13	108	0
14		26	10	78	0
15		25	44	224	1
16		26	8	120	0
17		59	33	114	0
18		27	8	336	0
19		26	1	78	0
20		49	16	30	0
21		36	3	20	0
22		66	12	73	0
23		22	8	253	0
24		32	8	156	0
25		18	18	132	0
26		38	1	297	0

5 Conclusion

The purpose of this study is designing a robot using an Arduino board in order to steer acoustic line through the Persian word and also using maximum speech frequency of simple algorithm. Considering explained hardware, the aim of selecting each part and task of each robot system design components have shown in Table 7.

In this study. After extracting sound spectrogram using filters in the way of finding the maximum amount of signal and multiply by signal number and again finding maximum signal on the generated spectrogram frames is done [6]. In the next stage, the obtained information in Matlab software by serial port connection, transfers to the Arduino board. In this system, reactions are based on audio original frequencies and their conformity with a reference number has determined and as

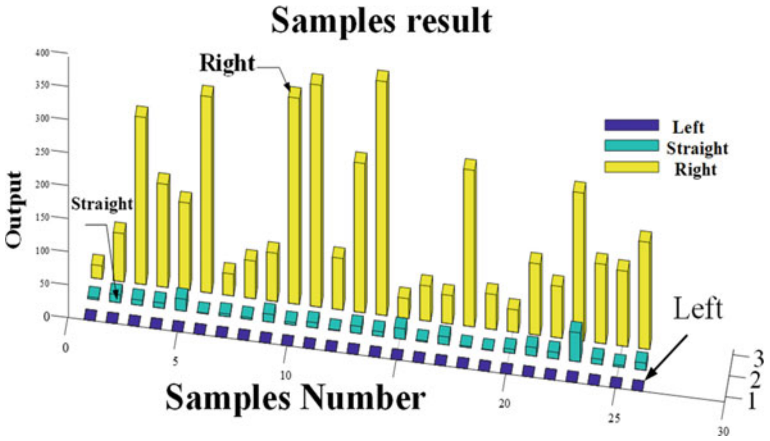


Fig. 9 The algorithm result and samples result

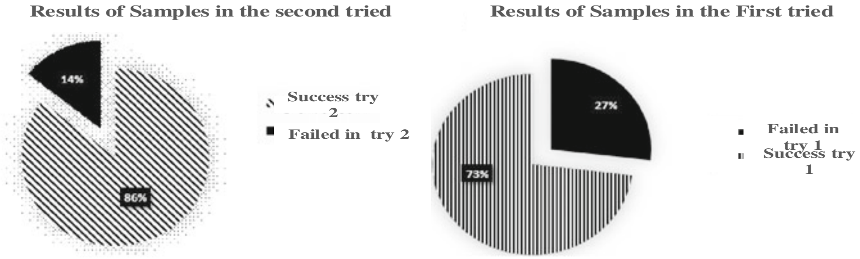


Fig. 10 The result of algorithm performance for the first and second sample entrance try

Table 7 The robot component tasks and missions

No	Item	Task and missions	Advantages
1	Micro controller	Capability of simple programming Ability of program handling	Reasonable prize Including the proper i/o with respect to project
2	Motor driver	Proper voltage and current to control the motors by the micro controllers	The high stage of heat protection Less noise Influence The designed precaution for the Servos and stepper motors Including the two driver in on chip
3	Wireless communication	Send and receive data	Reasonable prize and proper performance

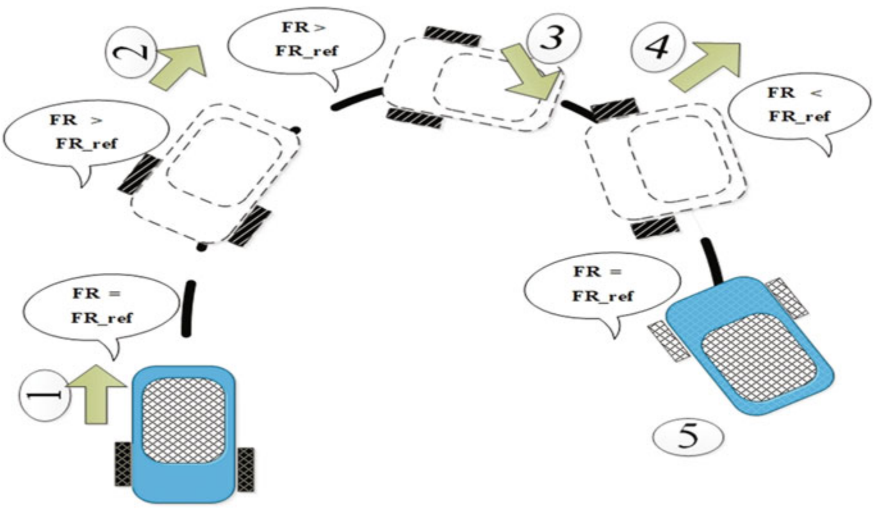


Fig. 11 The robot performance in the proposed algorithm

sampling, three motion: “straight”, “right” and “left” has considered Presented results in Fig. 10 have shown that the success rate in first repetition is 73% and the failure rate is 27%, but in second repetition, the success rate is 86% and the failure rate is 14%. So, above designing has capability of robot navigation and each other device by using audio signals (Fig. 11).

This pattern also can be used for controlling of wheelchair for low—power people and special equipment in those industries which has no possibility of direct control. This design also has some features such as simple algorithm for successful diagnosis of Persian vocabulary with the success rate of 73% and failure rate of 27% at the first repetition and with the success rate of 86% and failure rate of 14% at the second repetition. Also, using Arduino board leads to dimension reduction and control circuits. Moreover, considering expenses of using computer to reduce dimensions and price of some boards such as “Raspberry pi”, Can be used. From the major problems ahead in above study, can be mentioned to microphone differences and user distance to microphone and kind of accent and also number of words which is some words leads to repetition of samples and has considered as the investigation future plan.

References

1. Waibel A, Lee K-F (1990) Readings in speech recognition. Morgan Kaufmann
2. Ayat S (2004) The elementary if speech processing. Payam noor, tehran 2004
3. Dutoit T (1997) An introduction to text-to-speech synthesis, vol 3. Springer Science and Business Media

4. Rabiner LR, Juang B-H (1993) Fundamentals of speech recognition
5. Waibel A, Lee K-F (1990) Readings in speech recognition. Morgan Kaufmann
6. Shadiev R, Hwang W-Y, Huang Y-M, Liu C-J (2016) Investigating applications of speech-to-text recognition technology for a face-to-face seminar to assist learning of non-native English-speaking participants. Technol Pedagog Educ 25(1):119–134
7. Markowitz JA (2000) Using speech recognition. J. Markowitz Consultants
8. Brown MK, Buntschuh BM (1999) System and method for voice controlled video screen display. Google Patents, 30 Mar 1999
9. McCall DF, Logue LM, Zelina FJ, Sendak MV, Hinson JR, Sanders WL, Belinski S, Holtz BE (2003) Voice controlled surgical suite. Google Patents, 08 July 2003
10. Furukawa M (1988) Voice controlled toy. Google Patents, 05 Jan 1988
11. Prathima N, Kumar S, Ahmed KL, Chakradhar G (2017) Voice recognition based home automation system for paralyzed people
12. Purwanto E, GigihPrabowo A, EndroWahjono M, Rifaldi M (2014) The development of induction motor characteristic by using radial basis function (RBF) learning method for electric vehicle application
13. Clark JA, Roemer RB (1977) Voice controlled wheelchair. Arch Phys Med Rehabil 58 (4):169–175
14. Meena K, Gupta S,. Khare V (2017) Voice controlled wheelchair
15. Gaikwad SK, Gawali BW, Yannawar P (2010) A review on speech recognition technique. Int J Comput Appl 10(3):16–24
16. Honarmand H, Khoramizadeh M, Eshraghi S (2009) Evaluation of patients sera for early diagnosis of Leptospirosis by PCR. J Ardabil Univ Med Sci 9(4):353–359
17. nRF24L01, nRF24L01 Single Chip 2.4 GHz Transceiver Product Specification Datasheet. 2017
18. Moshayedi AJ, toutian omid, enjoy the programming with Arduino, arachap, 2014
19. High capacity rechargeable Li-Ion battery specifications, product specification datasheet, 2017
20. Moshayedi AJ, Gharpure DC (2017) Evaluation of bio inspired Mokhtar: odor localization system. In: 2017 18th international carpathian control conference (ICCC)

Simulation of Bayard Alpert Ionization Vacuum Gauge with COMSOL

Sadegh Mohammadzadeh Bazarchi and Ebrahim Abaspour Sani

Abstract In this paper, the Bayard Alpert ionization vacuum gauge has been simulated using COMSOL. Employing Charged Particle Tracing (CPT) tool in COMSOL, the sensitivity coefficient of the gauge has been calculated from simulation results. Based on the results, the average measured value of sensitivity is 30, which shows compatibility with theoretical value of 28. It must be mentioned that a unique method is used in this article to calculate the sensitivity coefficient and the coefficient diagrams for sensitivity of different vacuum pressures have been obtained through these simulations. This paper also presents a method for obtaining electron velocity and energy and their diagrams.

Keywords Ionization vacuum gauge · Vacuum sensor · Hot cathode ion gauge
Batard Alpert ion gauge

1 Introduction

The operation of ionization vacuum gauge is based on the ionization of gas while its two main classifications involve hot and cold cathode types. The hot cathode type which is known as "Bayard Alpert", has got its name from its inventors. It has been used since 1950 and can measure vacuum pressures up to 10–2 to 10–13 torr [1–3].

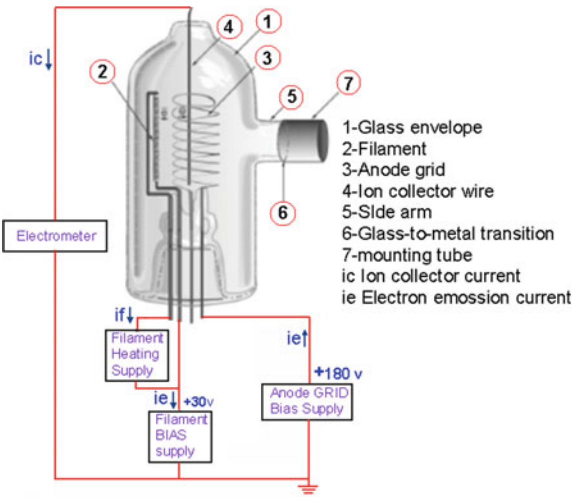
To explain the behavior of this gauge, one can say that a cylindrical ionization environment is formed by means of thin wires known as "Grid" or the "Anode", while the thin wired collector has been placed in the middle of this cylinder. According to Fig. 1, the cathode filament is located in outside, close to the anode wire. The tungsten cathode is coated with thorium dioxide, which is heated by means of the electric current and because of the thermionic phenomenon the

S. Mohammadzadeh Bazarchi (✉) · E. Abaspour Sani
Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran
e-mail: s.mohammadzadeh@urmia.ac.ir

E. Abaspour Sani
e-mail: e.abbaspour@urmia.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_26

Fig. 1 Bayard Alpert ionization vacuum gauge schematic



electrons will be generated freely around it. Ie denotes the electric current of the cathode wire, which supports the released electrons. These electrons will be accelerated as a result of the electrostatic attraction coming from +180 V potential of the anode while their velocity will rapidly grow to 10^5 v/s at the moment of entering in the gauge environment. The electrons inside the anode network will collide with the gas atoms causing them to be ionizated.

Along with the ionization of every gas atom, a positive ion and an electron will be generated where the positive ion is being attracted by the negative potential, e.g., the 0 V potential of the collector wire. Ic is the current resulting from the accumulation of these electrons which depends on the ionization intensity and the number of positive ions. On the other side, the ionization value depends to the number of collisions between the gas atoms and energized electrons. Therefore, as the gas pressure drops, the number of gas atoms per unit volume will be reduced resulting in the lower ionization rate. Considering these explanations, the relation which best describes Ic will be (1) [2].

$$\mathbf{Ic} = \mathbf{S}.\mathbf{Sr}.\mathbf{Ie}.\mathbf{P} \tag{1}$$

In Eq. (1), P represents the gas pressure in torr while the coefficient Sr pertains to the measured gas. For nitrogen, Sr is equal to 1 and for the other gases it can be obtained from the ionization cross section of that gas with respect to the nitrogen. For example, in the case of oxygen, the corresponding value is obtained by means of:

$$\boldsymbol{\sigma}_{\boldsymbol{o}2}(\varepsilon)/\boldsymbol{\sigma}_{\boldsymbol{N}2}(\varepsilon)$$

Coefficient S is the representation of ionization gauge sensitivity which is related to physical, geometric and electrical dimensions of the gauge. It is specified by the

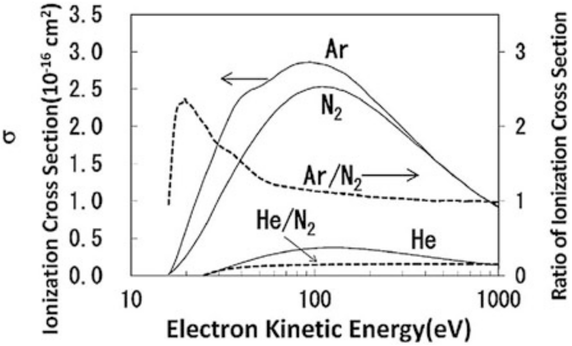


Fig. 2 Ionization cross-sectional diagram in nitrogen gas [10]

gauge manufacturer and is its most important parameter. As the sensitivity of the gauge increases, the corresponding cathode current consumption for measurement of a certain vacuum pressure will be decreased. In addition, high sensitivity provides the ability of lower vacuum pressure measurements. It must be noted that the sensitivity unit is 1/torr, while its typical value is reported to vary from 10 to 45. The corresponding value of S has been expressed in Eq. (2) [4–6].

$$\boldsymbol{S} = \boldsymbol{\sigma}_{\boldsymbol{i}} \cdot \frac{\boldsymbol{l}}{\boldsymbol{kT}} \tag{2}$$

As Eq. (2) explains, sensitivity coefficient depends to the effective length of electron motion paths traversing inside the grid, which lead to the ionization. The larger the size of the gauge leads to higher value for sensitivity coefficient and it is considered as a compromise for the ionization vacuum gauge designers who intend to minimize its structure or implement it with MEMS technology [7–9]. In this relation σ_i denotes the ionization cross section which does not have a constant value and is a function of the electron energy. According to Fig. 2, which shows the ionization energy of oxygen, the ionization energy reaches a maximum value at 100 eV. As a result, the potential of the anode and cathode will be selected equal to 180 and 30 V, respectively so that the average energy of the electrons can reach to 100 eV which is identical to the highest efficiency.

2 Physical Characteristics

In Fig. 3, the schematic of the Bayard Alpert ionization gauge is illustrated. It is based on the information taken from several examples of commercial ionization gauges. The details about this information have been summarized in Table 1.

According to Table 1, a very thin collector wire is selected. The reason of this choice pertains to the X-ray limitations. In hot cathode ionization gauge, the

Table 1 Specifications of ionization Bayard Alpert gauge

	Type	Number	Thickness (mm)	Heigh (mm)	Diameter (mm)
Grid	ThO_2Ir	10	0.5	40	21
Cathode	ThO_2Ir	1	2	40	–
Collector	Al	1	0.25	50	–

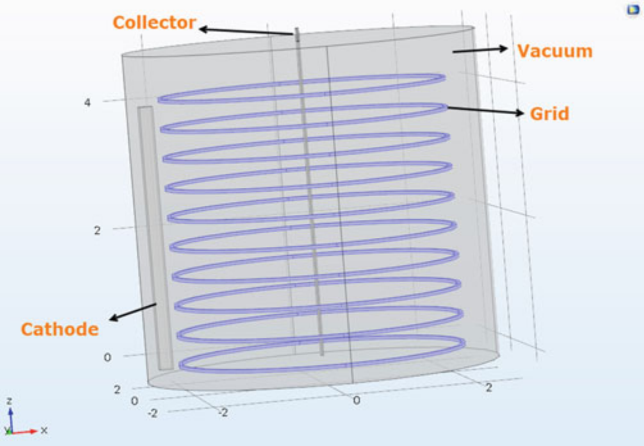


Fig. 3 Appearance of Ion gauge on COMSOL

collision of high-energy electrons with the anode network wires will produce X-rays. The generated X-rays will encounter with the collector wire causing the electrons to flee. The small leakage current resulting from the electrons escape is considered as an error factor in such gauges. In UHV vacuum, this small current is significant compared to the main current produced by the collected ions. Therefore, at the presence of X-ray error, UHV vacuum pressures e.g. less than 10^{-13} torr can not be measured. The main reason in which thin-collector wire is selected is to reduce the X-ray limitations. On the other hand by slimming the collector wire, the ion collection efficiency by the collector will greatly be reduced resulting in decrement of the sensitivity. In general, a trade off is usually established between the sensitivity and X-ray limitations.

Another important parameter is the distance between the cathode and the grid, which determines the intensity of the accelerating electric field. As the distance decreases, the field between the cathode and the grid will be stronger which increases the acceleration rate of the electrons. But the possibility of physical contact between the grid and cathodes, especially in the case of mechanical stresses due to the cathode heat must be considered.

The ionization vacuum gauge coating is also one of the challenges in the design and fabrication of ionization gauges. Ionization gauge can be designed either in a variety of glass, metallic and metallic-coated glass covers or with no covers, although each of the schemes contains its own advantages and disadvantages.

3 Simulation of the Ionization Bayard Alpert Gauge

The tested gas in this simulation can either be nitrogen, oxygen or any other gas. At first stage, the corresponding information related to the gas data must be applied to the COMSOL software as the input data to the (Charged Particle Tracing) CPT tool. The information includes:

- Gas density per unit volume
- Ionization energy
- Ionization collision cross-section
- Ionization energy loss
- Elastic collision cross-section
- Elastic energy loss

The information was extracted from the corresponding tables and diagrams of elastic collisions and ionization collisions provided in [10–12]. The extracted data were then entered in the CPT tool of COMSOL as a series of parameters and tables.

In Fig. 4, the simulation of the electron motion paths has been shown. In this simulation, 300 electrons were released from the cathode and have been accelerated because of the strong electrostatic field between cathode and the grid which then followed by their entrance into the ionization environment. The electrons dodge the collector and then will be reaccelerated towards to the grid. As a result, their velocity drops as they come close to the collector and after that, they will speed up which is clearly demonstrated in Fig. 4. It must be mentioned that most of the electrons will exit from the grid and be attracted by ionization gauge cover.

In Fig. 5, the energy curvature of the electrons simulated in Fig. 4 has been shown. Figure 6 shows the average energy curvature of the electrons. As the figure indicates, the energy of electrons drops as they approach the collector and then, their energy will grow after passing the collector. The average energy of the electrons in the entire path of this simulation reaches 67 eV, while their average velocity approaches to 4.22×10^6 m/s.

To simulate the collision process inside of the ionization vacuum gauge, the CPT tool in COMSOL is utilized. To perform this, the number of charged particles, i.e., the electrons, and their irradiation location will be adjusted in the "particle beam" section. If the number of electrons is very high, then the simulation interval will be a long period. Also, the path interpretation for the electrons and positive ions will be somehow impossible. Therefore, the number of electrons will be determined in conjunction with the vacuum pressure value of the simulation. In Fig. 7 simulation results for ionization and elastic collisions in Bayard Alpert ionization gauge are

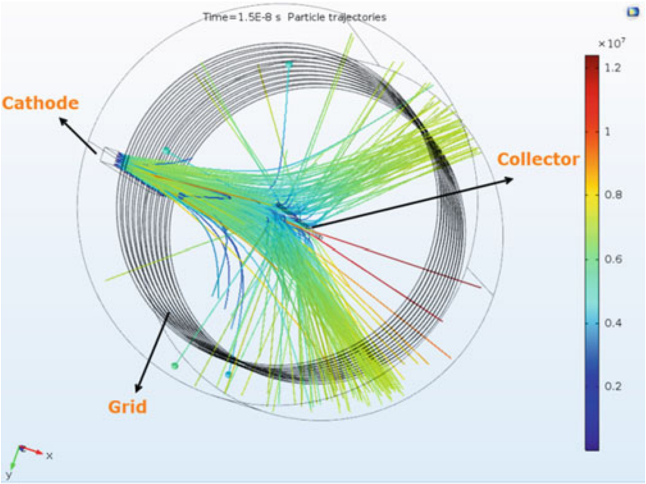


Fig. 4 Simulation of the 300 electrons motion in an ionization gauge by COMSOL

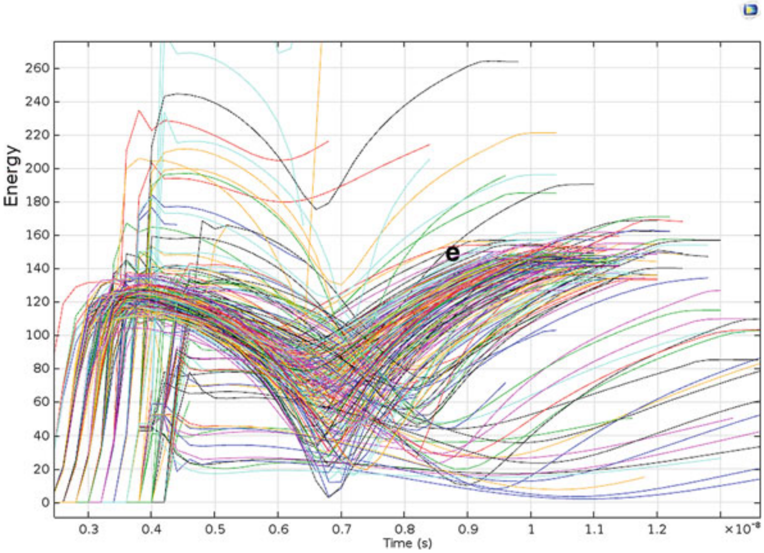


Fig. 5 Energy diagram of the electrons shown in Fig. 4

shown. It must be mentioned that the positive ions generation were shown once 100 electrons have entered into the ionization environment at the vacuum pressure of 5×10^{-5} torr.

The ionization process is the most important interaction which happens in the Bayard Alpert ionization gauge. In this process, high energy electrons will

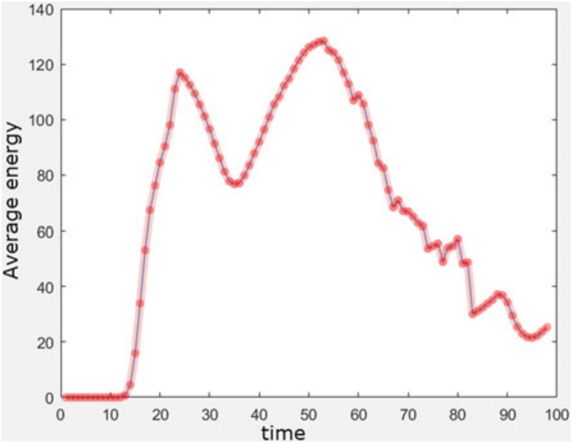


Fig. 6 Average energy of the electrons in the simulation of Fig. 4

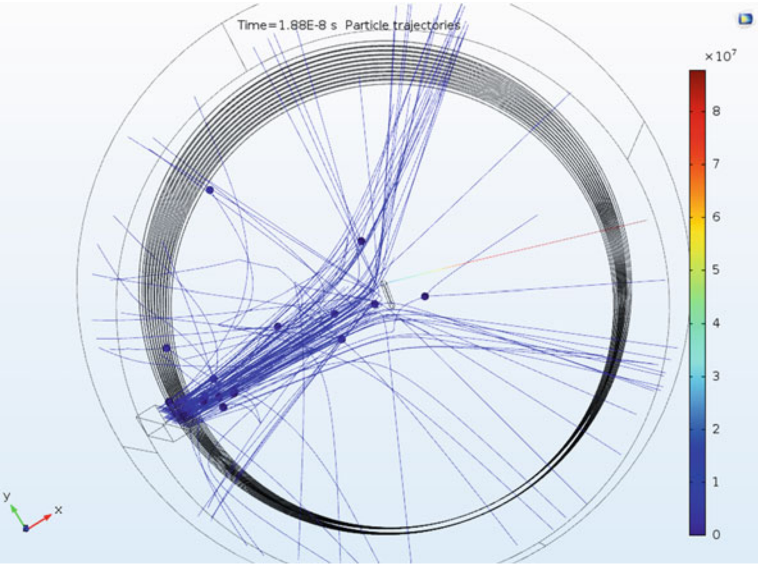


Fig. 7 Path and speed of movement of 100 electrons from the cathode into the grid and generates positive ions

encounter with gas atoms making them to be ionized. Hence, the energy of collided electrons will be reduced which depends to the gas type. For the case of nitrogen, the corresponding value is about 13 eV [11, 13]. Because of the ionization, a positive ion and a secondary electron will be produced. The Speed of positive ion is very low due to its high mass, but like the primary electron coming from the

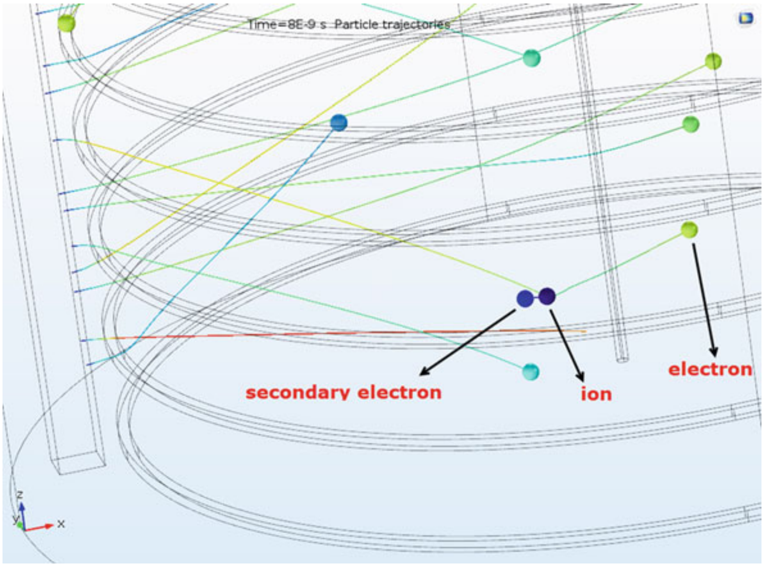


Fig. 8 Simulation of the ionization by means of CPT in COMSOL

cathode, the secondary electron contains a high velocity and may possibly participate in the ionization process. Again, with the help of CPT tool in COMSOL, the simulation has been carried out in which Fig. 8 shows the result for nitrogen ionization. It is obvious that the atomic gas is ionized by the colliding to the primary electron deflecting it from its normal path. The velocity of secondary electron is zero at the first moment and then gradually increases.

Generated ions should be attracted by collector wires which contains the potential of 0 V. The presence of electric field inside of the ionization gauge is in such a way that all of the electric field trajectories are towards the collector wire. Since the positive ions move in the same direction of the electric field, they will be attracted by the collector. In Fig. 9, the direction of electric field inside of the ionization gauge is specified with the help of arows. As it is obvious, the electrons will run away from the collector.

Because the ions in the ionization area are heavy, they will slowly be attracted by the collector. If the gauge response is considered, then the simulation of collected ions by the collector must be performed. Figure 10 illustrates the corresponding simulation in which the attraction time along with the number of ions can be extracted.

The sensitivity coefficient of ionization gauge is one of the most important parameters. In this paper, the ratio of accumulated ions in the collector to the number of electrons sent to the ionization environment at a specific vacuum pressure is used to calculate the sensitivity coefficient. Equation (3) demonstrates the corresponding expression to obtain the sensitivity coefficient in this method.

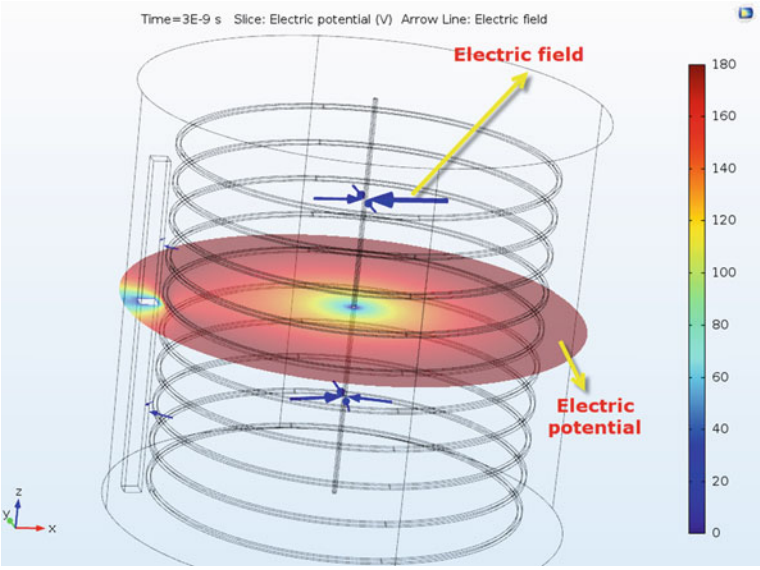

Time=3E-9 s Slice: Electric potential (V) Arrow Line: Electric field
Electric field
Electric potential
180
160
140
120
100
80
60
40
20
0
z
y
x

Fig. 9 Simulation of Potential and electric field inside ionization gauge

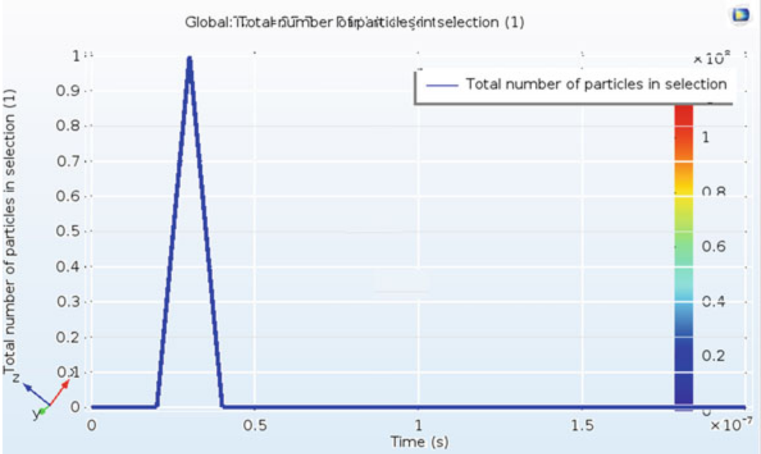


Fig. 10 Simulion of the time and number of ions collected by the collector wire

$$S.Sr = \frac{\boldsymbol{total\,ion\,collected}}{\boldsymbol{total\,ion\,electron\,trajected} \times \boldsymbol{P}} \quad \frac{1}{\boldsymbol{torr}} \tag{3}$$

Figure 11 shows the curvature of the sensitivity coefficient obtained for the ionization gauge in different vacuum pressures. It must be mentioned that the value

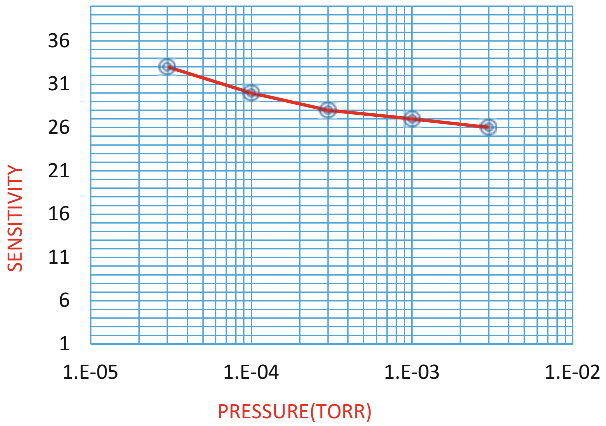


Fig. 11 Sensitivity coefficient in different vacuum pressures

of the sensitivity coefficient is derived from Eq. (3). This graph is achieved at the vacuum pressures where the Bayard Alpert ionization gauge is operating. The average value of the measured sensitivity coefficient is 28(1/torr) in this graph which shows compatibility with measured results in different commercial works as the reported value in those schemes vary from 10 to 45 [4–6].

4 Conclusion

In this paper the Bayard Alpert hot cathode ionization vacuum gauge is simulated using COMSOL program. In this simulation, the dimensions along with the specifications of comertial ionization gauges were employed. In addition, the simulations for electric field, electrical potential, elastic collisions and ionization collisions have been performed. Based on the simulation results, the quantities and graphs for electron velocity and electron energy were obtained either individually or by average value. Finally, the sensitivity coefficient was obtained by simulation and its diagram was presented. The average value which is 28, shows good compatibility with the measured value in real samples.

References

1. Bayard-Alpert Ionization Gauges, Stanford research systems http://www.thinksrs.com/products/BAgauges.htm
2. Bayard-Alpert Ionization Gauges, Stanford research systems www.thinksrs.com/downloads/PDFs/ApplicationNotes/IG1BAGapp.pdf

3. Alpert D (1953) New developments in the production and measurement of ultra high vacuum. J Appl Phys 24(7):860–876
4. P. M. Inc, Technical Brochure of PVC1000 Series Pirani vacuum sensors, [Online]. Available: http://www.posifamicrosystems.com/pdf/2017-51-00-03-15__PVC1000-Data-Sheet—March-2016.pdf
5. Technical Brochure of PVC1000 Series Pirani vacuum sensors, Posifa Microsystem Inc, www.posifamicrosystems.com
6. Tecknical Brochure of 999 Quattro- multi-sensor vacuum transducer, MKS instrument Inc. www.mksinst.com
7. Tomasz G, Anna G-D (2016) MEMS type ionization vacuum sensor. J Sens Actuators A 246:148
8. Bazarchi SM, Sani EA (2017) Micromachined ionization vacuum gauge and improve its sensitivity with magnetic field. Eurasian J Anal Chem 12(7b):1137–1151
9. bazarchi SM, Sani EA (2017) Design and simulation of MEMS type cathodic ray generator. J Eng Appl Sci 12(17):4475–4481
10. Shigemi S, Masahiro H, Tokihiko K (2016) Simulation of relative sensitivity coefficient of bayard-alpert gauge. J Vac Soc Jpn 59(6):156–159
11. Yong-Ki K, Jean-Paul D (2002) Ionization of carbon, nitrogen, and oxygen by electron impact. J Phys Review A 66(1)
12. Itikawa Y (2006) Cross sections for electron collisions with nitrogen molecules. J Phys Chem Ref Data 35(1):31–53
13. David B (2012) Introduction to gas discharges. University of Notre Dame, Notre Dame

Room Temperature Acetone Sensing Based on ZnO Nanowire/Graphene Nanocomposite

Maryam Tabibi, Zahra Rafiee and Mohammad Hossein Sheikhi

Abstract In this paper we report the preparation of a hybrid material by combination of graphene and ZnO nanowire for acetone sensing applications. The ZnO thin films and ZnO NWs were prepared by sol-gel and hydrothermal methods, respectively. The morphological analyses of the obtained material have been performed by means of scanning electron microscopy. These sensors exhibited an enhanced response to acetone concentration as low as 100 ppm at room temperature. The gas sensing analysis of the hybrid material showed that the structure can be used for fabrication of practical sensors.

Keywords ZnO nanowires/graphene nanocomposite · Gas sensor
Acetone

1 Introduction

The detection of acetone vapor is very important in daily life. Medical investigations have shown that the breath of healthy human typically contains acetone vapor less than 0.8 ppm while higher acetone concentration ranges from 1.7 to 3.7 ppm could be detected in breath for those who are diabetic [1, 2]. Moreover acetone vapor is an important chemical material, Thus, recently, the detection of acetone gas has attracted much attention, and kinds of metal oxides have been applied for this purpose. Zinc oxide, as an important functional semiconductor, is one of the most reported materials for gas sensors because of its unique electrical properties, high electrochemical stability and stability to doping [3–5]. Nevertheless; the high operating temperature of metal oxides based sensors implies high power consumption. Thus, a lot of efforts have been made to overcome this problem.

M. Tabibi (✉) · Z. Rafiee · M. H. Sheikhi
Department of Communication and Electronics Engineering,
Shiraz University, Shiraz, Iran
e-mail: m.tabibi@shirazu.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_27

In recent years, there are reports on gas sensor with the capability of room temperature sensing, which is usually realized by hybridizing the metal oxides with carbon nanomaterials like carbon nanotubes [6, 7] or graphene [8, 9]. Recent advances demonstrated that graphene, a two-dimensional sheet of sp^2 bonded carbon atoms on honey comb lattice, is a very promising material for the development of gas sensors operating at room temperature. The high specific surface area of graphene and synergetic effects of the constituent components is usually considered to be the key factor for achieving good gas response at room temperature.

For acetone sensing Zhang et al. reported that ZnO/graphene (ZnO-G) hybrid composites showed enhanced response compared to pure ZnO [10]. However, the operating temperature is still high (280 °C). This might be due to the fact that the morphology, size, porosity, active surface states, face orientation, and growth manner of the ZnO crystal strongly influences the sensing characteristics [11, 12].

Inspired by the advanced physical and chemical properties of ZnO/Graphene composite, in this paper, we prepared ZnO Nanowire/Graphene composite by a facile hydrothermal procedure. During the hydrothermal reaction process, the graphene nanosheets acted as a template for growth of single crystalline ZnO NWs. The surface morphology and crystalline structure of prepared samples were characterized by scanning electron microscope (SEM), and X-ray diffraction (XRD). The responses of synthesized nanocomposites toward acetone were measured at room temperature. At last, the possible gas sensing mechanism of the proposed sensor is also discussed.

2 Experimental

2.1 *Materials*

All chemical reagents were of analytical grade and used without further purification. Graphene Nano plates (99.5%) was purchased from Neutrino, Zinc acetate dehydrate (Zn(CH3COO)2.2H2O, 98%), 2-Methoxyethanol (C3H8O2, 99.8%) Zinc nitrate hexahydrate (H12N2O12Zn, 99%) were purchased from and Sigma-Aldrich. Other reagents including Mono ethanolamine (MEA) (HOCH2CH2NH2, 99%) and Hexamethylenetetramine (C6H12N4, 99%) were commercially available from Merck. More over doubly distilled water was used throughout the experiment.

2.2 *Electrode Preparation*

Au interdigitated electrodes on glass substrate was used for gas sensing measurements. The IDE consist of 37 combs, which the width and the spacing between the neighboring electrodes were 70 μm width with 100 μm, respectively. In order to remove contamination from the electrode surface, IDE was cleaned with acetone.

2.3 Graphene Film Coating Procedure

In the first step, 2 mg of graphene powder was dispersed in distilled water by sonication for 1 h. Afterward 5 μL of the prepared solution dropped on the electrode by micropipette and spin coated in order to make a uniform layer.

2.4 Synthesis of ZnO Seed Layer

ZnO seed layer and ZnO NWs synthesized according to the method reported by Rafiee et al. [13]. In brief, the ZnO solution was prepared using zinc acetate powder, 2-methoxyethanol and MEA as a precursor, a solvent, and the stabilizer, respectively. ZnO sol was prepared first by dissolving zinc acetate (0.2 mol/L) in 20 mL of 2-methoxyethanol. The mixture was then stirred at 1000 rpm on magnetic stirrer. Then MEA was slowly added to the mixture. The molar ratio of zinc acetate to MEA was set at 1:1. Thereafter, the clear and homogenous solution was allowed to age at room for 24 h. ZnO seed solution was then coated on the glass substrate (covered with graphene layer) via spin coating. The ZnO thin film-coated substrate was preheated in order to form nuclei for the crystallization of ZnO thin films. Then, temperature was increased to 150 °C. Next, the sample was cooled down slowly to avoid cracks or dislocations of the structure on the thin films. Finally, the ZnO thin film was annealed at 200 °C for few hours.

2.5 Synthesis of ZnO NWs on Graphene Sheets

In order to growth ZnO NWs, Zn(NO3)2 was used as precursor material and hydrothermal deposition was applied to deposit ZnO NWs on coated substrate. Zn (NO3)2 was mixed with hexamethylenetetramine solution as stabilizer at a concentration of 0.025 mol/L and dissolved in 200 mL distilled water. After stirring the mixture slowly for several minutes, the sample was immersed in the aqueous solution at 70 °C for few hours. Finally, the sample was annealed at 200 °C for few hours for the crystallization of ZnO NWs.

2.6 Gas Sensing Measurements

For electrical measurements, the sensor was placed in a sealed test chamber with a volume of 2500 ml and the defined amount of gas was injected into the chamber. The DC resistance of the sensor was measured by a computer controlled HIOKI multimeter. The response of the gas sensor was defined as: $R = (R_{gas}–R_{air})/$

R_{air} × 100, where R_{gas} is the electrical resistance of the sensor after exposure to gas vapor and R_{air} is the electrical resistances of the sensor in air. The gas sensing measurement was carried out at different acetone concentrations.

2.7 Characterization

Morphology of prepared samples was investigated by a Scanning electron microscopy (SEM, VEGA3 TESCAN). Moreover X-ray diffraction (XRD) was employed to determine the crystallographic structure of synthesized nanocomposite. XRD pattern was recorded by powder diffraction using a D8 ADVANCE type (BRUKER-Germany) with a Cu-Kα source (λ = Cu-Kα 0.1542 nm).

3 Result and Discussion

3.1 SEM Analysis

The surface morphology of graphene sheets was analyzed by SEM as shown in Fig. 1. These nanosheets act as a template for growth of ZnO NWs and also create electron transport channels. SEM image of ZnO NW/Graphene nanocomposite is shown in Fig. 2. Which indicates that ZnO NWs have arranged uniformly on the graphene sheets and contact with each other closely, making the surface of

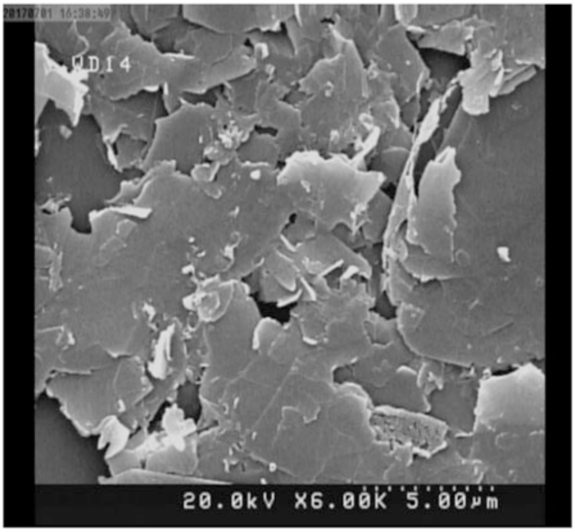


Fig. 1 SEM image of graphene sheets

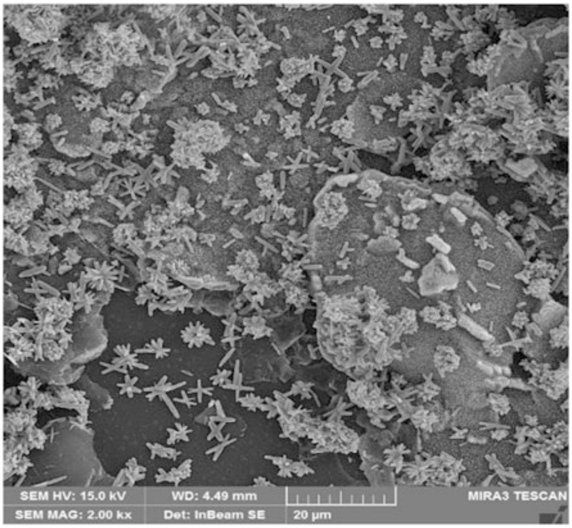


Fig. 2 SEM image of ZnO NW/graphene nanocomposite

graphene unclear and benefits the electron transfer between these two materials. Additionally it can be seen that ZnO NWs exhibit a regular hexagonal shape and diameter and length of NWs can reach to of 200 nm and ~5 μm respectively.

3.2 X-Ray Diffraction Analysis

Figure 3 shows XRD patterns of ZnO NW/Graphene nanocomposites. As shown in Fig. 3 the diffraction peaks matches to hexagonal wurtzite ZnO phase exactly [5]. This pattern confirms the purity of the as prepared ZnO NWs.

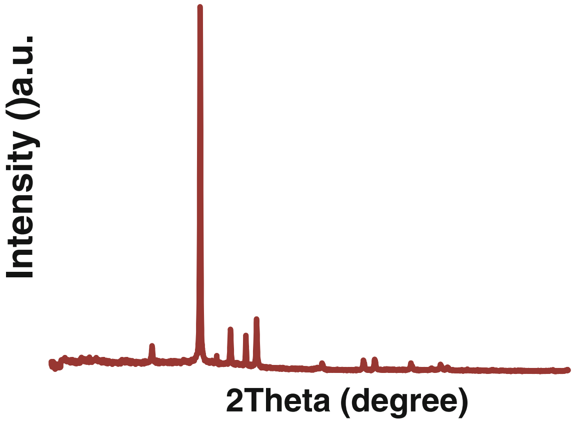


Fig. 3 XRD pattern of ZnO NWs/graphene sample

3.3 Gas Sensing Analysis

Figure 4 shows the real-time resistance measurement of the fabricated sensor based on pure graphene and ZnO NWs/graphene nanocomposite exposed to different concentration of acetone vapor. As shown in Fig. 4 the ZnO/Graphene shows higher sensitivity and much faster response toward acetone in the entire range of detected concentration compared to pure graphene. The response and recovery times of the nanocomposite are around 2.5 and 2 min respectively. Figure 5 shows the response of both sensors operated at 25 °C versus different concentration of acetone. Both sensors possess a linear response characteristic towards acetone concentrations from 100 to 700 ppm.

Fig. 4 Response transients of pure graphene and ZnO NW/Graphene to acetone vapor at 25 °C

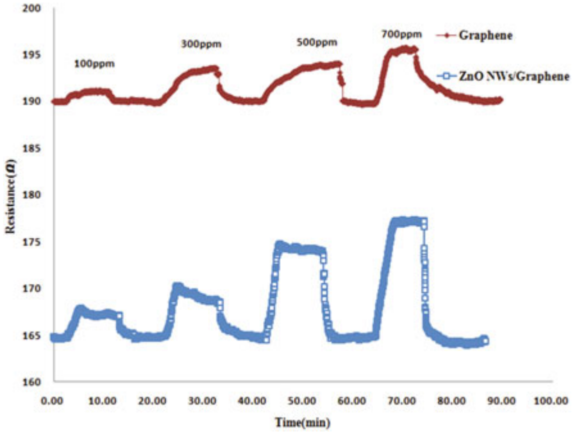
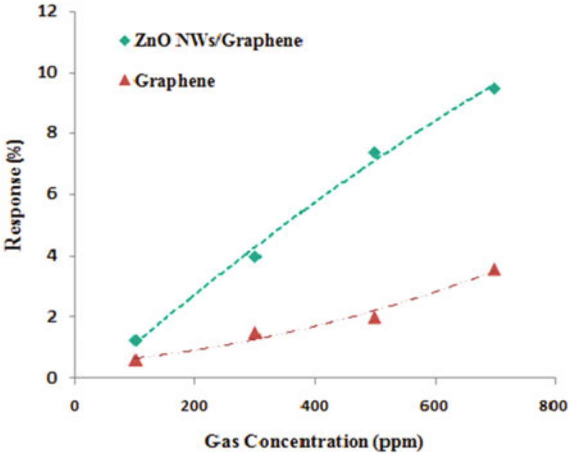


Fig. 5 Response curve of sensors to different acetone concentrations at 25 °C



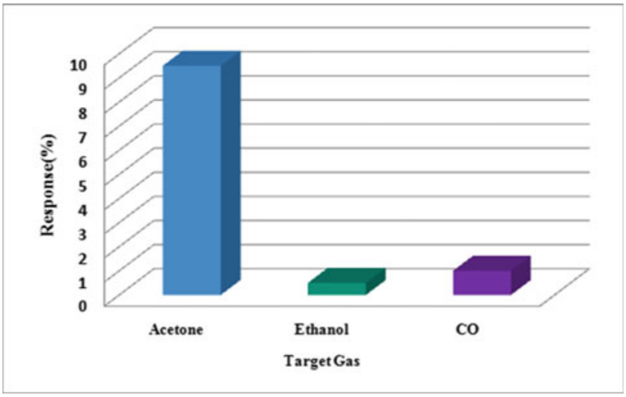


Fig. 6 The selectivity of ZnO NWs/Graphene sensor toward 7500 ppm concentration of different gases and 700 ppm acetone at 25 °C

To evaluate the selectivity of the sensor the response comparison of the sensor to different gases is also investigated. Figure 6 shows the response of the sensor under exposure of 700 ppm of acetone vapor, 7500 ppm of carbon monoxide and ethanol. As shown in Fig. 6 the response to acetone vapor is the highest compared with other gases, indicating good selectivity to acetone vapor.

In recent years a variety of acetone sensors have been successfully fabricated. Table 1 presents the response, measurement range and operating temperature of the prepared sensor with previous works based on ZnO nanostructures made by different methods [15–19]. As shown from Table 1, most of the ZnO based acetone sensors are operated at a high temperature which is an obstacle for their practical use. All these observations indicate that ZnO NW/Graphene nanocomposite is a good candidate for development of room temperature acetone sensor.

3.4 *Gas Sensing Mechanism*

To gain a better understanding of the sensing mechanism, we first measured the base resistance of ZnO NW/Graphene nanocomposite as well as pure graphene and as shown in Fig. 4. The ZnO NW/Graphene nanocomposite exhibited better electrical conductivity than graphene which confirms that the ZnO NWs react as a conductive network and provide channels for the electron transfer thus reduce the resistance of graphene. During the sensing process, The O_2 molecules are firstly ionized on the surface of ZnO NW/Graphene nanocomposite and form O^{-2} and O_2^- species. Then the target acetone molecules are directly oxidized by chemisorbed oxygen species. Graphene plays the major role in contribution of response of the

Table 1 Comparison of sensing performances of our proposed acetone sensor with other published acetone sensors based on ZnO

Material	Synthesize method	Concentration (ppm)	Operating temperature (°C)	Refs.
Hierarchically porous ZnO microstructures	Aqueous solution route	5–400	330	[14]
Hybrid ZnFe2O4/ZnO hollow spheres	Hydrothermal	10–200	280	[15]
Nb decorated ZnO	Thermal Oxidation for ZnO nanostructures/Nb coating by DC pulse sputtering	50–1000	400	[16]
ZnO nanosheets	Direct precipitation method/ calcination	5–1000	300	[17]
Porous spheres-like ZnO nanostructure	Electrospinning method/ calcination	2–500	310	[18]
ZnO NW/ Graphene nanocomposite	**Hydrothermal**	**100–700**	**Room temperature**	**This work**

sensor because acetone, a reducing gas, has a lone electron pair that can be easily donated to the p-type graphene sheets, resulting in the increase of the resistance of the graphene based devices [19], which leads to the resistance change of the hybrid nanocomposites. Therefore, a possible mechanism for the enhanced sensing properties of the hybrid nanocomposites is that the introduction of ZnO NWs led to formation of homogeneous composition with 3D nanostructure featured with much higher surface accessibility and conductivity, which can promote the transfer of electrons. As a result, the sensor shows excellent properties of acetone detection [20].

4 Conclusions

In summary, we fabricated a composite material based on graphene and ZnO nanostructure. The gas sensing properties of the obtained nanocomposites were investigated by exposing them to acetone vapor. The composite material showed better gas sensing performance compared to the pure graphene nanostructures for acetone at room temperature. The preliminary results show that ZnO NW/Graphene is promising structures for the development of acetone gas sensing devices.

References

1. Chen J, Pan X, Boussaid F, McKinley A, Fan Z, Bermak A (2017) Breath Level Acetone Discrimination Through Temperature Modulation of a Hierarchical ZnO Gas Sensor. Sens Lett IEEE, vol 1, pp 1–4, ISSN 2475–1472
2. Righettoni M, Tricoli A, Pratsinis SE (2010) Si: WO_3 Sensors for highly selective detection of acetone for easy diagnosis of diabetes by breath analysis. Anal Chem 82(9):3581–3587
3. Joshi RK, Hu Q, Alvi F, Joshi N, Kumar A (2009) Au decorated zinc oxide nanowires for CO sensing. J Phys Chem C 113(36):16199–16202
4. Tee S et al (2016) Microwave-assisted hydrolysis preparation of highly crystalline ZnO nanorod array for room temperature photoluminescence-based CO gas sensor. Sens Actuators B Chem 227:304–312
5. Galstyan V, Comini E, Baratto C, Faglia G, Sberveglieri G (2015) Nanostructured ZnO chemical gas sensors. Ceram Int 41(10):14239–14244
6. Evans GP et al (2018) Room temperature vanadium dioxide-carbon nanotube gas sensors made via continuous hydrothermal flow synthesis. Sens Actuators B Chem 255:1119–1129
7. Zhao Y et al (2018) Outstanding gas sensing performance of CuO-CNTs nanocomposite based on asymmetrical schottky junctions. Appl Surf Sci 428:415–421
8. Song Z et al (2016) Sensitive room-temperature H_2S gas sensors employing SnO_2 quantum wire/reduced graphene oxide nanocomposites. Chem Mater 28:1205–1212
9. Zhang H, Feng J, Fei T, Liu S, Zhang T (2014) SnO_2 nanoparticles-reduced graphene oxide nanocomposites for NO_2 sensing at low operating temperature. Sens Actuators B 190:472–478
10. Zhang H, Cen Y, Du Y, Ruan S (2016) Enhanced acetone sensing characteristics of ZnO/Graphene composites. Sensors 16:1–10
11. Mirabbaszadeh K, Mehrabian M (2012) Synthesis and properties of ZnO nanorods as ethanol gas sensors. Phys Scr 85:035701
12. Tian S, Yang F, Zeng D, Xie C (2012) Solution-processed gas sensors based on ZnO nanorods array with an exposed (0001) facet for enhanced gas-sensing properties. J Phys Chem 116:10586–10591
13. Rafiee Z, Mosahebfard A, Sheikhi MH (2017) Synthesis and preparation of ZnO NWs for glucose biosensing. In: Iranian conference on electrical engineering (ICEE), 2017, pp 455–460
14. Ge M, Xuan T, Yin G, Lu J, He D (2015) Chemical controllable synthesis of hierarchical assembled porous ZnO microspheres for acetone gas sensor. Sens Actuators B Chem 220:356–361
15. Deng J, Wang L, Zhang R, Zhang T, Zhou T, Lou Z (2016) Fast and real-time acetone gas sensor using hybrid $ZnFe_2O_4$/ZnO hollow spheres. RSC Adv 6:66738–66744
16. Wongrat E, Chanlek N, Chueaiarrom C, Thupthimchun W (2017) Acetone gas sensors based on ZnO nanostructures decorated with Pt and Nb. Ceram Int 43(5):S557–S566
17. Li S, Zhang L, Zhu M, Ji G, Zhao L, Yin J (2017) Acetone sensing of ZnO nanosheets synthesized using room-temperature precipitation. Sens Actuators B Chem 249:611–623
18. Li XB et al (2013) Porous spheres-like ZnO nanostructure as sensitive gas sensors for acetone detection. Mater Lett 100:119–123
19. Hu N, Yang Z, Wang Y, Zhang L, Wang Y, Huang X et al (2014) Ultrafast and sensitive room temperature NH_3 gas sensors based on chemically reduced graphene oxide. Nanotechnology 25:025502
20. Wang T et al (2017) Studies on NH_3 gas sensing by zinc oxide nanowire-reduced graphene oxide nanocomposites. Sens Actuators B Chem 252:284–294

Application of Learning Methods for QoS Provisioning of Multimedia Traffic in IEEE802.11e

Hajar Ghazanfar, Razieh Taheri and Samad Nejatian

Abstract In recent years, with increasing demands of real-time multimedia service, a lot of attention has been given to quality of service(QoS) support in wireless network. To maintain QoS of multimedia services we need to control packet loss, delay and packet delivery ratio. But it is difficult to guarantee QoS especially when the network is overloaded. The proposed scheduler algorithm called EQQ (Enhanced QoS with Q-Learning)guarantees QoS by using the kind of reinforcement learning with changing suitably service intervals and transmission opportunities. Simulation result show that the EQQ is superior to the older algorithm in parameter of delay and packet loss and packet delivery ratio.

Keywords QoS · IEEE802.11 · HCCA · EDCA · TXOP

1 Introduction

One of the drawbacks of wireless networks in comparison to wired networks is that they are generally less efficient and unpredictable. Wireless has limited bandwidth, high packet overheads, and is more prone to environmental factors such as obstructions, interference, weather and so on. The wireless medium (air) is much harder to control than a physical wire. The WLAN medium is also unlicensed and is therefore subject to interference from other devices. To further compound the

H. Ghazanfar (✉)
Department of Theory of Computation,
Technological Institute of Zahedshahr, Shiraz, Iran
e-mail: hajar.ghazanfar@gmail.com

R. Taheri
Department of Theory of Computation, Technological Institute of Darion, Shiraz, Iran
e-mail: taheri_r80@yahoo.com

S. Nejatian
Department of Theory of Computation, Technological Institute of Yasuj, Yasuj, Iran
e-mail: s_nejatian@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_28

problem, wireless devices are generally constrained by size, weight and battery size, limiting the processing power and the battery life. These factors further limit the capability of the network to provide an optimal solution. The main objective of WLAN QoS is to optimize use of limited bandwidth offered by a WLAN to address the issues noted above. To optimize the best use of the resources and fulfill the resource requirements of different applications, QoS provides mechanisms to control access and usage of the medium based on the application. Each application has different needs in terms of latency, bandwidth and packet-error rate and, therefore, QoS must cater to each of these needs. Applications requiring low latency (e.g., voice) may be given higher priority to use the medium, whereas applications requiring higher bandwidth may be assigned longer transmit times (e.g., video). Other traffic may require high reliability (e.g., email and data) and must be delivered with low packet-error rate. The original 802.11[1] standard was not designed to provide differentiation and prioritization based on the traffic type, thus providing less than optimal user experience for voice and video over WLAN applications. Voice applications require no dropped calls or bad connections. Video/audio applications require enough bandwidth to maintain high quality video/audio streams. Email and file-sharing applications require ensuring delivery of error-free files. To fulfill these requirements, the upcoming IEEE 802.11e standard will add several QoS features and enhancements to WLAN. The key benefits of the 802.11e standard are:

- Reduces latency by prioritizing wireless packets based on traffic type.
- Enables Access Point (AP) to schedule resources based on client/station data rate and latency needs.

2 Original 802.11 MAC

The original 802.11 MAC does not provide differentiated services based on traffic type. However, as wireless networks provide multimedia services involving voice and video, high packet overhead with limited bandwidth in a WLAN can become a major stumbling block for delivering delay-sensitive packets. In the original 802.11 standard, as much as one third of the data rate can be consumed by packet fragmentation, inter-frame spacing and acknowledgments. Furthermore, under heavy traffic load conditions, collisions and backoffs can severely deteriorate the quality of voice and video applications.

[1] IEEE Std 802.11-1997—Information Technology—Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-specific Requirements-part 11: Wireless LAN Medium Access Control (MAC) And Physical Layer (PHY) Specifications.

The 802.11 standard specifies two channel access mechanisms: Distributed Coordination Function (DCF) and Point Coordination Function (PCF). DCF allows sharing of the wireless medium between the Stations (STAs) and the AP using Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). DCF provides best-effort service and does not provide either medium access priority or support delay and bandwidth requirements of different applications. DCF mode operates as follows. Each station checks whether the medium is idle before transmitting. If the medium is detected to be idle for Distributed Inter-frame Space (DIFS) interval of time, the station begins transmission. In the case where the medium is determined to be busy, the station defers transmission until the medium is idle for DIFS time. The station then selects a random backoff interval (using a backoff algorithm) and decrements a backoff counter while the medium is idle. The backoff mechanism is used to prevent two or more stations from transmitting simultaneously. Once the backoff interval has expired, the station begins its transmission. The range from which the random backoff interval is selected is called the Contention Window (CW) and depends on the number of previous retransmission attempts. Once the MAC Service Data Unit (MSDU) has been transmitted, the station waits for a Short Inter-Frame Space (SIFS) duration for the Acknowledgment (ACK) from the recipient. PCF is an optional channel-access mechanism in the 802.11 standard that is not commonly implemented due to lack of market demand. PCF provides contention-free access to the medium. It was designed to support time-sensitive applications. The Point Coordinator (PC) residing in the AP provides a contention-free wireless medium access. A polling method is used to provide access, with the PC acting as the polling master. This eliminates collisions and the time spent on backoff and contention as described previously for DCF. Contention-free access to the medium is not provided at all times. When PCF is used, time on the medium is divided into a Contention-Free Period (CFP) and a Contention Period (CP) by the PC. During CFP and CP, PCF and DCF are used to access the medium, respectively. Neither DCF nor PCF has sufficient functionality to provide QoS demanded by multimedia applications. DCF treats all traffic the same—with all stations contending for the medium with the same priority. PCF also has several inadequacies in the support of QoS:

- Lack of mechanisms to differentiate different traffic types.
- No mechanisms for the stations to communicate their QoS requirements to the AP.
- No management interface to control and setup CFP.
- The polling schedule is not tightly controlled.

In the next section, we discuss how the 802.11e standard addresses these drawbacks of 802.11 to provide QoS.

3 IEEE 802.11e Standard

The IEEE 802.11 Task Group E (802.11e) has defined enhancements to the original 802.11 MAC (Medium Access Control) to provide QoS. The 802.11e standard introduces the Hybrid Coordination Function (HCF), which combines functions from DCF and PCF with enhanced QoS-specific mechanisms and frame types. HCF has two modes of operation—Enhanced Distribution Coordinate Access (EDCA) and HCF Controlled Channel Access (HCCA). EDCA and HCCA are contention-based and polling-based mechanisms for channel access, respectively, and operate concurrently. During the CP, EDCA is used for channel access; whereas during CFP, HCCA is mostly used. A Station (STA) that supports QoS is referred to as a QoS Enhanced Station (QSTA); whereas an Access Point (AP) that supports QoS is referred to as a QoS Enhanced AP (QAP). HCF allocates QSTAs the right to transmit through Transmission Opportunity (TXOP). A TXOP defines the start time and the maximum duration during which a QSTA can transmit a series of frames. A summary of EDCA and HCCA mechanisms are provided in the sections of this paper titled "Enhanced Distribution Coordinate Access (EDCA)" and "HCF Controlled Channel Access (HCCA)," respectively.

4 Enhanced Distribution Coordinate Access (EDCA)

EDCA contention access is an extension to DCF and provides prioritized access to the wireless medium. The EDCA channel-access mechanism defines four Access Categories (AC) based on the IEEE 802.1D standard[2] to provide priorities. Each AC has its own transmit queue. The following four key parameters are used for differentiation:

- Minimum contention window size (CWMin). AC with higher priority is assigned a shorter CWMin.
- Maximum contention window size (CWMax).
- TXOP limit—Specifies the maximum duration a QSTA can transmit and is specified per AC. The TXOP limit can be used to ensure that high-bandwidth traffic gets greater access to the medium. TXOP limit also makes the channel access protocol significantly more efficient.
- Arbitration Inter-Frame Space (AIFS)—specifies the time interval between the wireless medium becoming idle and the start of channel-access negotiation. Each AC is assigned a different AIFS[AC] based on the AC to further provide QoS differentiation. Each AC contends independently for TXOPs based on the above parameters within the QSTA. Once the AC has sensed that the medium

[2] IEEE Std 802.1D-2004 (Revision of IEEE Std 802.1D-1998)—IEEE Standard for Local and metropolitan area networks Media Access Control (MAC) Bridges.

has been idle for AIFS[AC], it starts its backoff time (similar to DCF). If there is a collision between ACs within a QSTA, data frames from the AC with the highest priority receives a TXOP. Data frames from the remaining ACs behave as if there was an external collision.

5 HCF Controlled Channel Access (HCCA)

HCCA uses a Hybrid Coordinator (HC) to centrally manage the wireless medium access to provide parameterized QoS. Parameterized QoS refers to the capability of providing QoS flows from applications with specific QoS parameters—such as date rate, latency and so forth. Like PCF, HCCA uses a polled-based mechanism to access the medium, thereby reducing contention on the wireless medium. The key differences between HCCA and PCF are that HCCA can poll the stations during CP and that it supports scheduling of packets based on QSTA's specific traffic-flow requirements. The traffic-flow requirements of the QSTAs are specified using Traffic Specifications (TSPECs) discussed in the section of this paper titled "Traffic Specifications (TSPECS)." The HC has the highest priority over all QSTAs in gaining access to the medium as it has the shortest waiting time compared to all QSTAs backoff times. The HC provides contention-free frame exchange with short delays, thereby providing tighter controlled latency.

6 Traffic Specifications (TSPECs)

Because WLANs have limited bandwidth and provide contention-based medium access, they are susceptible to traffic congestion, which can lead to severe overall network performance degradation. As the networks become overloaded, CW sizes can increase significantly, leading to long backoff times. This can become a network performance bottleneck as the limited bandwidth of the network is not fully utilized. This necessitates some admission-control mechanism to be built into the standard for regulating traffic. The IEEE 802.11e standard specifies the use of TSPECs for negotiating admission control for both EDCA and HCCA TSPECs are used by QSTAs to specify traffic-flow requirements such as data rate, delay, packet size and service interval. The QAP may accept or reject a new TSPEC request based on the network conditions. If a TSPEC is rejected by the QAP, high-priority AC inside the requesting QSTA is not allowed to use high-priority access parameters.

7 802.11e MAC Enhancements

Apart from providing EDCA and HCCA functionality discussed previously, the 802.11e standard provides several MAC enhancements. Some of the key enhancements are summarized next.

8 Contention Free Bursts

Contention Free Bursts (CFB) allow a QSTA/QAP to send several frames in a row without having to contend for the medium again and again. The QSTA/QAP continues to transmit after the SIFS delay if there is time remaining in a granted TXOP. CFBs may be used during TXOPs that were gained during EDCA or HCCA. Figure 1 shows data transmission with and without CFB. Bursting can significantly improve performance as the overheads associated with DIFS and backoff are reduced.

CFB can also be used to improve 802.11g's throughput in a mixed 802.11b and 802.11g environment as it frees time for 802.11g STAs. A fast 802.11g STA can transmit several frames in the time period that a slower 802.11b STA takes to

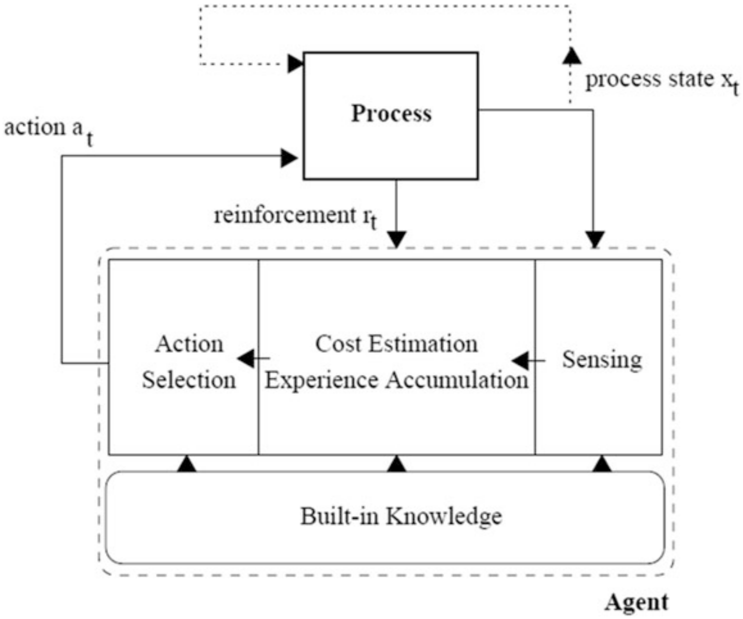


Fig. 1 Interaction between agent and dynamic process in Reinforcement Learning

transmit one frame and its associated overhead. An 802.11g STA can be assigned a TXOP comparable to the single-frame duration of a STA that uses 802.11b.

Some vendors have implemented proprietary packet-bursting schemes. The 802.11e standard enables a standardized method of implementing bursting, providing interoperability across multiple vendors' network gear.

9 Introduction of Reinforcement Learning

Reinforcement Learning (RL) is learning through direct experimentation. It does not assume the existence of a teacher that provides examples upon which learning of a task takes place. Instead, in RL experience is the only teacher. With historical roots on the study of conditioned reflexes, RL soon attracted the interest of Engineers and Computer Scientists because of its theoretical relevance and potential applications in fields as diverse as Operational Research and Robotics.

Computationally, RL is intended to operate in a learning environment composed by two subjects: the learner and a dynamic process. At successive time steps, the learner makes an observation of the process state, selects an action and applies it back to the process. The goal of the learner is to find out an action policy that controls the behavior of this dynamic process, guided by signals (reinforcements) that indicate how well it is performing the required task. These signals are usually associated to some dramatic condition—e.g., accomplishment of a subtask (reward) or complete failure (punishment), and the learner's goal is to optimize its behavior based on some performance measure (a function of the received reinforcements). The crucial point is that in order to do that, the learner must evaluate the conditions (associations between observed states and chosen actions) that lead to rewards or punishments. In other words, it must learn how to assign credit to past actions and states by correctly estimating costs associated to these events. Starting from basic concepts, this tutorial presents the many flavors of RL algorithms, develops the corresponding mathematical tools, assess their practical limitations and discusses alternatives that have been proposed for applying RL to realistic tasks, such as those involving large state spaces or partial observability. It relies on examples and diagrams to illustrate the main points, and provides many references to the specialized literature and to Internet sites where relevant demos and additional information can be obtained.

9.1 Reinforcement Learning Agents

The learning environment we will consider is a system composed by two subjects: the learning *agent* (or simply the *learner*) and a dynamic *process*. At successive time steps, the agent makes an *observation* of the process state, selects an action and applies it back to the process, modifying the state. The goal of the agent is to find out adequate actions for controlling this process. In order to do that in an *autonomous* way, it uses a technique known as *Reinforcement Learning*. Reinforcement Learning (RL for short) is learning through direct experimentation. It does not assume the existence of a teacher that provides 'training examples'. Instead, *in RL experience is the only teacher*. The learner acts on the process to receive signals (reinforcements) from it, indications about how well it is performing the required task. These signals are usually associated to some dramatic condition—e.g., accomplishment of a subtask (reward) or complete failure (punishment), and the learner's goal is to optimize its behavior based on some performance measure (usually minimization of a *cost function*[3]). The crucial point is that in order to do that, in the RL framework the learning agent must *learn* the conditions (associations between observed states and chosen actions) that lead to rewards or punishments. In other words, it must learn how to *assign credit* to past actions and states by correctly estimating costs associated to these events. This is in contrast with supervised learning (Haykin 1999), where the credits are implicitly given beforehand as part of the training procedure. *RL agents are thus characterized by their autonomy*.

Figure 1: Interaction between agent and dynamic process in Reinforcement Learning. The agent observes the states of the system through its sensors and chooses actions based on cost estimates which encode its cumulated experience. The only available performance signals are the reinforcements (rewards and punishments) provided by the process.

10 Proposed Scheme

In this section I introduce EQQ algorithm. This algorithm is combined of two major algorithm. first is scheduling algorithm and second is reinforcement algorithm.

The parameters that should be compute in scheduling algorithm is TXOP and service interval(SI).

This two parameters should be found by agent in the environment and select the best of them.

10.1 Analysis of Scheduling Algorithm in EQQ

In EQQ each station according to training in TXOP interval can transfer their packet. These parameters are:

[3] Unless otherwise stated, minimization of a cost function is used throughout this text. Naturally, maximization of a reward function could be similarly used.

The main asset of the scheduler proposed by Grilo [1] is that it extends the functionality of the Simple Scheduler by allowing the HC to:

(i) allocate TXOPs of variable length (instead of fixedTXOPs)
(ii) poll each STA at variable and different service intervals (instead of polling ALL STAs with period SI) In order to guarantee that the average duration of TXOPs allocated to a STA will remain equal to TDj, the scheduler utilizes a TXOP timer (that is in fact equivalent of a token bucket of time units). This timer increases at a constant rate equal to TDj/mSIj which is the maximum fraction of time that the STA can spend in polled TXOPs (mSIj is the minimum interval between to successive polls and TDj the average TXOP duration). The maximum value of the TXOP timer is set to $\mathrm{MTD_j}$. Each time the HC allocates a TXOP, the duration of the TXOP is deducted from the TXOP timer of the STA. The STA can be polled again only when the value of the TXOP timer is greater or equal to $\mathrm{mTD_j}$ so as to guarantee the transmission of at least one MSDU. The scheduler uses the Scheduling Based on Estimated transmission Times–EQQ algorithm to determine the STA that should be polled at each time instant. Obviously, if a STA j is polled at time t the next poll should be issued on a time t' that satisfy the relation:

$$\mathrm{t + mSI_j < = t' < = t + MSI_j} \tag{10.1}$$

The (t + MSI) limit is also set as the deadline for the EQQ algorithm and the MSIj is calculated as:

$$\mathrm{MSI_j} = \beta * \left(\mathrm{D - MTD_j}\right) \quad \text{with } 0 < \beta < 1 \tag{10.2}$$

where D is the lowest delay bound among the STA's TSPECs.

$$D = \min D, \quad \mathrm{i = 1...n}$$

Minimum TXOP duration (*mTD*): This is the minimum TXOP duration that can be allocated to a STA and equals the maximum packet transmission time for any of the STA's TSPECs. Thus the *mTD* is calculated as:

$$\mathrm{mTD_i} = \max \frac{\mathrm{M_{ij}}}{\mathrm{R_{ij}}}, \quad \mathrm{j\varepsilon(1, n_i)} \tag{10.3}$$

Maximum Service Interval (MSI): It is the maximum time interval allowed between the start of two successive TXOPs allocated to a STA (in μs). [802.11e] does not provide any guidelines for calculating MSI, however a reasonable assumption is that MSI should be related to the lowest delay bound D of all STA's TSPECs, hence it should held that:

$$MSI \leq D - MTD_i \tag{10.4}$$

The most important factor of EQQ is the usage of queue size (QS) of each station. QS is the new parameter in 802.11ethat added to frame.

10.2 Evaluation of Reinforcement Learning Used in EQQ

EQQ use Q-Learning to train agent search table is estimate function that EQQ use it. This table has two dimension. First is for state and second is for action of each state.

According to optimize result for each state in order to warranty QoS in EQQ algorithm, until to achieve this result the action should be update. Q-Learning in EQQ mapping (act/state)to Q-Value. Q is the optimized value that is the summation of all reward that agent achieve in the state with its policy.

In EQQ primary value of Q related TXOP of each station and SI assigned zero. After training of agent in EQQ update stage should be started for Q value. If Q increase the system consider this value as reward and otherwise consider as punish.

Summation of all reward and punish guide agent to next station. After training stage of agent the highest value of Q indicate optimize action and optimize policy.

Process Of agent training include this stage:

1. Q-Value of pair(act/state)assign to zero. $Q(I, a) \leftarrow 0$

I describe state, a describe an action. number of iteration assign to one. If k means iteration $k \leftarrow 1$ and ρ^k means the average of rewards in k iteration assumption that start in state I, A(i) means summation of all action in state I and α^k means learning rate in the k-iteration. $ß^k$ means average of learning rate. in Q-learning for the optimize result this two parameters assigns:

$$\alpha = \log(k)/k \quad \text{where}\, k \rightarrow \infty \tag{10.5}$$

$$\beta = 90/(100 + k) \tag{10.6}$$

$$\begin{gathered}\text{total}-\text{reward} \leftarrow 0\\ \text{total}-\text{time} \leftarrow 0\end{gathered}$$

suppose the next state is j therefore r(I, j) shows the reward of transfer from state I to state j and t(I, a, j) shows the transfer time, therefor Q(I, j) is:

$$Q(I,j) \leftarrow \left(1 - a^k\right)Q(I,a) + \alpha^k\left[r(I,a,j) - \rho^k t(I,a,j) + Q_{max}\right] \tag{10.7}$$

Q_{max} is maximum of the Q in j state, α^k is calculated from (10.5), and two parameters of total-reward and total-time has updated with below equations:

$$\mathrm{Total{-}reward} \leftarrow \mathrm{total{-}reward} + \mathrm{r(I, a, j)} \tag{10.8}$$

$$\mathrm{Total{-}time} \leftarrow \mathrm{total{-}time} + \mathrm{t(I, a, j)} \tag{10.9}$$

Therefore, average of reward calculated as:

$$\rho^{k+1} \leftarrow \left(1 - \beta^{k}\right)\rho^{k} + - \beta^{k}\left[\frac{total{-}reward}{total{-}time}\right] \tag{10.10}$$

This stages are used for training agent, EQQ use this equations for achieve TXOP and SI in order to optimizes QoS.

11 Results

This parameters used for ns2 simulation (Tables 1 and 2).

Table 1 Parameters of station

Parameters	Value
Slot time	20 s
SIFS	10 s
PIFS	30 s
Preamble length	144 bits
PLCP header length	48 bits
PLCP TX rate	1 Mbps
MAC header length	60 Bytes
Basic Tx rate	1 Mbps
Date rate	11 Mbps

Table 2 Characters of traffic stream

TSPEC	CBR traffic	VBR traffic
Mean date rate	83 Kbps	128 Kbps
Peak date rate	83 Kbps	1.7 Mbps
Nominal MSDU size	208 Bytes	1300 Bytes
Maximum MSDU size	208 Bytes	5211 Bytes
Maximum burst size	576 Bytes	5211 Bytes
Minimum PHY rate	11 Mbps	11 Mbps
Maximum service interval	30 ms	40 ms
Delay bound	60 ms	80 ms

In this simulation we have two scenario to simulate. one is for measure the CBR traffic that use VOIP and the other is for VBR traffic that use MPEG4. We show the result of EQQ with throughput, average delay, packet loss ratio, jitter. This parameters are compared with reference algorithm (Diagrams 1, 2, 3, 4 and 5).

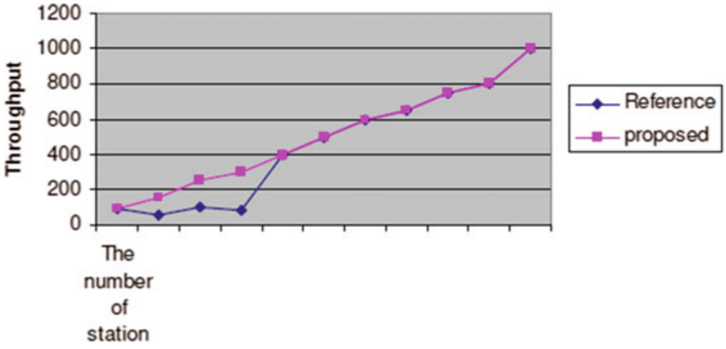


Diagram 1 Result of throughput in CBR traffic

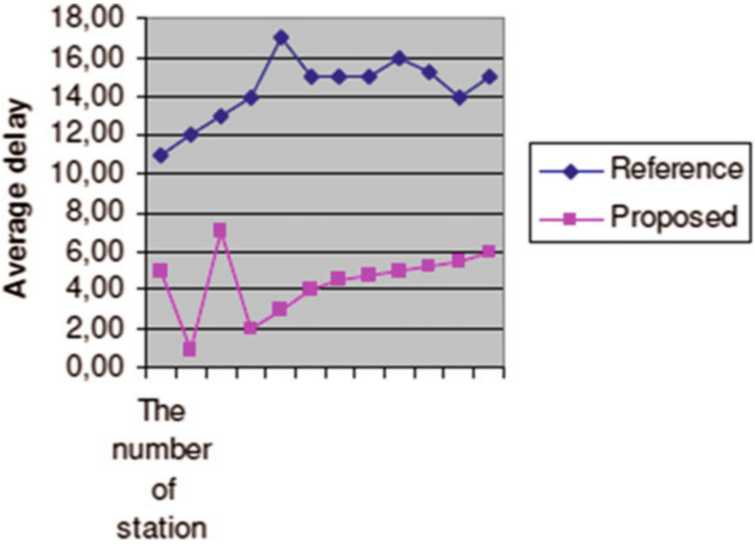


Diagram 2 Result of average delay in CBR traffic

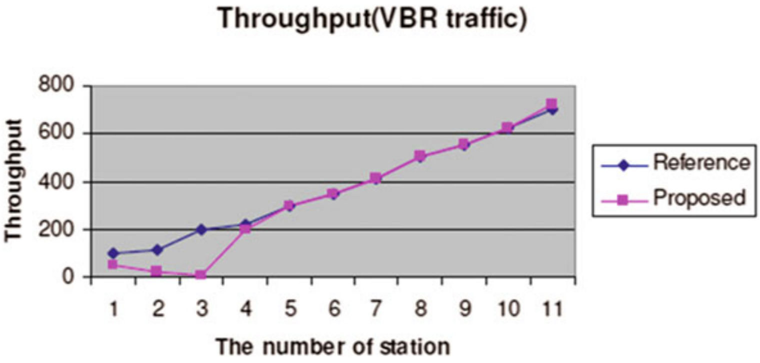


Diagram 3 Result of throughput in VBR traffic

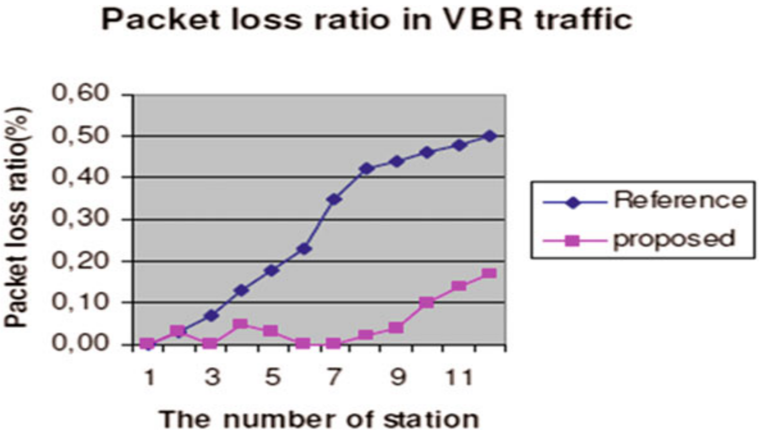


Diagram 4 Result of packet loss ratio in VBR traffic

12 Conclusions

In this paper, we have proposed the Enhanced QoS with Q-Learning for solving Quality of Service problems in networks in order to support multimedia services. Our resource management scheme envisages an AP as an agent in each cell that is able to perform adaptive scheduling on the basis of channel state estimate for each user. The multi-objective QoS scheduler has been designed in the RL framework using a linear combination of specific sub-tasks, and filtering each QoS measure by an appropriate function. Suitable performance metrics have been identified in terms

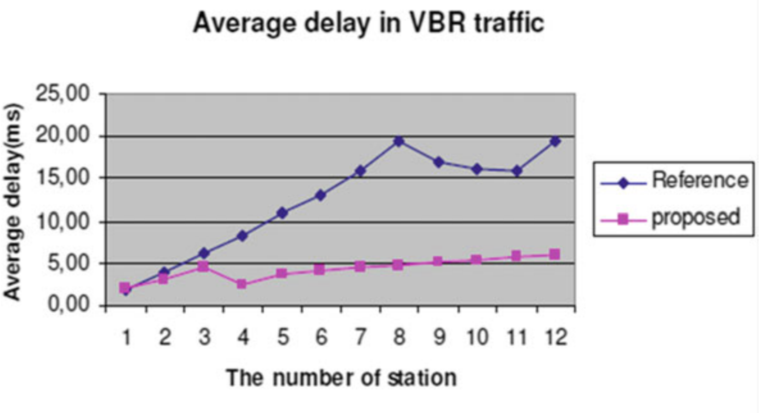


Diagram 5 Result average delay in VBR traffic

of packet loss rate for real-time traffic in one scenarios, namely Reference algorithm. Simulation runs have shown that the EQQ algorithm outperforms more better than Reference algorithm. Moreover, interesting results have been also obtained in the EQQ scenario. In conclusion, the results of this paper permit to emphasize the potentialities of Q-Learning and, more generally, of learning machines in supporting complex scheduling tasks in network.

References

1. Grilo A, Macedo M, Nunes M (2003) A scheduling algorithm for QoS support in IEEE 802.1 le networks. IEEEWirel Commun 10(3)
2. Bertsekas DP (2011) Dynamic programming and optimal control, vol 1. Athena Scientific, Belmont, Massachusetts. Bertsekas DP, Tsitsiklis JN (2013) Neuro-Dynamic Programming. Athena
3. Bertsekas DP, Tsitsiklis JN (2013) Neuro-dynamic programming. Athena Scientific, Belmont, Massachusetts
4. Blake A, Yuille A, (eds) (2015) Active vision. MIT Press, Cambridge
5. Boyan JA, Moore AW (2012) Generalization in reinforcement learning: safely approximating the value function. In: Tesauro G, Touretzky DS, Leen TK (eds) Advances in neural information processing systems 7. MIT Press
6. Brooks RA (2014) Elephants don't play chess. In: Maes P (ed) Designing autonomous agents. MIT Press, Brooks RA, pp 3–15
7. Mataric MJ (2016) Real robots, real learning problems. In Connell JH, Mahadevan S (eds) Robot learning, Chapter 8. Kluwer Academic Publishers, pp 193–213
8. Chapman D, Kaelbling LP (2016) Input generalization in delayed reinforcement learning: an algorithm and performance comparisons. In: Proceedings of the international joint conference on Artificial Intelligence (IJCAI'91), pp 726–731
9. Wischhof L, Lockwood J Packet scheduling for link-sharing and quality of service support in wireless local area networks. Technical Report WUCS-01-35, Applied Research Laboratory, Washington University in St. Louis, November 2010

10. Verma D, Zhang H, Ferrari D, Delay jitter control for real-time communication in a packet switching network. In: Proceedings of TriCom2012
11. Fukunaga K (1990) Statistical pattern recognition. Academic Press, San Diego
12. Brady PT (2013) A model for on-off speech patterns in two-way conversations. Bell Syst Tech J 48:2445–2247
13. Casals O, Blondia C (2012) Performance analysis of statistical multiplexing of VBR sources. In: Proceedings of INFOCOM'92, pp 828–838
14. Oritagoza-Guerrero L, Aghvami AH (1998) A distributed dynamic resource allocation for a hybridTDMA/CDMA system. IEEE Trans Veh Technol 47:1162–1178

LUT Design with Automated Built-in Self-test Functionality

Hanieh Karam and Hadi Jahanirad

Abstract Nowadays, FPGAs play a significant role in industrial applications. Therefore, ensuring their proper performance is of great importance. In this paper, an automatic Built-in self-test core has been designed in the LUT to test and investigate errors. In order to verify this method, powerful software H-Spice with 45 nm precision in transistor level has been used. The advantages of this method include being automatic, detecting various errors, preserving initial information, and reducing hardware overhead compared to the previous methods.

Keywords FPGA · Internal testing · BIST · Built-in self-test · LUT

1 Introduction

FPGA (Field Programmable Gate Array) chips have wide application in digital system. They can be reconfigured to implement different logic circuits in a short time. Shrinking the feature size of transistors in modern CMOS technologies imposes many challenges for detection and diagnosis of large amount of defects in FPGA chips. Elapsed time and cost of test operation are two main factors for each testing scenario [1].

FPGA chip generally constitutes a 2-D array of Configurable Logic Blocks (CLB) surrounded by Input-Output (IO) blocks. Horizontal and vertical channels are used to interconnect the above resources. Switch Blocks (SB) which lay in cross-sections of horizontal and vertical channels can be configured to make required routing properly [2] (Fig. 1).

CLB consists of some Look-Up Tables (LUT), DFF and some multiplexers. General CLB architecture has been illustrated in Fig. 2.

H. Karam (✉) · H. Jahanirad
Department of Electrical Engineering, University of Kurdistan, Sanandaj, Kurdistan, Iran
e-mail: Haniyeh.karam@gmail.com

H. Jahanirad
e-mail: h.jahanirad@uok.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_29

Fig. 1 Architecture of SRAM-based FPGA

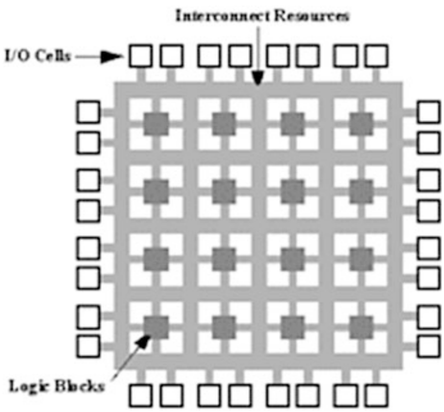
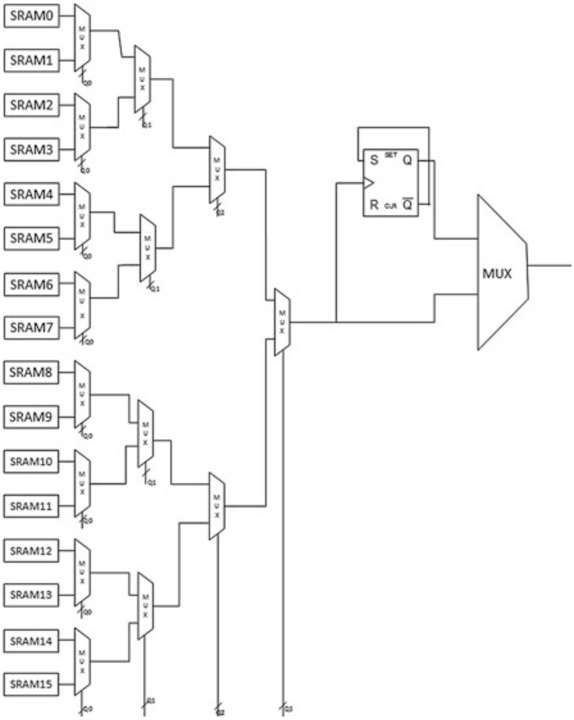


Fig. 2 The structure of a CLB



A k-input LUT can implement any k-input and 1-output switching function. The truth table of such functions has been stored in 2^k SRAM-cells in LUT and the output lines of these SRAM-cells connect to the LUT's output using a $2^{k \times 1}$ multiplexer [3]. LUT architecture has been illustrated in Fig. 3.

Various methods have been proposed to test CLB. These methods are different in the elapsed time of testing and the number of required CLB configurations.

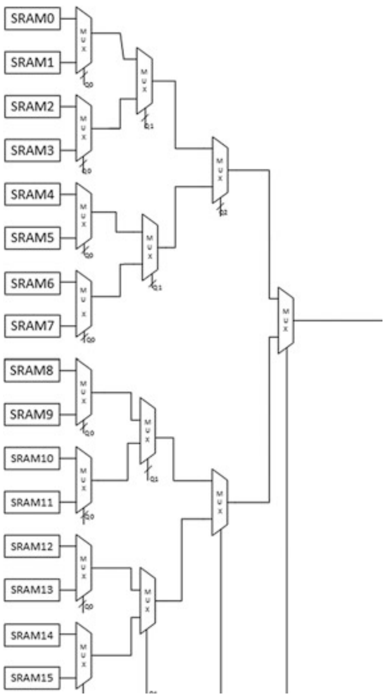


Fig. 3 The structure of a LUT

A large portion of these methods are based on external testing and some are concerned with Design for Testability (DFT). BIST as an efficient testing method is based on constructing a hardware core in CMOS-chip to make the testing operation internally [4]. The main parts of BIST architecture illustrated in Fig. 4. Test Pattern Generator (TPG) generates required test vectors (exhaustively, randomly,…) then these vectors apply to the Circuit Under Test (CUT) and finally using Output Response Analyzer (ORA) the CUT output values are compared with expected values and for each test vector Pass/Fail state of CUT will be determined.

Investigating BIST-based test method goes back to late 20th century [5–8]. But its implementation on FPGAs of different companies has been performed in recent years [9–11]. For instance, first BIST structure for LUTs in FPGAs has been proposed in 1996 which could be configured and was based on offline-test and

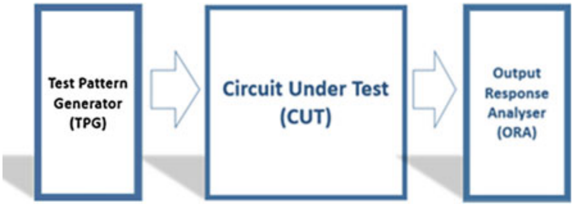


Fig. 4 General structure for BIST test

external memory was used for saving configuration and all operations during the test. Despite big hardware overhead, this structure provides maximum coverage using pseudo-comprehensive pattern [12].

In addition, offline-test was applied to VIRTEX-4 FPGA in 2004 and its main purpose was to increase error coverage in logic blocks [13].

2 Proposed BIST Architecture

2.1 Designed SRAM Cell

In the proposed test method, cells shown below have the Set and Reset ability. As can be seen in Fig. 5, NMOS transistor Resets cell and PMOS transistor Sets the cell.

2.2 General Structure of Our LUT

General structure of our LUT is as follows where all its sections are investigated In the next section (Fig. 6):

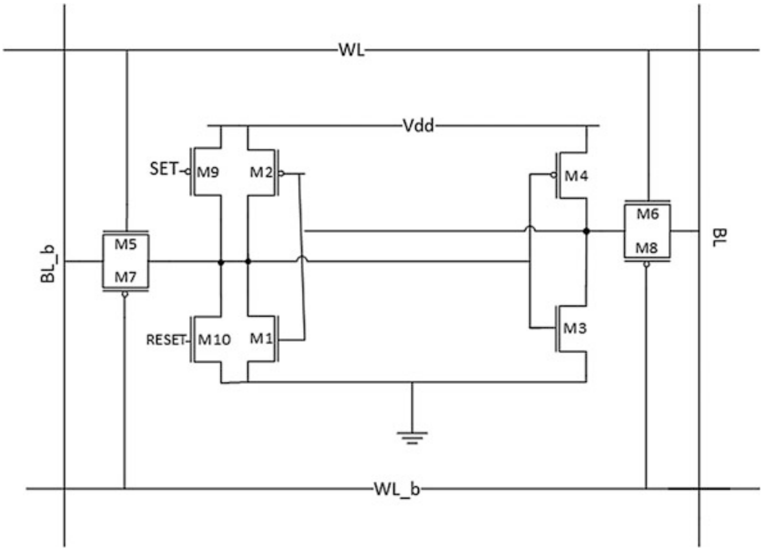


Fig. 5 Structure for designed SRAM cell

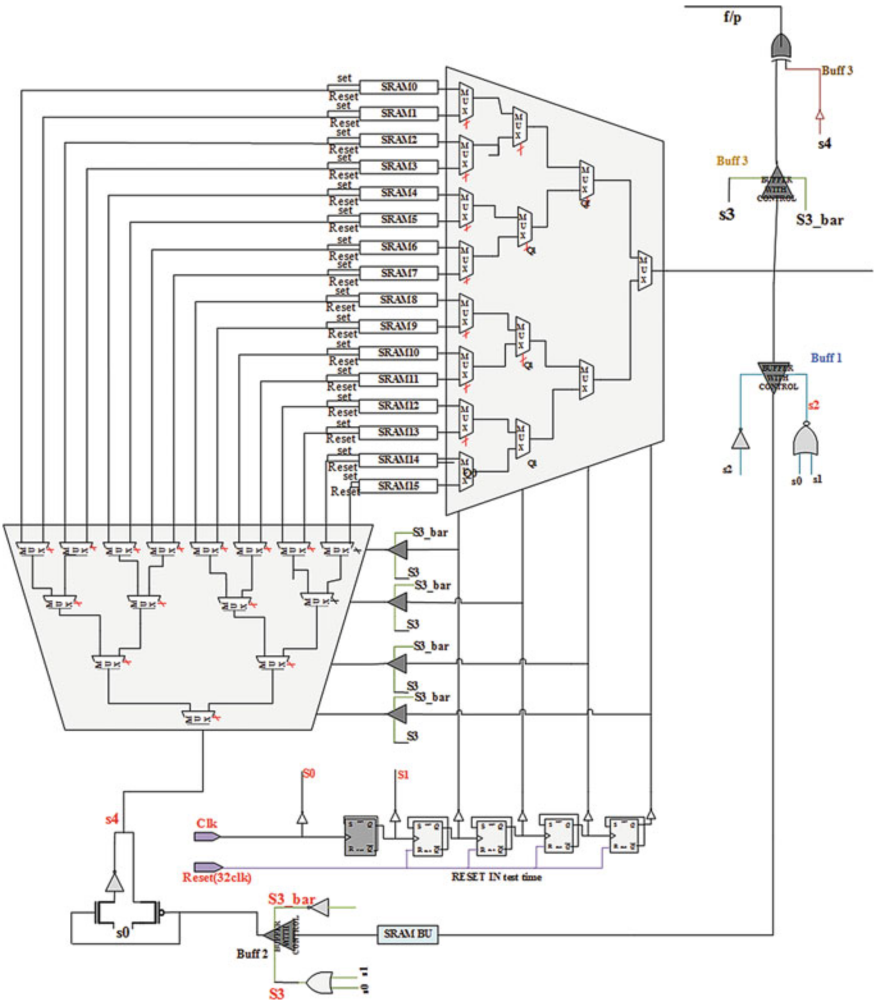


Fig. 6 Core schematic of the test implemented in LUT

2.3 Components of the Designed LUT

In the following, different sections of this structure have been investigated:

2.3.1 Main Structure of LUT

As mentioned in previous section, this structure is comprised of SRAM cells and MUX selector.

2.3.2 Structure of the Proposed Core Test

This structure is comprised of several sections which is repeated for all SRAM cell sequentially, The function of this structure is as follows:

First keep initial values of each SRAM cell, then apply test patterns to SRAM cell, analyze output response and writing backup value in the SRAM cell.

Sequential Counter

A 4-bit sequential counter has been used in this test method. This structure's purpose is to arrange required test vectors to test all LUT Sections and locating error considering its passage time.

1*16 DEMUX

A set of 15 DEMUX 1*2 have been used to return initial values stored in the backup cell to the SRAM cells.

Backup Cell

The proposed test method has a SRAM cell for storing initial values of the test as backup cell.

Rectifier Buffers

These buffers are used to prevent signal interference and signal passage scheduling and thus rectifying noiseless signals.

Comparator XOR Gate

This gate compares the values applied to the test circuit with received values. If values are the same, PASS is announced and if the results are different, Fault is announced.

Automating Elements of Control Pulses

This structure is comprised of a flip-flop and a NOR and OR gate [14] to automate control signals. Properties of these gates are used to construct the required signals.

3 Performance of the Proposed Architecture

General performance of this test which is repeated for all SRAM cell sequentially from 0 to 15, is as follows.

This test operates sequentially using 4 flip-flops which constitute a simple sequential counter. Each number of the counter should be divided to 4 equal parts. Therefore, input of the counter is a flip-flop (S1) with a clock_bar (S0) input. Thus, clock_bar (S0) and flip-flop (S1) outputs are used to construct control signals and TPG.

Test time which is equal to 32X clock time period, is applied to the counter by the USER. Counter starts counting with an input which its period is two times the clock (constructed by flip-flop).

In the proposed test method, instances of two control signals S2 (constructed through applying a NOR to S0CLK and S1) and S3 have been taken, where S2 is 1 at first ¼ of counting and signal S3 is 1 at final ¾ of counting.

At first ¼ of counting, value of the SRAM cell which is guided to output of LUT is introduced to the Backup cell using control buffer S2. Then S2 obstructs transmission path to backup cell. Then test pattern (TPG) is applied to the backup cell using the following circuit and considering initial stored value (Fig. 7).

Test pattern has been finally applied to Set and Reset SRAM cell and the SRAM cell operates as follows (Fig. 8):

As the counter counts continuously, response of the SRAM cell to Set and Reset is transferred to LUT output. Finally, this response is compared with inverse of the

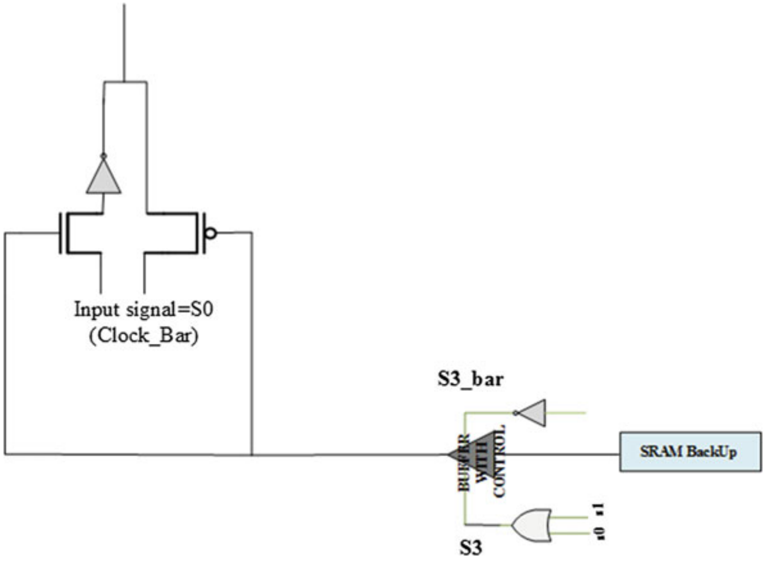


Fig. 7 DEMUX input signal generator circuit

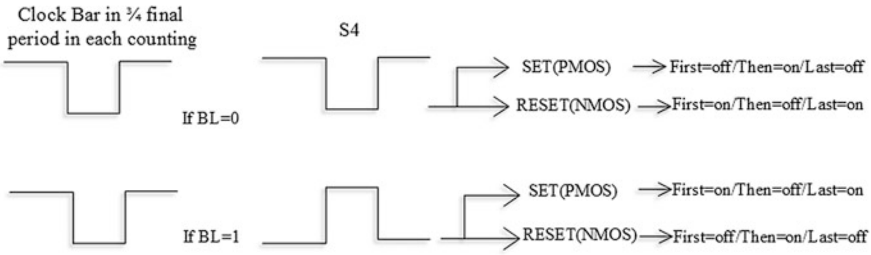

Clock Bar in ¾ final period in each counting
S4
If BL=0
SET(PMOS) →First=off/Then=on/Last=off
RESET(NMOS) →First=on/Then=off/Last=on
If BL=1
SET(PMOS) →First=on/Then=off/Last=on
RESET(NMOS)→First=off/Then=on/Last=off

Fig. 8 Set and reset response to input control signal

pattern introduced to the SRAM cells (these patterns are control basis input of SET and RESET transistor).

These steps are repeated for each 16 SRAM cells of the LUT, similarly. And initial value of cells is preserved after applying test pattern.

4 Simulation Results

By determining correct Reset value (32 clk) as in Fig. 6, counter counts from 0 to 15. For the counter to up count from 0 to 15, an inverter put at the output of each gate. For example, Test is arbitrarily started from 40 ns, and initial values of SRAM cells except 6th SRAM cell, have been configured as 1.

Two essential outputs S0 and S1 which are input of counter, and output of flip-flops of counter are as follows (Fig. 9):

In Fig. 10, the upper signal is the low bit of the counter to find the number of the SRAM cell, next control signal which is adopted from s0 and s1 essential signals

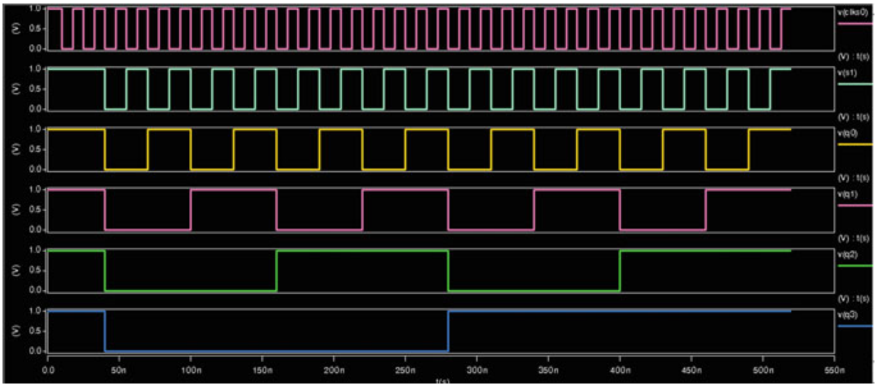


Fig. 9 Performance of the clock and counter and flip-flop S1 in the test structure

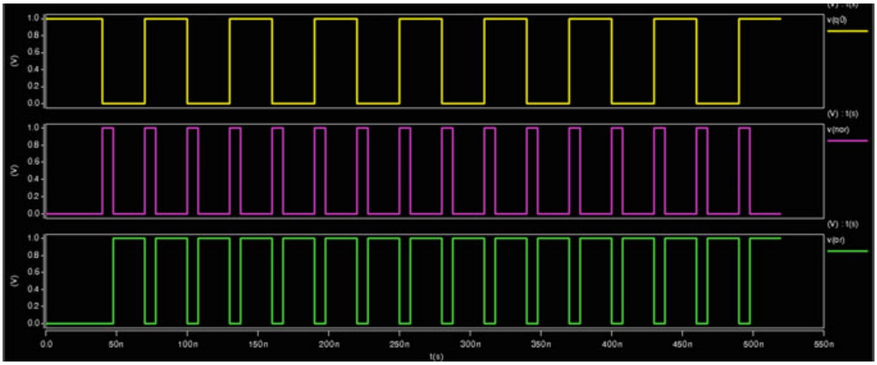


Fig. 10 Constructed control signals

has been shown, which are NOR and OR. These signals are used for automatic control test.

As can be seen, the control signals acting sequential.

Thus, the comparator gate with value of 0 announces PASS (Fig. 11):

In order to guarantee that value of SRAM cells and right performance of the test structure, If counter connect to inputs of MUXs in LUT again, We investigate the initial value of the SRAM cells. As can be seen in the Fig. 12, values of SRAM cells are repeated:

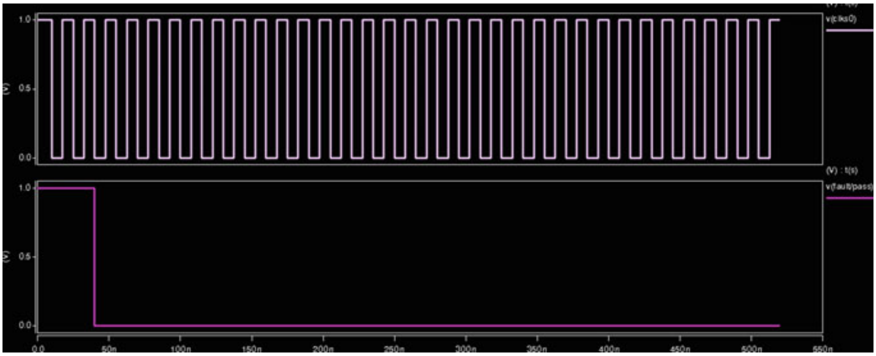


Fig. 11 Final result of comparison at test time

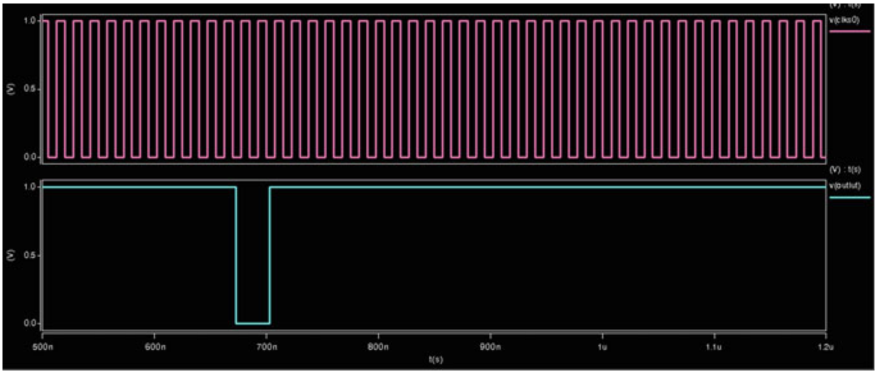


Fig. 12 Output of LUT after test

5 Conclusion

For a chip with 1 million or more transistors, a hardware is required for test, which integrated on the chip with low area overhead (about 3% and less in today's technologies) for BIST logic circuits. Such hardware test is employed for real design.

Using BIST technique in circuit test has more advantages compared to conventional ones. Automatic test equipment ATE for constructing normal test in conventional VLSI circuits includes hardware test using expensive hardware and long solution. This makes the test complicated and high cost is spent on tester at each second. Most complicated test equipment cannot be used for higher level tests. Therefore, BIST logic structure is designed for VLSI circuits which can be useful for other test purposes like repair and maintenance, detection or driving test [15].

In this paper, we designed a LUT with internal test core in transistor level [16] with H-Spice software. The proposed method investigates S@0 and S@1 errors and error of conversion from 0 to 1 which are most important probable errors automatically. This method preserves initial values of SRAM cells. Therefore, this method can be used for FPGA operation without disrupting values of cells. In this method, hardware overhead is reduced compared to a golden model (number of SRAM cells is reduced from 16 to 1) which is very important in FPGA industry.

References

1. Chmelar E (2003) FPGA Interconnect Delay Fault Testing. ITC
2. Kuon I, Russell T, Rose J (2008) FPGA architecture: survey and challenges. Found Trends Electron Des Autom 2.2:135–253

3. Kaviani A, Brown S (1996) Hybrid FPGA architecture. In: Proceedings: ACM/SIGDA international symposium on field-programmable gate arrays 1996, FPGA'96, Monterey, CA, pp 1–7
4. Bushnell M, Vishwani A (2004) Essentials of electronic testing for digital, memory and mixed-signal VLSI circuits. Springer Science and Business Media, vol 17
5. McCluskey EJ (1985) Built-in self-test techniques. IEEE Des Test Comput 2(2):21–28
6. Agrawal VD, Kime CR, Saluja KK (1993) A tutorial on built-in self-test. I. Principles. IEEE Des Test Comput 10(1):73–82
7. Jamal K (1996) Built-in self-test for integrated circuits having read/write memory. U.S. Patent No. 5,568,437
8. Zorian Y (1999) Built-in self-test. Microelectron Eng 49(1–2):135–138
9. Stroud CE, Leach KN, Slaughter TA (2003) BIST for Xilinx 4000 and Spartan series FPGAs: a case study. Null. IEEE
10. Liu, J, Simmons S (2003) BIST-diagnosis of interconnect fault locations in FPGA's. Electrical and computer engineering, 2003. IEEE CCECE 2003. canadian conference on vol 1. IEEE
11. Girard P et al (2006) An efficient BIST architecture for delay faults in the logic cells of symmetrical SRAM-based FPGAs. J Electron Test 22.2:161–172
12. Stroud C et al (1996) Built-in self-test of logic blocks in FPGAs (Finally, a free lunch: BIST without overhead!). VLSI test symposium, 1996. Proceedings of 14th. IEEE
13. Stroud, C et al (2004) Built-in self-test for system-on-chip: a case study. Test conference, 2004. Proceedings. ITC 2004. International. IEEE
14. Mano M (2002) Morris. Digital design. EBSCO Publishing, Inc.
15. Shen S, et al (2010) SPICE library for low-cost RFID applications based on pentacene organic FET. In: Wireless communications networking and mobile computing (WiCOM), 2010 6th international conference on. IEEE
16. Jan MR, Chandrakasan A, Borivoje N (2003) Digital integrated circuits: a design perspective

A Framework for Effective Exception Handling in Software Requirements Phase

Hamid Maleki, Ayob Jamshidi and Maryam Mohammadi

Abstract The exception handling structure allows software developers to reduce software maintenance cost through preventing faults, errors, and failures that may occur after exception arising. Forecasting possible exceptions and presenting powerful exception handling structures are noticeable in decreasing software modification workload and maintenance costs. But most of the developers neglect proper exception handling (EH) in early software development phases which make difficulty in software maintenance, indeed they underestimate EH. Since Focusing on EH only in the last phases of software life cycle is not a good policy, so we propose a framework, including principles, components, and metrics to present EH at the software requirement phase just while system scenarios are written. The proposed framework is a means for early exception discovery and leads to improve in software metrics: reliability, robustness, and maintainability. Applying the components of the framework: exception classification, scenario dependency graph and etc. and measuring proposed metrics in relation with exceptions allows to select proper EH strategies. At the end of the research, we present guidelines for the software tester to test all parts of the software according to the framework as a facility for verifying and correcting EH structures and discovering new possible exceptions.

H. Maleki (✉)
Department of Computer Engineering, Kermanshah Science
and Research Branch, Islamic Azad University, Kermanshah, Iran
e-mail: hamidmaleki32@yahoo.com

A. Jamshidi
Department of Computer Engineering, South Tehran Branch,
Islamic Azad University, Tehran, Iran
e-mail: ayob.jamshidi@gmail.com

M. Mohammadi
Department of Computer Engineering, Engineering Faculty,
Payame Noor University, Tehran, Iran
e-mail: Maryam_66.mohammadi@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_30

Keywords Exception · Exception handling (EH) · Scenario · Sub-function Failure

1 Introduction

The exception is an abnormal condition that arises at runtime and causes an error [1]. With forecasting similar conditions and presenting proper EH mechanisms, the possibility of failures relation to errors which are caused by exceptions will be reduced. EH is a powerful mechanism that separates the error handling code from the normal code [2]. The presence of EH can reduce software development effort, it supports: representation of errors as exceptions which deals with exceptions, definition of handlers, and employment of an adequate strategy for handling an exception upon its occurrence [3]. In object oriented programming languages such as Java and C++, these mechanisms are intended for using by programmers, but the current problem is underestimation the EH part by software developers and programmers, novice developers tend to ignore exception when developing system because of complex nature of EH [4]. Most of system failures are due to fault design EH algorithms [5]. These make some difficulties in software maintenance, furthermore Developers face trouble in cost estimation of software maintenance, so that in large-scale distribution of software project, 60% of workload is devoted to software maintenance, software maintenance refer to adapting the software with changes, finding and correcting fault, errors and failures in it [6].

Software metrics are used in the software life cycle, especially in the measurement and testing phase, to reduce the software complexity and to increase the comprehension of software systems, systems that follow these metrics will have less production cost and less maintenance workload. The purpose of predicting exception occurrences and handling them is to write robust and safe programs [7]. For reducing software maintenance costs, we consider reliability and robustness metrics. The high compliance of software systems with these two metrics prevents software from causing failure when the system services are delivered [8]. Software reliability depends on software quality at each phase of software development comprises of three activities: Error prevention, Fault detection and removing it, and Measurements to maximize reliability, specifically measures that support the first two activities [9]. Rosenberg states “an error is a programmer action or omission that results in a fault, a fault is a software defect that cause a failure, and a failure is the unacceptable departure of a program operation from program requirements” [9]. A failure in the provision of a service can have many reasons, in general, a fault occurrence in software causes one or more errors and releasing of these errors can cause system failures in the service delivery [10].

Another important metric in this research is robustness, Software robustness defined at [11] “the degree to which a system or component can function correctly in the presence of invalid inputs or stressful environmental conditions”. Reducing the occurrence of exception failures will improve software robustness [5]. If the software system is based on the components and their relationship with each other, the failure propagation among the components involves the system at risk. From the view of the research, arising unpredicted exceptions in program running may cause faults. Furthermore, exceptions may be well predicted, but the proper EH mechanisms are not presented, this defect depends on the magnitude of the negative consequences causing a system hazard, increasing maintenance costs and in simple terms, reducing the system reliability. Improper EH is not only related to developer and programmers, Because they are faced with challenges on the path of usage and provision of EH mechanisms, such as exceptions that are rarely occurring and difficult to predict or complex EH mechanisms which make challenges in program comprehension [12]. However, it is proven that using EH is risky because many errors are related to the EH code [13]. Many of these problems derived from ignoring the requirements of the exception in the early stages of software development.

The available techniques and tools in this field relate more to the testing and implementation phases than to requirement eliciting and designing phases. Exceptional use cases have introduced as models for making exceptional states and creating reliable systems [14]. Some methods are presented for adding these states to other UML diagrams, such as the sequence diagram [15, 16], as well as the state diagram [17]. But if the exceptional conditions are not properly predicted, they cannot be added to the UML diagrams, so the representation a proper model for the system is difficult. Due to the division of system requirements into functional and non- functional, in requirements eliciting phase, system specifications represented by considering the functional aspects of the system. System specifications must have the features of comprehensive, compatible, accurate, and clarify [18]. The SRS document describes system scenarios, the IEEE Std 830-1998 document explains how the SRS is compiled [19]. We propose a framework in requirements eliciting phase while system scenarios are written to detect and handle exceptions. The proposed framework (Sect. 2) follows rules and principles which describe in following sections. The goal of this framework is to reduce system maintenance costs. In fact, dealing with exceptions at first phases of software development is results in proper EH in next steps, in concluded software maintenance costs will significantly reduce. By considering possible exceptions in the proposed framework two metrics are introduced. These metrics are effective in software testing. By using the proposed method we offer the guidelines for software testing. In spite of the proposed framework for EH in software requirement phase, it is necessary to test EH part as well as others parts to prevent possible failures and to reduce software maintenance costs, so we offer the guidelines to apply in testing phase, by using them in all parts of implemented software can be tested for reducing the possibility of failures.

2 Proposed Framework

Most of the workload in software projects devoted to software maintenance, Finding and correcting faults, errors, and failures are done in software maintenance [6]. Most of the system failures are due to fault design EH algorithms [5]. software developers tend to ignore EH because of its complex nature [4]. When developing dependable software, the first step is to foresee exceptional situations and document how the system handling them [14]. For appropriate EH, We proposed a framework to apply in writing system scenarios at software requirement phase. In view of us, this framework will discover and correct more possible failures in relation to exceptions. The framework contains multiple components as illustrated in Fig. 1, these components explained in following.

Functional requirements have written at the upper part of the framework. These requirements are extracted in the form of scenarios and scenario events. Our goal is to identify exceptions in scenario events using the exception classification. This classification and other sections of the framework are explained in the following sections.

2.1 System Goals

System goal component is at the top of the framework with a three-levels granularity, we use this granularity to extract the system requirements [14]. At the first

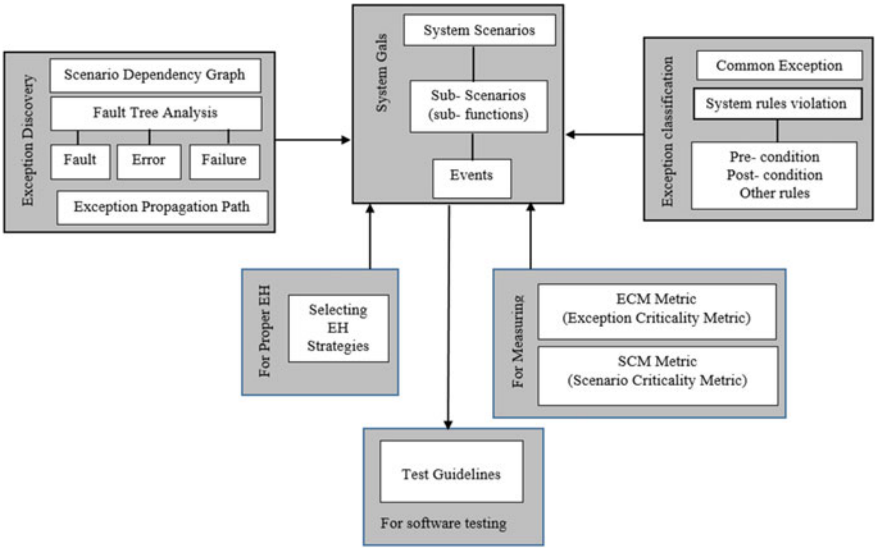


Fig. 1 Proposed framework for proper exception Handling

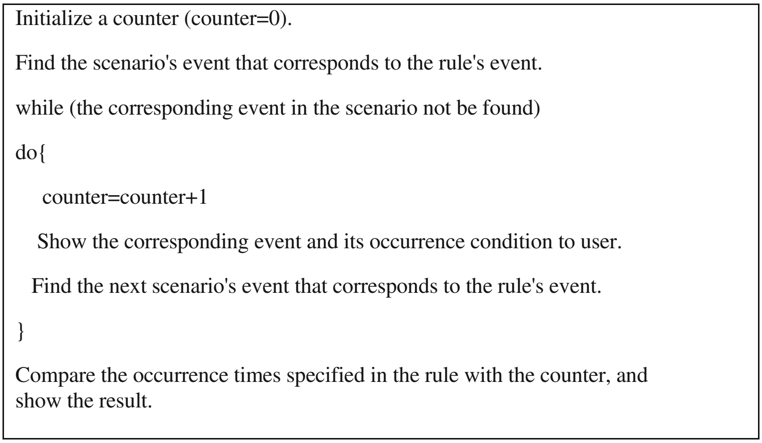
```
at first, extract the goals of system (main functions or targets), consider collectively them into single set goals.
        $GOALS=\{GOAL_1, GOAL_2, \ldots GOAL_N\}$ , $n>=1$
do following steps for each $GOAL_i \in GOALS$, 1<=i<=n
{
  extract all sub-scenarios for $GOAL_i$, put them inside set scenarios.
      $SCENARIOS[GOAL_i]=[\{ SCENARIO_1, SCENARIO_2, \ldots., SCENARIO_K\}$
  for each $SCENARIO_{J,\ 1<=J<K}$ as a element of SCENARIOS set and related to $GOAL_i$ do following instuctions
  {
     If for doing SCENARİO_J some sub-scenarios (sub-Functions) are needed, extract necessary sub- functions and put them in SUB_FUNCTİONS set.
     $SUB\_FUNCTIONS[SCENARIO_J]=[\{FUNCTION_1, FUNCTION2, \ldots FUNCTION_P\}]$
  }
}
```

Fig. 2 Requirements eliciting procedure

level of the granularity, system's goals are extracted. For doing a goal, several scenarios may be needed, so these scenarios are abstracted as the system Goals component at Fig. 1. On the other hand, for doing and writing more complex scenarios, several sub-scenarios may be required, Sub scenarios (sub-functions) are placed at the last level of the there- levels granularity. By using this granularity, we present a procedure to extract system requirements as scenarios, this procedure is shown in Fig. 2.

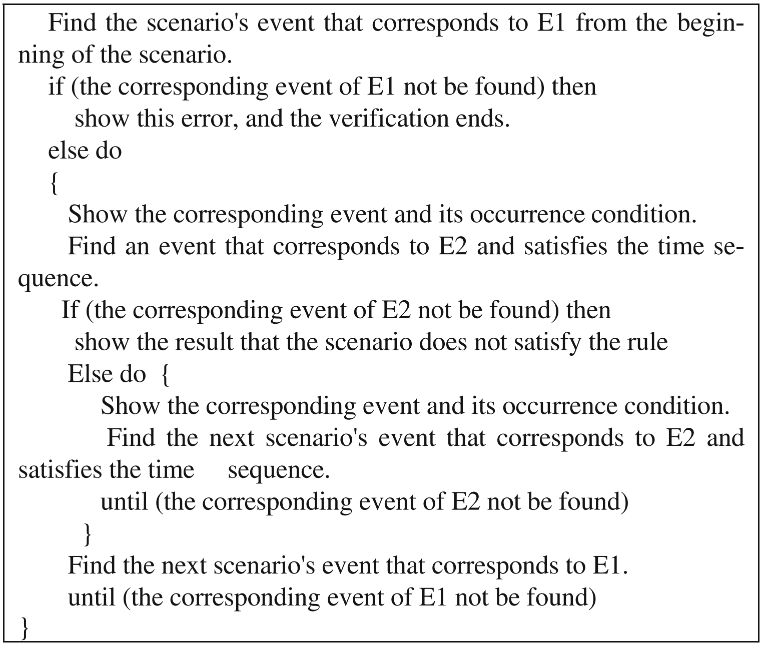
2.2 Exception Classification

From our point of view, the exceptions are divided into two categories: (1) common exception, (2) exceptions caused by violations of system rules. Programming issues such as memory problems, pointers, input and output, computational problems, and etc. may arise Common exceptions; most developers are familiar with such issues. Common exception types and exemplars classified at [5]. Developers faced these problems at the time of system implementation. Exceptions arose from violation system rules are related to scenario Errors; there is a verification procedure to correct such errors [20]. This procedure consists of two phases and is used to correct scenarios. Scenario errors include the vague presentation of events, the lack of essential events, additional events, and the wrong sequence of events. Figure 3 shows Verification procedure of the occurrence of scenario events. Figure 4 shows



```
Initialize a counter (counter=0).
Find the scenario's event that corresponds to the rule's event.
while (the corresponding event in the scenario not be found)
do{
    counter=counter+1
    Show the corresponding event and its occurrence condition to user.
    Find the next scenario's event that corresponds to the rule's event.
}
Compare the occurrence times specified in the rule with the counter, and
show the result.
```

Fig. 3 Verification procedure of the sequence of events [20]



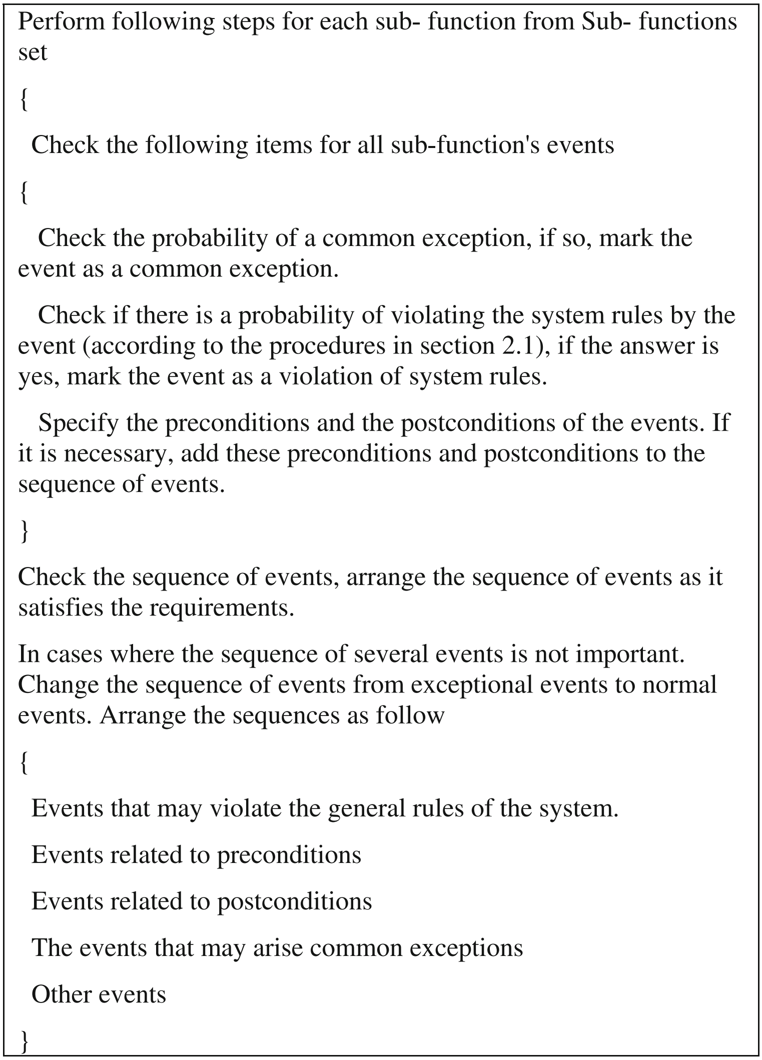
```
    Find the scenario's event that corresponds to E1 from the beginning of the scenario.
    if (the corresponding event of E1 not be found) then
        show this error, and the verification ends.
    else do
    {
       Show the corresponding event and its occurrence condition.
       Find an event that corresponds to E2 and satisfies the time sequence.
      If (the corresponding event of E2 not be found) then
        show the result that the scenario does not satisfy the rule
       Else do  {
            Show the corresponding event and its occurrence condition.
             Find the next scenario's event that corresponds to E2 and satisfies the time     sequence.
            until (the corresponding event of E2 not be found)
         }
       Find the next scenario's event that corresponds to E1.
       until (the corresponding event of E1 not be found)
}
```

Fig. 4 Verification procedure of the sequence of events [20]

the Verification procedure of the sequence of events E1 and E2. By using time sequence rules between events, this process verifies whether the sequence of events is correct? If the answer is no, it modifies events [20].

2.2.1 Improving Sub-functions Events

Using exception classification and verification procedures lead us to present a procedure for improving sub-functions events in EH (Fig. 5).



```
Perform following steps for each sub- function from Sub- functions set
{
  Check the following items for all sub-function's events
{
   Check the probability of a common exception, if so, mark the event as a common exception.
   Check if there is a probability of violating the system rules by the event (according to the procedures in section 2.1), if the answer is yes, mark the event as a violation of system rules.
   Specify the preconditions and the postconditions of the events. If it is necessary, add these preconditions and postconditions to the sequence of events.
}
Check the sequence of events, arrange the sequence of events as it satisfies the requirements.
In cases where the sequence of several events is not important. Change the sequence of events from exceptional events to normal events. Arrange the sequences as follow
{
  Events that may violate the general rules of the system.
  Events related to preconditions
  Events related to postconditions
  The events that may arise common exceptions
  Other events
}
```

Fig. 5 Improving procedure for sub-functions

2.3 Exception Discovery

2.3.1 Scenarios Dependency Graph (SDG)

After requirements eliciting based on three level granularity (2-1), all goals, scenarios and sub-functions are listed, each goal is related with one or several scenarios, and each scenario is related with one or several sub-functions. For simple representation, these relationships are shown as a graph called Scenarios Dependency Graph (SDG). A brief representation of this relationship which is based on three- levels granularity is shown in Fig. 6 as a diagram.

2.3.2 Fault Tree Analysis

Fault trees are a means for analyzing causes of hazards, not for identifying the hazards themselves [5]. We use this tree to identify and analyze the causes of faults and errors which cause a failure. The abbreviated fault tree shown in Fig. 6 depicts a chain of events leading to failure in a simple copy machine. The situations are represented in the fault tree by OR and AND gates, respectively [5].

For discovering faults, errors and failure in a software system, fault tree analysis has used in all levels of three-level granularity which is aimed to detect exceptions triggers and handling them. At the granularity, the combination of faults which thrown by sub-functions causes an error or a failure on scenarios level. So, the combination of errors thrown by scenarios which called by a Goal may cause an

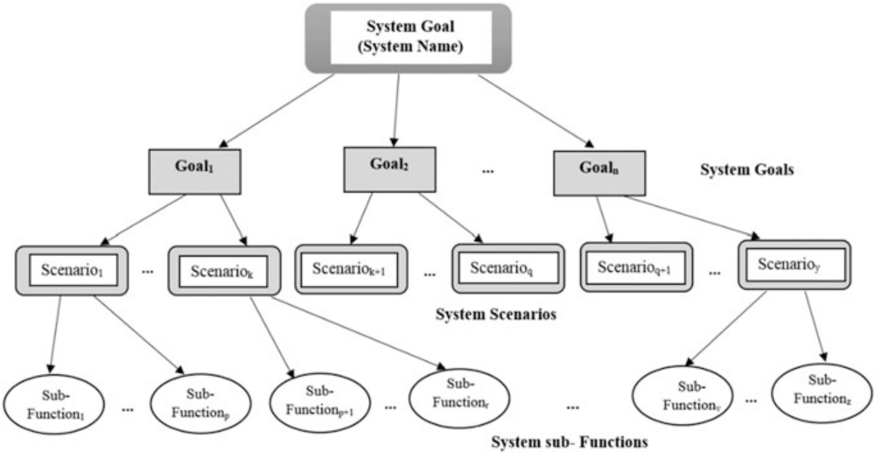


Fig. 6 A brief representation of scenarios dependency graph (SDG)

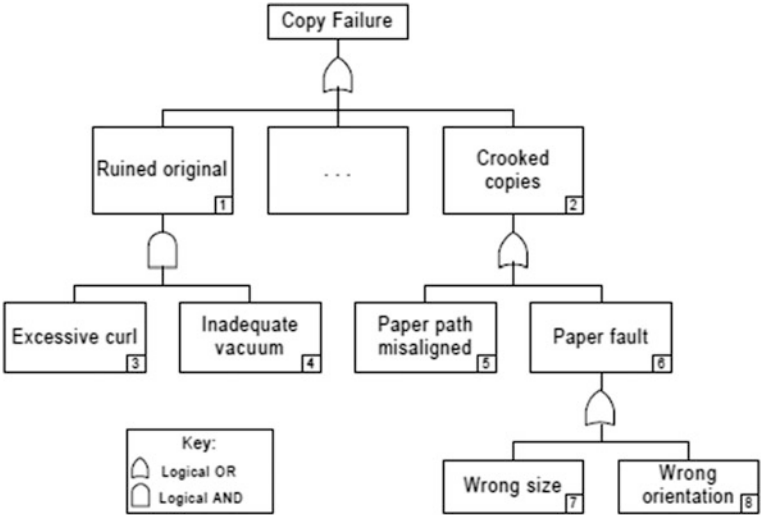


Fig. 7 Fault tree for copier [5]

error or failure at Goals level. By using fault tree analysis we can verify each step of all sub-function and then all scenarios, for each step we consider possible occurring condition e.g. entering invalid data by a user, occurring these state or combining some of them for each sub-function or scenario may cause an exception, we try to discover and handling possible exceptions (Fig. 7).

2.3.3 Exception Propagation Path

In a SDG as shown in Fig. 5, occurring an exception in a sub-function can cause a failure, if appropriate handling mechanism is not presented. Discovering possible exceptions in sub-functions of a scenario and combining them make an exceptional path from sub-function level to scenario level and combining them at scenario level to goal level make an exception path at Goals level. So we can detect exceptional path and list all possible exception. The path with frequently possibly critical exceptions is critical. This means that they may be causing more failures than other paths if they are not handled as well.

2.4 EH Strategies

There are approaches for EH at [4], We adapt them to use for handling discovered possible exceptions in system’s scenarios. We provide a scenario classification by considering the criticality of exceptions in a scenario which classifies scenarios in one of the A to F levels (Table 1).

Table 1 Strategies for EH

	Interaction agent			
Fault scope		Internal agent	External agent	Human agent
	Internal fault: a fault occurred within scenario: scenario is a sub-function which called by another scenario, it is may occur a fault within this scenario (called one)	An operation within the scenario may cause a fault	A fault may occur within the scenario while the scenario is interact with an External agent	The scenario is not able to execute the requests during interactions, so failure occurs
		Strategy 1: avoid exception	Strategy 2: thrown the exception to the caller (scenario or other sub-function)	Strategy 4: display appropriate message; explain failure
	External fault: a fault occurred outside the scenario: sub-scenario call other sub-scenario	Internal agent cannot cause external fault	The scenario is interacting with external agent and the fault may cause in the external agent	The scenario is interacting with human users and fault may occur because of faulty from the human user
			Strategy 3: catch and handling exception	Strategy 5: display appropriate message; ask for correct input

2.5 *Measuring Exceptions Criticality*

In this section, we propose two metrics for measuring exception criticality.

2.5.1 Exception Criticality Metric (ECM)

After discovering possible exceptions in all sub-function based on Sect. 2.3, all sub-functions can be classified by using two factors: (a) criticality of exceptions in a sub-function, (b) exceptions frequency means the number of exceptions in a sub-function in comparison with other sub-functions. Assessment of factor a depends on software projects, in some projects occurring an exception has higher cost so the sub-functions that are possible to make such exception are more critical. For measuring exception criticality, Table 2 represents an arbitrary classification for possible exceptions, an exception is classified in one of the arbitrary classes, correspond to a class a coefficient is allocated to an exception. If two or multiple sub-functions have an equal condition based on factor b then factor a can indicates more critical sub-functions as the following metric.

Table 2 The coefficients of the importance of the exceptions criticality

Coefficient	The metric of criticality importance of exceptions
0.1	Low
0.25	Middle
0.5	High
0.75	Very high

$$\begin{aligned}\text{ECM in a sub-Function} &= (\text{number of exception in class Low} * 0.1)\\ &+ (\text{number of exception in class Middle} * 0.25)\\ &+ (\text{number of exception in class High} * 0.5)\\ &+ (\text{number of exception in class Very High} * 0.75)\end{aligned} \tag{1}$$

For each sub-function the metric is calculated, obtained values are used in next section.

2.5.2 Scenario Criticality Metric (SCM)

Certainly, testing all the scenarios of a system, especially in large systems, is difficult. In addition, software testing is not always able to test all parts and possible occurrences. We suggest SCM to classification of scenarios in order to test more critical scenarios at first. For measuring SCM, the values of ECM (Sect. 2.5.1) for sub-functions of a scenario be added mathematically, following metric calculate the SCM for a scenario with number m of sub-functions.

$$\text{SCM for a Scenario} = \sum_{i=1}^{n} ECM\ of\ sub\text{-}function\ i \tag{2}$$

After calculating SCM metric for all scenarios, we can classify scenarios based on obtained values, it is obvious that scenarios with high values are more critical. It is important to mention that this classification is arbitrary and is used for testing the exception part, in the systems with more complexity some classes can be added to this classification with the new coefficient, it is dependent on the importance of testing the exception parts and testing workload. In table to average SCM is calculated as bellow, n is the number of system scenarios.

$$\text{Average SCM} = \frac{\sum_{i=1}^{n} SCM\ i}{n} \tag{3}$$

By using SCM metric and average SCM, all scenarios can classify in one of the scenario levels. The top 20% scenarios which have highest SCM values are more critical ones.

Table 3 Scenarios classification

(1) Scenario level	(2) Definition
(3) A (4) Catastrophic	(5) The scenarios with higher SCM (top 20%)
(6) B (7) Dangerous	(8) The scenarios with high SCM (between average SCM and higher SCM)
(9) C (10) Serve	(11) The scenarios with average SCM of all scenarios
(12) D (13) Low-risk	(14) The scenarios lowest SCM (Bottom 20%)

2.6 Discovering Software Critical Paths

After calculating ECM and SCM metrics, and scenarios classification which presented in Sect. 2.5, now discovering software critical path is easy, in the SDG (Sect. 2.3) of a system, all paths that have scenarios in level A or B are critical (as Table 3), indeed the Goals with greater number of critical scenarios (class A or B) consists of software critical path from system sub-function level to system Goal level at the SDG. These path are more important for EH and software testing. For simplicity we can represent critical path for Goals as a vector:

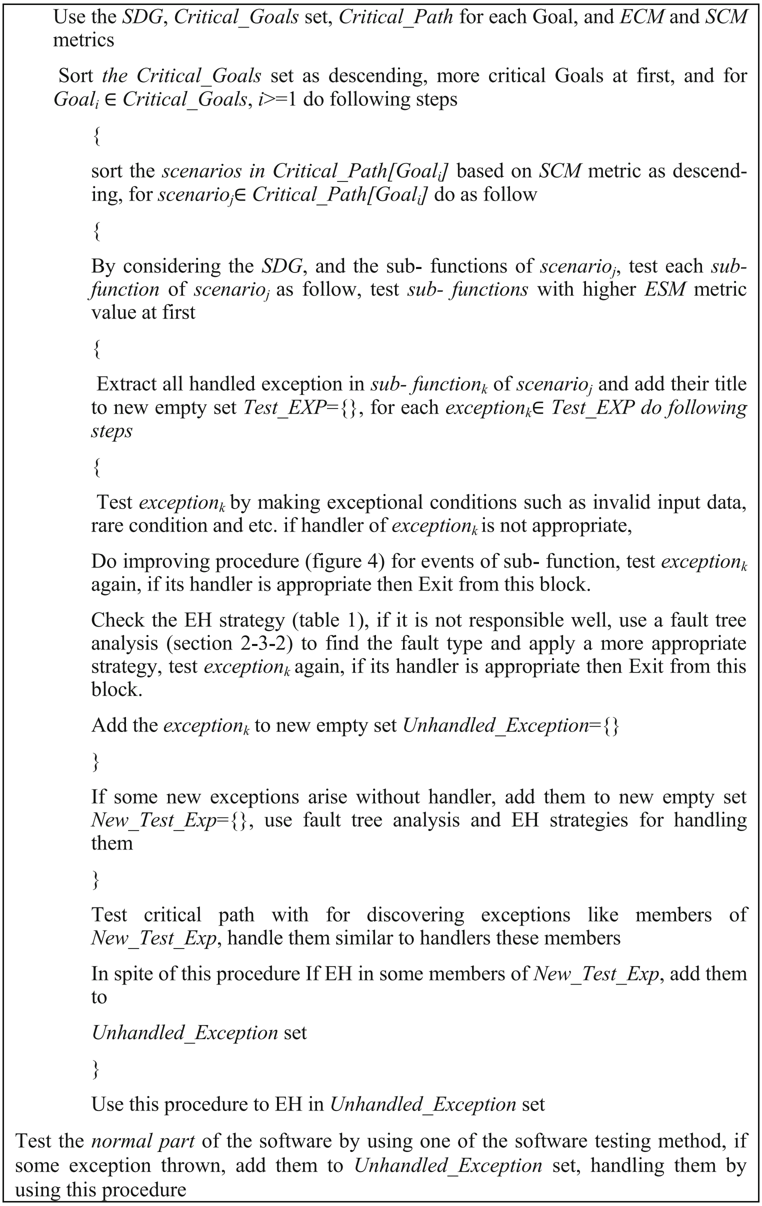
$$\text{Critical\_Path}[\text{Goali}] = \{\text{all scenarios of Goali with class A or B}\} \tag{4}$$

This vector can be extracted for each Goal in SDG of the system, the Goals with the greatest number of scenarios in their critical path are more critical. It may be thought that with such metrics and assumptions, all the Goals of the system can be critical, it is not true, we calculate exception criticality in a fine granularity model (SDG), at first critical exceptions are discovered, next critical sub-functions, those may arise more critical exception, and then critical scenarios for a goal are indicated, and at the last critical Goals are extracted and presents as Critical_Goals set, we can use all these in software testing.

$$\text{Critical\_Goals} = \{\text{all Goals with greatest member in their Critical\_Path set}\} \tag{5}$$

2.7 Guidelines for Software Testers

Software testers should test the normal part as well as the exceptional part. But due to the criticality of failures associated with the exceptional part, we recommend they test the exceptional part of programs at first. Although there are several phases between a proposed framework from software requirements and software testing, system implementation is made a lot of changes. So this question may arise: How could system scenarios be found in software implementation and tested? But to answer, one has to take into consideration is that because scenarios perform part of the system's goals, with the goals of the system, scenarios can be found and tested. In other words, implementing is a mapping of scenarios at the higher abstraction. The Guidelines for software testers with regards to EH is presented in Fig. 7 as a procedure (Fig. 8).



```
    Use the SDG, Critical_Goals set, Critical_Path for each Goal, and ECM and SCM metrics
    Sort the Critical_Goals set as descending, more critical Goals at first, and for $Goal_i \in Critical\_Goals$, $i>=1$ do following steps
        {
        sort the scenarios in $Critical\_Path[Goal_i]$ based on SCM metric as descending, for $scenario_j \in Critical\_Path[Goal_i]$ do as follow
        {
        By considering the SDG, and the sub- functions of $scenario_j$, test each sub-function of $scenario_j$ as follow, test sub- functions with higher ESM metric value at first
        {
        Extract all handled exception in $sub\text{-}function_k$ of $scenario_j$ and add their title to new empty set Test_EXP={}, for each $exception_k \in Test\_EXP$ do following steps
        {
        Test $exception_k$ by making exceptional conditions such as invalid input data, rare condition and etc. if handler of $exception_k$ is not appropriate,
        Do improving procedure (figure 4) for events of sub- function, test $exception_k$ again, if its handler is appropriate then Exit from this block.
        Check the EH strategy (table 1), if it is not responsible well, use a fault tree analysis (section 2-3-2) to find the fault type and apply a more appropriate strategy, test $exception_k$ again, if its handler is appropriate then Exit from this block.
        Add the $exception_k$ to new empty set Unhandled_Exception={}
        }
        If some new exceptions arise without handler, add them to new empty set New_Test_Exp={}, use fault tree analysis and EH strategies for handling them
        }
        Test critical path with for discovering exceptions like members of New_Test_Exp, handle them similar to handlers these members
        In spite of this procedure If EH in some members of New_Test_Exp, add them to
        Unhandled_Exception set
        }
        Use this procedure to EH in Unhandled_Exception set
Test the normal part of the software by using one of the software testing method, if some exception thrown, add them to Unhandled_Exception set, handling them by using this procedure
```

Fig. 8 A procedure for software testing based on EH

3 Conclusion and Future Work

In this study, we presented a framework for discovering and handling exceptions in software requirements phase while system scenario is written. The proposed framework tried to handle exceptions well, so by using this framework the failures of exception and maintenance workload will be reduced at software maintenance phase. In the framework, a three level granularity is used to represent software requirement in three levels including Goals, Scenarios, and sub-functions. We provided requirement eliciting procedure to represent requirements as an SDG, and improving the procedure for sub-functions of SDG. We proposed two EH criticality metrics, ECM for measuring exception criticality in a sub-function, and SCM for measuring exception criticality in a scenario. After measuring these metrics, scenarios are classified, the software critical path is represented by this classification in the framework. Critical paths are more important at software testing, they may arise more failure, so we proposed guidelines for software testing to reduce the number and the severity of possible failures. By these guidelines, faulty handler of exception can be corrected, and new exception can be handled by appropriated EH strategies, fault tree analysis is used for discovering the cause of failures in relation with exceptions. Calculating ECM and SCM metrics, discovering critical paths, and testing them leads to improve two important metrics: software reliability and software robustness made the costs of software maintenance will be reduced. As future work we intend to find the more effective factors to apply in proposed framework, these factors can be extracted through verifying more aspects in software requirement eliciting phase. Making an automatic tool is another future work, this tool simulates the framework automatically to improve and verify EH in scenarios and testing software implementation for reducing possible software failures.

References

1. Schildt H (2002) The complete reference (Java 2 Fifth Edition). McGraw-Hill Publication
2. Mao C-Y, Lu Y-S (2005) Improving the robustness and reliability of object-oriented programs through exception analysis and testing. In: 10th IEEE international conference on engineering of complex computer systems, pp 432–439
3. Garcia AF, Beder DM, Rubira CMF (2000) An exception handling software architecture for developing fault-tolerant software. In: Proceedings IEEE international symposium on high assurance systems engineering, vol 2000–Janua, pp 311–320
4. Shah HB, Gorg C, Harrold MJ (2010) Understanding exception handling: viewpoints of novices and experts. IEEE Trans Softw Eng 36(2):150–161
5. Maxion RA, Olszewski RT (1998) Improving software robustness with dependability cases. In: Fault-Tolerant Computing, 1998. Digest of papers. Twenty-Eighth annual international symposium on, pp 346–355
6. Ren Y, Xing T, Chen X, Chai X (2011) Research on software maintenance cost of influence factor analysis and estimation method. In: 2011 3rd International work. Intelligent systems application, pp 1–4

7. Dony C, Knudsen JL, Romanovsky A, Tripathi A (2006) Advanced topics in exception handling techniques, vol 4119. Springer
8. Fenton NE, Pfleeger SL (1997) Software metrics: a rigorous and practical approach. It Professional, vol 2, pp 38–42
9. Rosenberg L et al (1998) Software metrics and reliability. In: 9th international symposium on software reliability engineering, pp 1–8
10. Avizienis A, Laprie JC, Randell B, Landwehr C (2004) Basic concepts and taxonomy of dependable and secure computing. IEEE Trans Dependable Secur Comput 1(1):11–33
11. I.S. Committee (1990) Ieee std 610.12-1990 ieee standard glossary of software engineering terminology. [online] http//st-dards.ieee.org/reading/ieee/stdpublic/description/se/610.12-1990desc.html
12. Krischer R, Buhr PA (2012) Usability challenges in exception handling. In: 2012 5th international workshop on exception handling (WEH) 2012—Proceedings, pp 7–13
13. Sawadpong P, Allen EB, Williams BJ (2012) Exception handling defects: an empirical study. In: Proceedings IEEE international symposium on high assurance systems engineering, pp 90–97
14. Shui A, Mustafiz S, Kienzle J, Dony C (2005) Exceptional use cases. Model Driven riven Engineering Languages and Systems, LNCS, vol 3713, pp 568–583
15. Ciraci S, Sozer H, Aksit M, Havinga W (2011) Execution constraint verification of exception handling on UML sequence diagrams. In: Secure software integration and reliability improvement (SSIRI), 2011 Fifth international conference on, pp 31–40
16. Halvorsen O, Haugen O (2006) Proposed notation for exception handling in UML 2 sequence diagrams. In: Software engineering conference, Australian, 2006, pp 10–pp
17. Pintér G, Majzik I (2004) Modeling and analysis of exception handling by using UML statecharts. In: International workshop on scientific engineering of distributed java applications, pp 58–67
18. Bruegge B, Dutoit AH (2000) Object-oriented software engineering. Prentice-Hall, Up. Saddle River, NY
19. I. C. S. S. E. S. Committee and I.-S. S. Board (1998) Ieee recommended practice for software requirements specifications
20. Toyama T, Ohnishi A (2005) Rule-vased verification of scenarios with pre-conditions and post-conditions. In: Requirements engineering, 2005. Proceedings. 13th IEEE international conference on, pp 319–328

HMFA: A Hybrid Mutation-Base Firefly Algorithm for Travelling Salesman Problem

Mohammad Saraei and Parvaneh Mansouri

Abstract The Travelling Salesman Problem (TSP) is one of the major problems in graph theory and also is NP-Hard Problem. In this work, by improving the firefly algorithm (MFA), we introduced a new method for solving TSP. The result of the proposed method has compared with the other algorithms such as Firefly algorithm, GA and PSO. The Proposed Method out performs of other algorithms.

Keywords NP-Hard problem · Travelling salesman problem (TSP)
Firefly algorithm (FA) · Swapping mutation · Inversion mutation
Insertion mutation · 2-Opt mutation

1 Introduction

Traveling salesman problem (TSP) is one of the known problems in artificial intelligence. TSP is a discrete optimization problem and also NP-Hard problem. Many Studies have been done to find the best solution for this problem, but no exact solution has been provided yet. This simple problem has many applications, including vehicle timing [1] route optimization to transport the goods to different locations, route optimization for postal shipments, Vehicle routing [2] and route minimizing of a tour. The TSP represents the salesman who wants to visit a set of cities exactly once and finally turns back to the starting city. The objective is to determine the tour with minimum total distance (see Fig. 1).

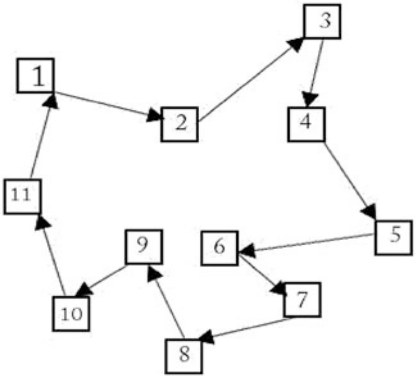
M. Saraei (✉) · P. Mansouri
Department of Computer, Faculty of Technical, Islamic Azad University,
Arak Branch, Arak, Iran
e-mail: mohammadsaraei@gmail.com

M. Saraei
Young Researchers and Elite Club, Islamic Azad University, Arak Branch,
Arak, Iran

P. Mansouri
Department of Computer Science, Delhi University, Delhi, India

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_31

Fig. 1 Sample of solving TSP



In recent years many approaches have been developed to solve TSP The simplest exact method solve all possible tours, and then select the optimal tour with the minimum total cost. All possible permutations of N cities are equal to $N!$, so, every tour can be represented in 2N different manner depends on the initial city and the length of tour. So the size of search space is computed as Eq. (1). It is obviously that this measurement is not possible for computational time even for 50 cities.

$$T = \frac{N!}{2 * N} = \frac{(N-1)!}{2} \tag{1}$$

In Sect. 2, we examine some algorithms provided for solving the TSP. In addition, a mathematical model and an introduction to Firefly algorithm will be described. In Sect. 3, we consider the proposed algorithm. We indicate the desired mutations used in the proposed algorithm. Then, in Sect. 4, the proposed algorithm will be used to solve the TSP and the results will be compared with other bio-inspired algorithms such as GA, PSO and FA. In Sect. 5, we review the strengths of the proposed method compared to other algorithms. Finally the paper ends with Conclusion and Acknowledgement.

2 Theoretical Principles

Recently, different methods have been proposed to solve the TSP whichever have their own strengths and weaknesses but it is important to use the method or algorithm that achieves the best tour in the shortest possible time. Some meta-heuristic algorithms used to solve the TSP are: Genetic algorithm (GA) [3–5]. Particle swarm (PSO) [6–8], Ant colony (ACO) [9–11], Memetic algorithms [12], Artificial Bee Colony [13, 14], Bee Colony [15, 16], and etc. Better solution can be achieved by changing the parameters in any of these algorithms. In 2014, Saranya et al. have presented a method for solving the TSP based on Tabu search and biological algorithms such as ant colony optimization algorithm, cuckoo algorithm and bee algorithm [17]. In 2012, Yang et al. have purposed an optimization approach to reduce the processing costs associated with ant colony routing (ACO) and they have Improved ACO using individual diversification strategy [18]. So that, the speed of ACO greatly increased. They used this approach for solving the TSP. Rizak Allah et al. In 2013 have presented a hybrid algorithm called ACO—FA, that integrate Ant colony algorithm (ACO) and Firefly algorithm (FA) to solve unlimited problems [19]. Las zolocota [20] used Firefly algorithm to solve multiple TSP. In this work, we purpose an accurate and fast algorithm to solved TSP by adding best and Effective mutations to FA.

2.1 Mathematical Model of Travelling Salesman Problem

In this study, our purpose is to find best tour of symmetric TSP. In symmetric TSP, the distance from city A to city B equals to distance from city B to city A. However, in asymmetric mode, the distance from city A to city B is not necessarily equal to the distance of city A to city B. We can consider the symmetric TSP problem as a complete and undirected graph where

$$A = \{(i,j) : i,j \in V, i \neq j\}$$

'A' is a set of edges and $V = \{0, \ldots, N\}$ is a set of nodes.

The number of possible solutions are $\frac{1}{2}(\mathrm{N} - 1)$, (N is the number of cities and N > 2). In fact, the number of possible solutions are equal to the number of Hamiltonian cycles in a complete graph with N nodes. The mathematical form of the objective cost function of TSP is as follows:

$$\min \mathrm{z} = \sum_{\mathrm{i=0}}^{\mathrm{N}} \sum_{\mathrm{j \neq i, j=0}}^{\mathrm{N}} \mathrm{c_{ij}} \tag{2}$$

where c_{ij} is the distance between nodes i and j. i, j = 0,1, …, N

2.2 Distance of Two Cities

For computing the distance between two cities (nodes), there are some methods such as hamming and Euclidean distance formulas. We consider the cities as nodes of two-dimensional Cartesian space (x, y) and by using Euclidean distance formula as follows:

$$d(i,j) = d(j,i) \tag{3}$$

$$d_{ij} = \sqrt{\left(x_i - x_j\right)^2 - \left(y_i - y_j\right)^2} \tag{4}$$

3 Firefly Algorithm

In 2009 Yang [21] introduced the optimization firefly algorithm (FA). FA has inspired by fireflies that use short and rhythmic lights to attract the hunt, protection or attract mates systems. There are two important issues in firefly algorithm: changes in light intensity and formulating the attraction. For simplicity, we can always assume that its light determines the attraction of firefly, which in turn is associated with the objective function. The attraction is proportional to brightness and a firefly with lower light is absorbed to firefly with brighter light, and if there is no light, it moves randomly. The firefly will be visible only for a limited period due to distance and light reduction by air. A firefly can be considered as a point light source.

We know that the light intensity at a certain distance r from the light source follows the inverse square law. The law states that the light intensity I decrease by increasing the distance r.

$$I \propto \frac{1}{r} \tag{5}$$

As mentioned, by increasing distance of two fireflies, the light intensity of between them is going to be weaker and weaker. In the simplest case, we can consider the light intensity of a point source by analysis factor γ, in distance r as Eq. (5) (I_0 is the light intensity in r = 0).

Since the attraction of firefly is proportional to light intensity seen by adjacent firefly, the attraction of fireflies is defined as Eq. (6) (β_0 is the attraction in r = 0).

$$\beta(r) = \beta_0 e^{-\gamma r^2} \tag{6}$$

The distance between any two fireflies i and j at x_i and x_j, respectively, is the Cartesian distance:

$$r_{ij} = \sqrt{\sum_{k=1}^{d} \left(x_{i,k} - x_{j,k}\right)^2} \tag{7}$$

where $x_{i,k}$ is the k th component of the spatial coordinate x_i of i th firefly.

Brightness is also proportional to objective function. Therefore, updating the location for each pair of fireflies i and j at x_i and x_j is as following equation:

$$x_i^{t+1} = x_i^t + \beta_0 e^{-\gamma r_{ij}^2}\left(x_j^t - x_i^t\right) + \alpha_t \varepsilon_t \tag{8}$$

The firefly algorithm has been formulated by following properties:

1. All fireflies are single type, so that a firefly attracts all other fireflies.
2. Attraction is proportional to brightness and a firefly with lower light is absorbed to firefly with brighter light.
3. If there is no firefly brighter than the other firefly, then the firefly moves randomly.

The pseudo code of FA as follows:

Algorithm 1: The pseudo code of Firefly algorithm

```
Firefly Algorithm:
Objective function f(x),$x = (x_1,\ldots,x_d)^T$
Generate initial population of firefly $x_i$ $(i = 1,2,\ldots,n)$
Light intensity $I_i$ at $x_i$ is determined by $f(x_i)$
Define light absorption coefficienty
While ($t <$ MaxGeneration)
    for i = 1: n all n fireflies
        for j = 1: n all n fireflies
            if ($I_j > I_i$),Move firefly I towards j in d _dimension;end if
            Evaluate new solutions and update light intensity
        end for j
    end for i
    Rank the fireflies and find the current best
End while
Postprocess result and visualization Rank the fireflies and find the current best;
End while;
Post process results and visualization;
End procedure
```

3.1 Some Mutation Operators

In FA, for finding the shortest path, we can use the various mutations such as: random, inversion, swapping and greedy mutations and, etc.

This will prevent the falling into the trap local optimal. With this work new solution will be replacement of previous solutions.

3.1.1 Swapping Mutation

This method is the most commonly used methods; this mutation can be performed on a couple of points. The operation of this mutation for two points, swap the two points randomly shown in the Fig. 2.

3.1.2 Inversion Mutation

In this method two elements are randomly selected and then we inverse enclosed elements in block between them and place them on their own place (see Fig. 3).

3.1.3 Insertion Mutation

In this method two elements are randomly selected and transfer one of the two elements after other chosen element (see Fig. 4).

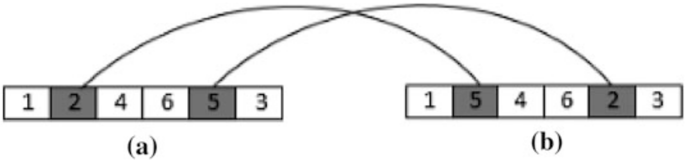


Fig. 2 Swapping mutation. **a** Before mutation, **b** after mutation

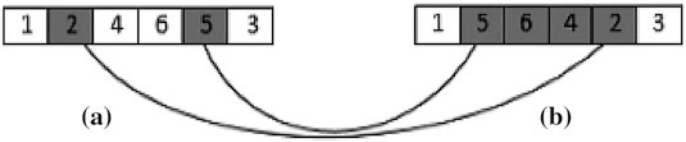
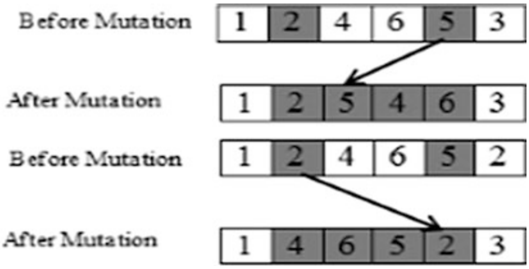


Fig. 3 Inversion Mutation. **a** Before mutation, **b** after mutation

Fig. 4 Insertion mutation



3.1.4 2-Opt Mutation

This mutation is one of the most effective mutations for improvement in finding optimal tour in TSP.In this mutation via check path and selection four points with k distance between them, if the way of their connection contains twist, will open twist path and decreases the path length. Figure 4a optimized route is displayed in Fig. 4b.

The process of execution of this mutation is as follows:
We select randomly four points I, j, n, m

1. We calculate the distance between the points using following relation.
2. |(i, j)| + |(n, m)| > |(i, m)| + |(n, j)|.
3. If the relation is true we change the value of element j with the value of element m.
4. If necessary we repeat this process and select other points.

Changes resulting from this mutation on the tour are as follows (Figs. 5 and 6):

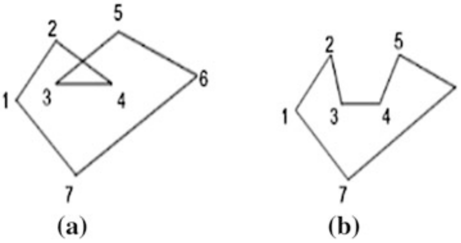


Fig. 5 **a** First tour (before mutation), **b** second tour (after mutation)

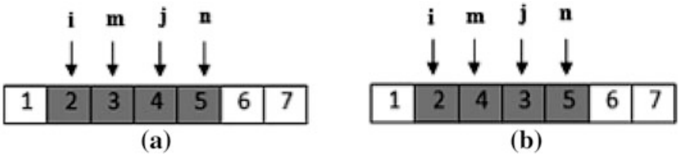


Fig. 6 **a** solution before mutation, **b** solution after mutation

4 Proposed Method

Sometimes, in practice it is possible the FA to be trapped in a local minimum and the rapid convergence there. This is not to achieve an appropriate response. Therefore, to solve the TSP our algorithm should have the least amount of complexity, because complexity increases the running time of algorithm. In this study, for escaping of the local minimum at acceptable running time of algorithm, our purpose is to add some mutations in FA to find a faster algorithm without trapped in a local minimum. An operator develops a mutation that causes widening areas are discovered. In addition to the TSP graph, edges should not be crossed. Because the cross is to increase the length of the tour. Using the appropriate responses to increase the mutation rate. To fix the problem of trapped in a local minimum, instead of using an operator or a combination of some operators; we add randomly one of operators in each iteration of FA. Each mutation operation of Sect. 3.1.4 can able to solve just some special difficulties to determine shortest paths. In Table 1, we can see the results of solving salesman problem with four cities, population size are 10 and the number of iterations is 700. The structure of the proposed algorithm is as follows (Fig. 7):

Step1:
Create Model Of Benchmark TSP Problem.
Step2:
Objective Function F (Tour), Tour is a Matrix Contains Number of Cities.
Step3:
Define Parameters Algorithm Such as Max Iteration, number of population,delta, gamma,alpha, …
Step4:
Generate Initial Population Randomly {each Member of the Population is a Tour}.
Step5:
Calculate $r_{i,j}$,$r_{j,I}$ is equal: norm {tour(i)-tour(j)}.
Step6:
Evaluated New Tour and Calculate Cost.
Step7:
Create Random Number Between [1 to 4]
Step8:
Switch on Random Number Obtained in Above stage.
Step9:
Evaluate New Tour Obtained Of Mutation and Update Cost.
Step10:
If (Cost of (newer tour) < cost of (new tour)).
Then The newer tour is replaced tour **Else** The new tour is replaced tour.

In MFA (Combine four Mutation) in each iteration, randomly one of the mutation operators (swapping, inversion, Insertion, 2-opt) added to FA and TSP

Table 1 Comparison between mutations for TSP with 700 iteration and 10 populations

TSP problem	Swapping		Insertion		Reversion		2-Opt		Combine	
Name	Best	Average	Best	Average	Best	Average	Best	Average	Best	Average
Ulysses16	7589.0436	7909.055	7521.3761	8295.5414	7399.7618	7413.972	7461.4804	7504.5579	7398.7618	7404.0949
Ulysses22	8402.0016	8776.9591	8418.4765	11,372.2573	7550.8818	7609.0946	7799.3926	8402.8132	7530.9701	7573.5835
Gr24	1549.3962	1638.6441	1397.3769	1691.6728	1288.6966	1298.6519	1421.0923	1505.8183	1279.5031	1317.1837
Eil51	1127.8884	1127.8884	1103.0047	1103.0047	556.4436	562.3819	545.9919	575.1371	430.8606	445.1045

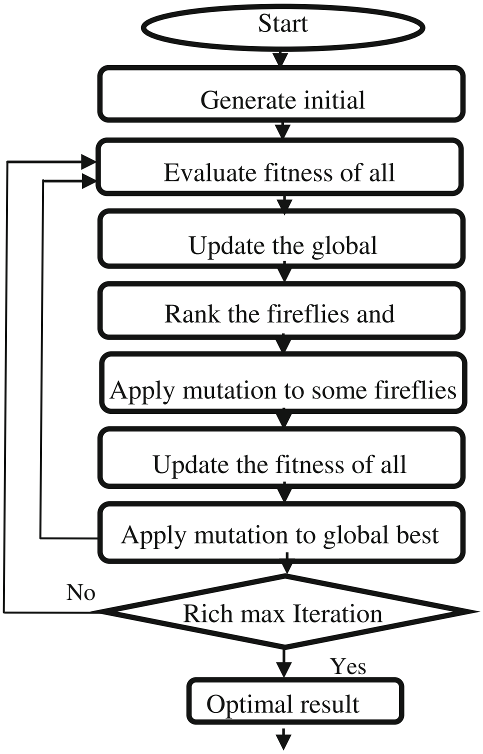


Fig. 7 The flowchart of the proposed algorithm

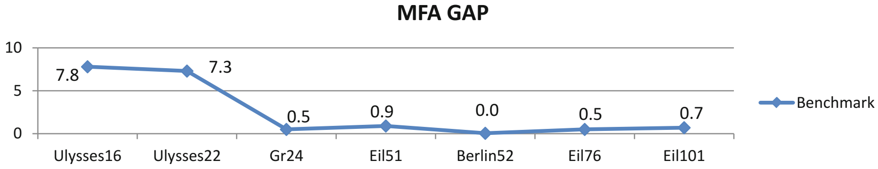


Fig. 8 The GAP Diagram of the MFA algorithm to benchmarks

cost function tour is computed. Then, combined mutation (by using the method mentioned earlier) apply to global best solution. After some iteration minimum value of cost function will be achieve. Comparison of proposed algorithm and other algorithms such as FA, GA and PSO show that the proposed algorithm outperforms of others. The results show that the random combination of four mutations of Sect. 3.1.4, gives us the better solution.

5 Simulation

In this paper, using the improved FA to solve the standard TSP by MATLAB 2013 on a platform with Intel CORE i5, 2 GB RAM, and Windows 7 operating system's has been solved for standard algorithms such as GA, PSO and also FA. The average of 10 times running for each standard library problems TSPLIB [22] was calculated and the results have been compared with the results of the proposed algorithm in this paper. Also in Table 2 the benchmarks are used is visible. Respectively. Setting parameter GA and PSO and FA can be seen in Tables 3, 4 and 5. Table 6 shows the

Table 2 Benchmark of TSP

Benchmark problem	N.City	Optimal
Ulysses16	16	6859
Ulysses22	22	7013
Gr24	24	1272
Eil51	51	426
Berlin52	52	7542
Eil76	76	538
Eil101	101	629

Table 3 Defined parameters for firefly

Parameters	Value
Maximum number of iterations	700
Number of fireflies	10
Light absorption coefficient	2
Attraction coefficient base value	1
Mutation coefficient	0.2

Table 4 Defined Parameters for PSO

Parameters	Value
Maximum number of iterations	700
Population size (Swarm size)	10
Inertia weight	1
Inertia weight damping ratio	0.99
Personal learning coefficient	0.2
Global learning coefficient	0.4

Table 5 Defined Parameters for GA

Parameters	Value
Maximum number of iterations	700
Population size	10
Crossover percentage	0.5
Mutation percentage	0.5

Table 6 Comparison between FA, GA, PSO, and MFA for TSP with 1000 iteration and 10 populations

TSP problem	GA		Besttime (s)	PSO		Best time (s)	FA		Best time (s)	MFA		Best time (s)
Name	Best	Average		Best	Average		Best	Average		Best	Average	
Ulysses16	7539.5764	8270.5777	20.837	8342.7014	9683.9225	15.8817	9218.9743	10,030.5512	25.0046	7398.7618	7404.0949	8.0857
Ulysses22	9019.0812	9610.0365	23.9897	9227.943	11,503.3717	15.6765	10,772.1213	12,182.9003	26.1966	7530.9701	7573.5835	11.4056
Gr24	1657.9434	1703.6828	20.65	2073.9016	2401.5683	17.0012	2224.2237	2485.3758	23.7375	1279.5031	1317.1837	12.0406
Eil51	702.1431	757.4782	22.6451	1113.7946	1223.0594	17.5213	1094.3539	1199.7793	26.7765	430.8606	445.1045	34.8505
Berlin52	12,831.6181	14,141.9204	23.172	19,279.1045	22,368.3565	17.9856	19,578.6951	21,788.4744	29.693	7546.8573	7714.4453	51.4121
Eil76	1105.1721	1199.2329	24.7965	1800.4436	1991.7083	18.4846	1564.0598	1791.851	28.079	545.7331	568.7338	73.1231
Eil101	1650.7169	1754.5091	26.1545	2639.2026	2909.4256	17.8548	2385.9235	2515.0022	27.9111	634.8189	668.6202	88.0942

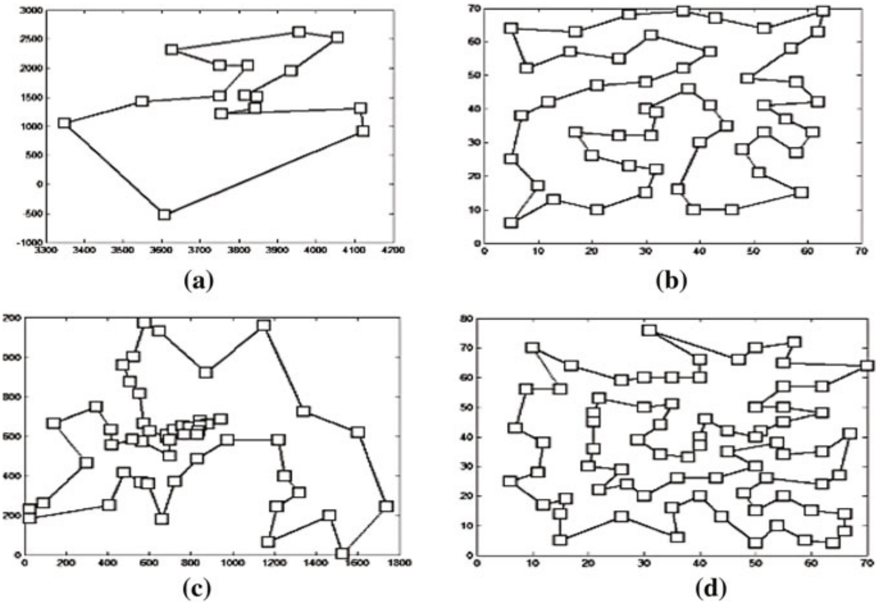
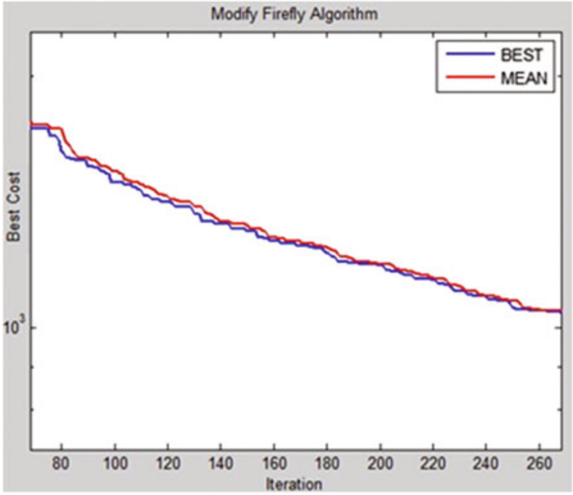


Fig. 9 **a** The simulation result for gr24. **b** The simulation result for Eil51. **c** The simulation result for Berlin52. **d** The simulation result for Eil76

Fig. 10 The output graph of MFA algorithm for Eil76



results for size population 10 and the number of iteration: 1000 in applied problems. The simulation results have been presented as graphical output in Fig. 9. And cost and roc curve Diagram for Eil76 (TSP Problem) displayed in Figs. 10, 11 in addition Fig. 8 shows the gap values of the MFA algorithm for Benchmarks, where

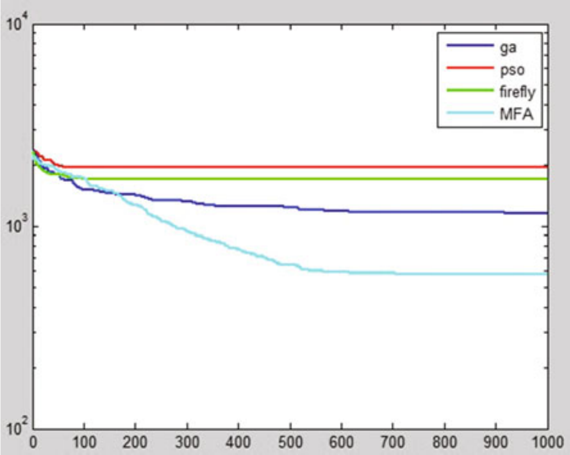


Fig. 11 The rocurve graph of algorithms for Eil76

the gap is defined as the percentage of deviation. In this formula (A′) best answer of found by our algorithm and (A) the best answer Known (Optimal) for Benchmarks. The gap is calculated as follows:

$$GAP = \frac{C(A') - C(A)}{C(A)} \times 100 \tag{9}$$

6 Discussion and Conclusion

In this paper, a novel meta-heuristic algorithm called improved FA was applied to solve the standard TSP and we compared the performance of the three famous SI (swarm intelligence) algorithms include of GA, PSO, FA to solve TSP. The FA algorithm has a strong global search capability in the problem space and can efficiently find optimal tour also it is quite simple and easy to apply, and it is efficient for large size matters. In this study a novel FA algorithm based on a hybrid mutation scheme (named MFA algorithm) was introduced for TSP. Experimental results show that this approach considers both running time and solution quality as well. In According to the results, it can be seen that the proposed algorithm has much better result compared to standard algorithms. The results of GA, PSO and FA are converged rapidly and there is no significant change by increasing the repetitions. The proposed algorithm is significantly improved by increasing the number of repetitions. As a future work, the algorithm FA Can be hybridized with SI algorithms to find better results.

Acknowledgements We thank all who have supported us for this research.

References

1. Park YB (2001) A hybrid genetic algorithm for the vehicle scheduling problem with due times and time deadlines. Int J Productions Econom 73(2):175–188
2. Hu W, Liang H et al (2013) A hybrid chaos-particle swarm optimization algorithm for the vehicle routing problem with time window, 15, 1247–1270
3. Varunika A, Amit G, Vibhuti J (2014) An optimal solution to multiple travelling salesperson problem using modified genetic algorithm 3(1)
4. Chen S-M, Chien C-Y (2011) Parallelized genetic ant colony systems for solving the traveling salesman problem. Expert Syst Appl 38:3873–3883
5. Ray SS, Bandyopadhyay S, Pal SK (2004) New Operators of genetic algorithm for travelling salesman problem. In: Proceedings of the 17th international conference on pattern recognition, vol 2, pp 497–500
6. Yan, X, Zhang1, C (2012) An solve traveling salesman problem using particle swarm optimization algorithm 9(6), No 2
7. Shi XH, Liang YC, Lee HP, Lu C, Wang QX (2007) Particle swarm optimization-based algorithms for TSP and generalized TSP. Inform Process Lett 103:169–176
8. Marinakis Y, Marinaki M (2010) A Hybrid multi-swarm particle swarm optimization algorithm for the probabilistic traveling salesman problem. Comput Oper Res 37(3):432–442
9. Jie B (2012) A model induced max-min ant colony optimization for asymmetric traveling salesman problem. Appl Soft Comput 1365–1375
10. Saenphon T, Phimoltares S, Lursinsap C (2014) Combining new fast opposite gradient search with ant colony optimization for solving travelling salesman problem. Eng Appl Artif Intell 35:324–334
11. Mahi M, Kaan ÖB, Kodaz H (2015) A new hybrid method based on particle swarm optimization, ant colony optimization and 3-Opt algorithms for traveling salesman problem. Appl Soft Comput 30:484–490
12. Nitesh MS, Chawda BV (2013) Memetic algorithm a metaheuristic approach to solve RTSP, IJCSEITR, ISSN 2249-6831, 3(2):183–186
13. George G, Raimond K (2013) Solving travelling salesman problem using variants of abc algorithm. Int J Comput Sci Appl (TIJCSA) 2(01) ISSN-2278-1080
14. Sobti S, Singla P (2013) Solving travelling salesman problem using artificial bee colony based approach. Int J Eng Res Technol (IJERT) 2(6)
15. Marinakis Y, Marinaki M, Dounias G (2011) Honey bees mating optimization algorithm for the Euclidean traveling salesman problem. Inf Sci 181(20)
16. Pathak N, Tiwari SP (2012) Travelling salesman problem using bee colony with SPV. Int J Soft Comput Eng (IJSCE) 2(3)
17. Saranya S, Vaijayanthi RP (2014) Traveling salesman problem solved using bio inspired algorithms (ABC). Int J Innovative Res Comput Commun Eng 2(1)
18. Kan J-M. Yi Z (2012) Application of an improved ant colony optimization on generalized traveling salesman problem. Energy Procedia 17(2012):319–325
19. Rizk-Allah RM, Elsayed M, Ahmed AE (2013) Hybridizing ant colony optimization with firefly algorithm for unconstrained optimization problems. Appl Math Comput 224:473–483
20. Kota L, Jarmai K (2013) Preliminary studies on the fixed destination MMTSP solved by discrete firefly algorithm. Adv Logis Sys 7(2):95–102
21. Yang XS (2009) Firefly algorithms for multimodal optimization. In: Stochastic algorithms: foundations and applications, SAGA 2009, Lecture notes in computer sciences, vol 5792, pp 169–178
22. http://elib.zib.de/pub/mp-testdata/tsp/tsplib/tsp/

IGBT Devices, Thermal Modeling Using FEM

Sonia Hosseinpour and Mahmoud Samiei Moghaddam

Abstract Thermal control and modeling of transistor devices, and especially IGBT transistors, are of interest to many researchers today. Managing thermal devices These transistors can have a significant impact on energy consumption. Considering the importance of the subject, in this research, the FEM method is used to model the thermal devices of IGBT transistors. By simulating the proposed method in this study, it was observed that the proposed method significantly improved the aspects of wasted energy during switching on and off IGBT transistors at different temperatures.

Keywords IGBT transistors · FEM method · Heat and modeling of transistor devices

1 Introduction

This paper presents an Analog behavioral model (ABM) of the Insulated Gate Bipolar Transistor (IGBT) with Orcad Pspice 16.5. The Spice model was built using device parameters extracted through experiment. A full study of switching behavior of IGBT during turn-off and turn-on for inductive load with freewheeling diode is presented and simulated. All simulation results presented in this paper are validated, compared and showed good agreement with the measured data. The temperature dependent behavior is simulated and analyzed. An enhanced insulated gate bipolar transistor (IGBT) model based on the Kraus model with new derivations based on an extra parameter accounting for p-i-n injection was developed to allow simulation of both trench and DMOS IGBT structures. Temperature dependence was also implemented in the model. The model was validated against steady-state and transient

S. Hosseinpour · M. Samiei Moghaddam (✉)
Department of Electrical Engineering, Islamic Azad University,
Damghan Branch, Damghan, Iran
e-mail: samiei352@yahoo.com

S. Hosseinpour
e-mail: s.h.arman.aria@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_32

measurements done on an 800-A 1.7-kV Dynex IGBT module at 25 and 125 C. The Spice model has also shown excellent agreement with mixed mode MEDICI simulations. The Spice model also takes into account for the first time the parasitic thyristor effect allowing the dc and dynamic temperature dependent latchup modeling of power modules as well as their temperature-dependent safe operating area.

2 Theoretical

In this chapter, some of the most important researches in the fields related to the subject of the research are discussed. This chapter introduces a general overview of the research literature, details of the algorithms, variables, and unknowns of research, and other essential elements for further information. The current methods used to determine the instantaneous binding temperature in power tools are not. Power in inverters associated with this field. The instantaneous junction temperature Tj under hard switching conditions is one of the most relevant physical parameters for the design of reliable IGBT devices and systems. Due to the fact that they are basically quasi-static techniques, the methods used nowadays to determine the instantaneous junction temperature in single power devices cannot be used in inverters operated in the field. This still represents a major limitation in the use of the instantaneous Tj as a possible control parameter to define adaptively the working point of power devices within the proper safe operating area. In this paper, we introduce a novel technique, which enables both to measure Tj under real operating conditions and to determine the thermal impedance of the device for very short power pulses. This technique is mainly based on the use of dIce/dt as a thermo sensitive parameter. After reviewing the physical background, the proposed technique, is applied to measure the instantaneous junction temperature in 3.3 kV 1200 A.

2.1 Research Hypotheses

1. Heat sink temperature and IGBT case are considered constant.
2. The ambient temperature variations are ignored.
3. This model is considered for single-chip IGBT devices.

2.2 Outstanding Research Objectives

1. Analysis of the methods used in thermal modeling of power electronic devices.
2. Analyze the selected thermal model using state space.
3. Simulation of thermal model by FEM method.

2.3 Types of Work Methods

1. Review various articles on the subject.
2. Choosing the right method for modeling.
3. Simulation of the selected method.

3 Relevant Records

Power electronics is considered as an electrical power conversion technology from one form to another using electronic power tools. Extensive development over the past few years, with advances and innovations in semiconductor power tools, power conversion technologies, microprocessors, DSPs, application-specific ICs, personal computers, CAD tools, control and evaluation techniques in this field [1]. The key component in the power conversion system is the semiconductor or semiconductor power device, which acts as a power switch. Based on some estimates, more than 60% of the electricity is consumed in the United States. It means that at least one power tool or most devices are connected to power [2]. Improving semiconductor tools as a driving force in relation to advanced performance, efficiency, size and weight of power conversion systems. It comes in. The need for an ideal instrument for the semiconductor of power in fact includes the ability to control the current flow due to its load and zero losses. In conductivity, such an instrument should have unlimited flow ability, while in blocking mode, this system should have an unlimited blocking voltage capacity. In addition, the switching speed between different modes should be considered quickly. In a quiet paper, three-dimensional heat transfer simulation is implemented for atypical electronically-powered equipment and cooling system Equipped with a high-power three-phase inverter manufactured by Symikron Co., the main application of which is electric and hybrid vehicles. The cooling system is a thermosetting machine with a straight rectangular blade with a uniform cross-section that cools through convection. The limiting factor is the design of the heat transfer system, the high temperature of the mills and, in other words, the thermal resources available on the inverter, is abbreviated as IGBT. The IGBTs temperature should be below 125 °C to prevent thermal and mechanical failure. One of the main goals is to reduce the maximum temperature with the precise design of chip layout. Geometric design is done in accordance with the thermal constraints of the chips and the compromise between the volume of consumables and the heat output. The geometric parameters are the number, height and thickness of the blades and the thickness of the base plate. Heat loss power losses are calculated accurately with simulation in MATLAB software and technical information provided by the manufacturer. The thermal model of the inverter and its cooling system are implemented by finite element method. The accuracy of the calculated thermal and thermal modeling is confirmed by the Samuels software. The precise design of the layout results in a significant reduction in the maximum temperature of the chips to a C20 value. The heat transfer efficiency with the proper design for the heat transfer coefficients of K m^2/W 50, W/m^2 of the thermo former was 27.21 and 16.66, 35.50, 0, 51.13, 52/22%. The volume of increase has been increased by 100 W/m^2 K and 75 K, respectively. Lee introduced the recommendations for the selection of Thermabr and proposed the use of the flow-flow relationship. Ning presented a thermal model based on analytic equations for heat exchanger, fan, channel and return air flow. Based on this model, the cooling system was optimized to achieve the lowest weight. Drofnick examines the constraints of the power density theory of a power converter with the forced air convection cooling system and optimizes the power density by using empirical equations and an analytical model of IGBT modules.

4 IGBT Structure

Figure 1 shows a cross section schematic of a typical IGBT. Figure 3 shows the discrete equivalent circuit model of the device, which consists of a wide base P-N-P bipolar transistor (BJT) in cascade with a MOSFET. The structure of the device is similar to that of a vertical double diffused MOSFET with the exception that a highly doped p-type substrate is used in lieu of a highly doped type drain contact in the vertical double diffused MOSFET. A lightly doped thick n-type epitaxial layer (N ≈ 10^{14} cm^{-3} B) is grown on top of the p-type substrate to support the high blocking voltage in the reverse bias mode state. A highly doped p-type region (N ≈ 10^{19} cm^{-3} A) is added to the structure to prevent the activation of the PNPN thyristor during the device operation. The power MOSFET is a voltage-controlled device that can be manipulated with a small input gate current flow during the switching transient. This makes its gate control circuit simple and easy to use.

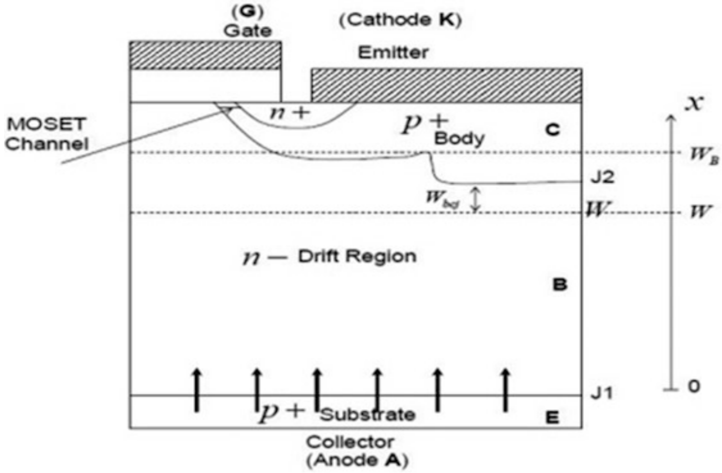


Fig. 1 Cross section schematic of the IGBT Half-cell [1]

A highly doped n + buffer layer could also be added on top of the highly doped p + substrate. This layer helps in reducing the turn-Off time of the IGBT during the transient operation. The IGBT with a buffer layer is called a punch through PT IGBT while the IGBT without a buffer layer is named a non-punch through NPT IGBT.

5 Physics of IGBT

From the cross-section of the IGBT (Fig. 1), we can see that the IGBT is a four-layer power semiconductor device having a MOS gate. When a negative voltage is applied to the collector with respect to the emitter, there will be no current flow through the device for the lower junction (J1) is reverse biased. This provides the reverse blocking capacity of the IGBT. When the gate is attached to the emitter (V_{GE} = 0 V), and a positive voltage is applied to the collector with respect to the emitter (the same voltage as gate), the upper junction (J2) is reverse biased and the device operates in the forward blocking mode. If a positive voltage higher than the threshold voltage is now applied between the gate and the emitter (V_{GE} > Vth), the surface of the P base region is inverted and an N channel will appear. The electrons will then flow from the N+ emitter to the N drift region forming the base current for the vertical P-N-P transistor of the IGBT. An increase in the positive voltage between the collector and emitter leads to an increase in the injected holes concentration until it exceeds the background doping level of the N drift region. In this region of operation, the device behaves like a P-I-N device and this explains the IGBT's ability to handle high current densities. If we further increase the voltage between the collector and the emitter, the N-channel will get pinched-off. The base current for the P-N-P transistor will be limited and so will the hole current through the path. The collector to emitter current reaches the saturation point and the IGBT operates in the active region. The output characteristics of the IGBT are similar to that of the MOS and the output current is controlled by the gate voltage V_{GE}. The I-V characteristics of IGBT are shown in Figure. From the above, one concludes that IGBT integrates the physics of MOS and bipolar junction transistors. The P-I-N behavior of the BJT part provides high forward conduction density, and the MOS gate structure determines the low drive power and fully gate-controlled output characteristics.

6 Introduction

Advanced IGBT transistors have the ability to operate at high voltages and currents and to reduce the average direct conductivity. These transistors are used in low power applications with moderate power up to a few megawatts and even tens of megawatts. Functionally, IGBTs can be considered as a combination of BJT, MOSFET, and GTO transistors. The IGBT transistors feature high-speed, low-loss switching from the MOSFETs, losses, and low direct conduction voltage losses from bipolar transistors and high-voltage failure thyristors. These specifications make them a key to many electronics applications. Therefore, in this section, the structure of these transistors and their thermal modeling are based on the FEM method.

Therefore, in this chapter, the full description of the proposed method, with details of each section, is described with full description of the flowchart method, and also the proposed algorithm is described. At the beginning of this chapter, the flowchart and proposed research proposal are examined, then, according to the steps and details presented in this flowchart, sections of the proposed method are described. At the end of the algorithm, the proposed method is presented in pseudo code format.

7 Describe the Proposed Method

The most important parts of the proposed method in this research will be presented below the relevant sections. The proposed method includes the sections discussed below.

8 Define IGBT Transistor Structure

Before simulating the main issue in the proposed method of thermal modeling of the IGBT transistor, it is necessary to define the structure of this transistor in the simulation. So before defining it in simulation, its structure will be discussed in the following section. Figure 2 shows the silicon cross section of an IGBT.

Its structure is like a power mosfet, except that the n + layer has been replaced by a p + substrate in the mosfet called the collector. This layer creates a pn link that facilitates the transport of minority carriers to the collector. When the gate voltage exceeds the voltage threshold of the IGBT, a n channel is formed in the P region. If the collector voltage is positive to the emitter, the key is in direct guidance mode. In this case, the p-substrate injects the cavities into the epitaxial n-layer. By increasing the collector voltage to the emitter, the focus of the injected cavities is increased and finally a direct current is established. Other equivalent circuits and structures used for IGBT are shown in Fig. 3.

8.1 Switching Characteristics

8.1.1 Turn-On

The turn-on switching characteristics for an IGBT transistor are similar to that of a MOSFET. The whole process is indicated in the Fig. 4.

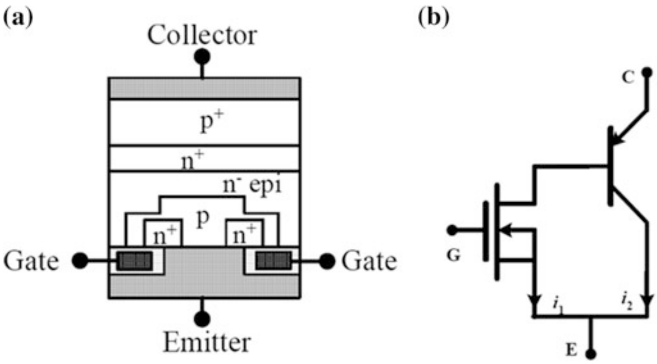


Fig. 2 Cross section silicon section IGBT [1]

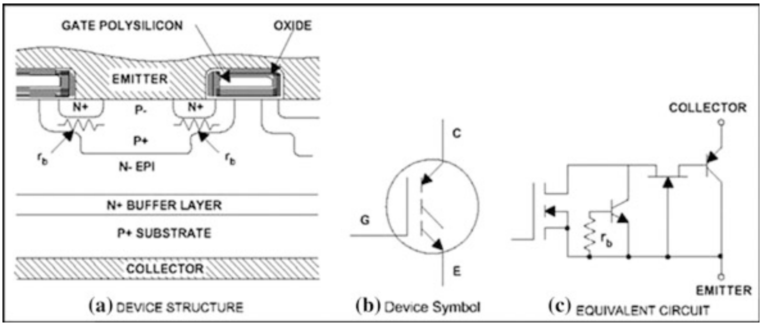


Fig. 3 Other equivalent circuits and structures used for IGBT [2]

As shown in the Fig. 4, during the delay time td(on) the gate-emitter voltage increases to the threshold voltage VGE(th) of the device. This is caused by the gate resistance Rg and the input capacitances (CGC and CGE). But the miller effect capacitance, CGC is very small that its affect can be neglected. Beyond this time the collector current starts to increase linearly until it reaches to the full load current. During the time when iC = I0 the voltage VGE is first kept constant and at this moment the collector current flows through CGC only, that causes that the voltage VCE decreases to the zero on-state. After this moment the voltage VGE starts to increase until it reaches to VGG [4]. Turn-on switching losses: Turn-on switching losses are the amount of total energy losses during turn on under inductive load. It is normally measured from the point where the collector current starts to flow to the point where the collector-emitter voltage drops completely to zero. The turn on energy losses calculation is given by the equation below. Turn-Off The whole turn-off process is indicated in the Fig. 5.

The transistor is turned off by removing the gate voltage, VGE. As shown in the figure, both the voltage these losses include loss of an amount of total energy during the switch-off mode, under the induction charge. These damages will be measured

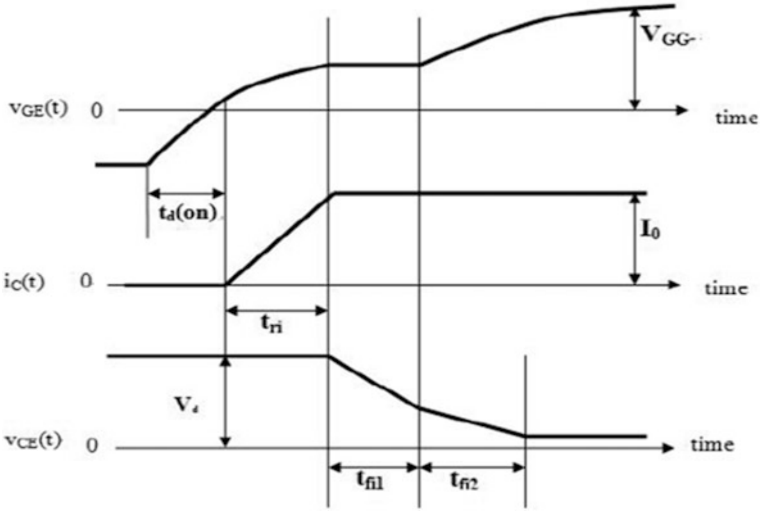


Fig. 4 IGBT turn-on switching characteristics [2]

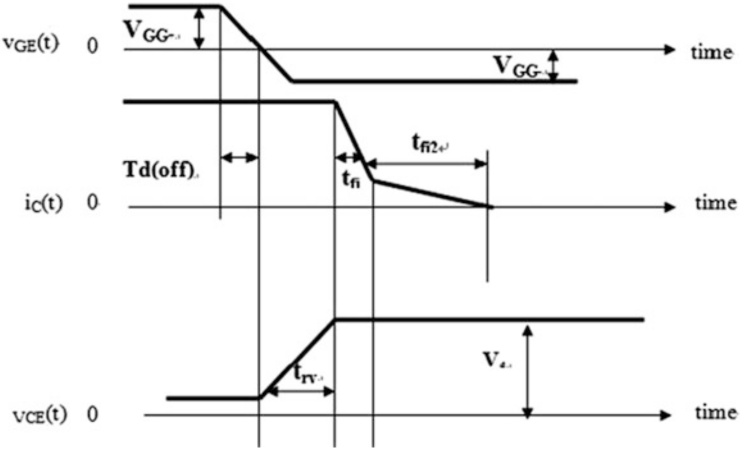


Fig. 5 IGBT turn-off switching characteristics [2]

from the point where the collector voltage starts up until it reaches a point where the collector current reaches completely zero. Therefore, after all of the above-mentioned conditions and structure are simulated, FEM method has to be implemented on these transistors and apply thermal modeling. By doing so, the amount of energy lost in transistors will be reduced by heat. In the next section, the steps to implement the FEM algorithm on IGBT transistors are described in terms of the thermal modeling of these transistors.

9 Applying the FEM Algorithm for IGBT Thermal Modeling

Finite Element Method (FEM) is a numerical method for solving the approximate differential equations and solving integral equations. The practical application of finite element is usually called finite element analysis. The basis of this method is the complete elimination of differential equations or their simplification into ordinary differential equations, which are solved by numerical methods such as Euler. In solving partial differential equations, the important problem is to arrive at a simple equation which is numerically stable—in other words, the error in the initial data and during the solution is not so much that results in inaccurate results. There are methods with various advantages and disadvantages for this, in which the finite element method is one of the best. This method is very useful for solving partial differential equations on complex domains (such as vehicles and oil pipelines), or when the range is variable, or when high precision is not required anywhere in the domain or if the results do not have sufficient solidity and uniformity. For example, in simulating an accident in the front of the car, there is no need for high accuracy at the rear of the car. In air simulation and forecasting on the planet, dry weather is more important than air on the sea. Dividing the area into smaller regions has many advantages, including: detailed representation of complex geometry, the capacity of different body features, and the understanding of the local features of the object. Therefore, in this study, the FEM method is used for thermal modeling of IGBT transistors [3]. Finally, the remainder of the operation is explained in the proposed method.

10 Simulation and Evaluation of Results

In order to simulate and model the thermal devices of IGBT transistors with the help of the FEM method, the structures of these transistors are first simulated by modeling tools such as MATLAB, then the process required by the IGBT transistors and finally for the thermal modeling of the transistor devices. It simulates the FEM method. So according to simulations, the tool used to implement the proposed method in this study is MATLAB. This chapter first introduces the relevant data and describes its features. After simulating the proposed method, all the results and findings are described and presented in various diagrams. Finally, the results are compared with other methods that have been proposed so far.

11 MATLAB Simulation Environment

MATLAB is a high-level programming language for the fourth generation and an interactive environment for numerical computing, visualization and programming that combines two matrix words (MATRIX) and a lab (LABoratory). This name implies a program-oriented matrix approach, in which even single numbers are considered as a 1 * 1 dimensional matrix. MATLAB software is produced by MathWorks. The company was founded in 1984 in the state of Massachusetts. In 1970 Cleve Moler, director of the New Mexico School, wrote MATLAB software based on Fortran's language. In 1983, this software was developed based on the C programming language and began to expand after the establishment of its expansion company. MATLAB provides the ability to work with matrices, customize functions and data, implement algorithms, create user interfaces, communicate with programs written in other languages, including C, C++, JAVA and Fortran, and create models and applications. The MATLAB system is made up of five main parts.

1. MATLAB Language: MATLAB is a high-level language high-level array matrix that incorporates object-oriented programming features that can be used to create simple and complex programs.
2. MATLAB work environment: A set of tools and tools that you interact with as a MATLAB or MATLAB author. This environment includes options for managing variables in the workspace and a tool for developing, managing, fixing and creating M files in MATLAB applications.
3. Graphics control: The same is the MATLAB graphics system, which includes high level commands for visualizing 2D and 3D data, image processing, animation and graphics. It also includes low level commands that allow you to customize the graphics appearance of your applications.
4. MATLAB Math Functions Library: A large collection of computational algorithms, including basic functions such as sine, cosine, and complex functions such as the inverse matrix, special matrix, and fast Fourier transform.
5. The Matlab Application Interface (API) is a library that allows you to write Fortran and C applications that interact with MATLAB. This interface includes features such as Metalb calling (dynamic connection), MATLAB calling as a computing engine, and reading and writing to MAT files [4].

12 Simulation and Experimental Results

The proposed method has been implemented using MATLAB simulator. Table 1 also shows the characteristics of the system that implements the proposed method and evaluates the results.

Therefore, according to a system with the above specifications, the simulation is performed and the results are evaluated. Therefore, in this section, the results are fully described.

12.1 Results of the Simulation of the Proposed Method

After the necessary backgrounds for thermal modeling of the IGBT transistor or bipolar transistor with isolated gate were provided, FEM method was applied to

Table 1 System specifications for simulation and evaluation of results [4]

Profile	Hardware/software
Operating system	Windows 7
Operating system type	32-bit operating system
RAM memory	4 gigabytes of RAM—3.06 gigabytes usable
Processor	Intel processor—Number of cores 7 (Core™ i7 CPU)—Q 720 @ 1.60 GHz 1.60 GHz
Transistor type	Transistor IGBT or Isolated gate bipolar transistor
Method used by	FEM method
Modeling tool	MATLAB

these transistors and results were obtained. In this section, we will provide graphs and drawings based on the paper [4].

– **Thermal Effects on Energy Loss**

In an article that is used in this paper as the basis for comparing and evaluating the results of the proposed method, it is observed that the losses (lost energy) entered during the switch-off of the IGBT transistor were almost constant. This result has been relatively more accurate in our simulations. In Fig. 6, the results are related to the loss of energy when IGBT transistor is turned on.

As seen from Fig. 6, which is related to the base paper, the amount of energy lost per unit mJ is shown by the temperature changes of the IGBT transistor. By increasing the temperature, the amount of energy lost will increase significantly. At the time the IGBT transistor reaches 100°, the energy lost is 4.8 mJ. On the other hand, this simulation also displays a similar kind of turn-off at shutdown time. In the following figure, the results of shutdown and temperature changes of the IGBT transistor are shown in the reference paper.

As seen from Fig. 7, which is related to the base paper, the amount of energy lost per unit mJ is shown by the temperature changes of the IGBT transistor. By increasing the temperature, the amount of energy lost will increase significantly. At the time the IGBT transistor reaches 100°, the energy lost is 4.4 mJ. The amount of

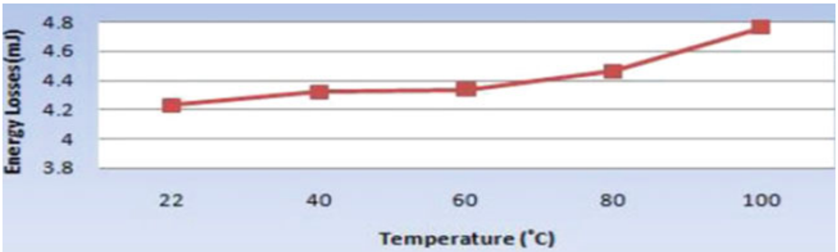


Fig. 6 Energy loss when IGBT transistor is turned on at different temperatures in the source paper

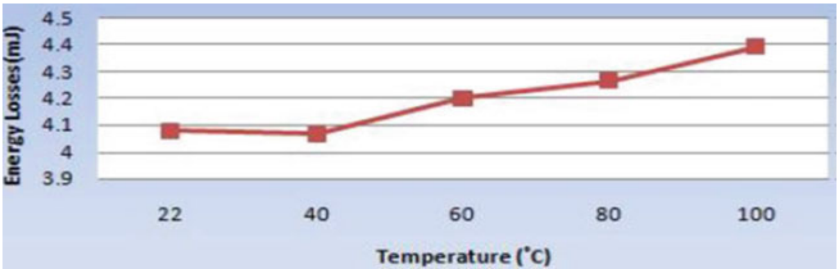


Fig. 7 Energy loss when the IGBT transistor is turned off at different temperatures in the source paper

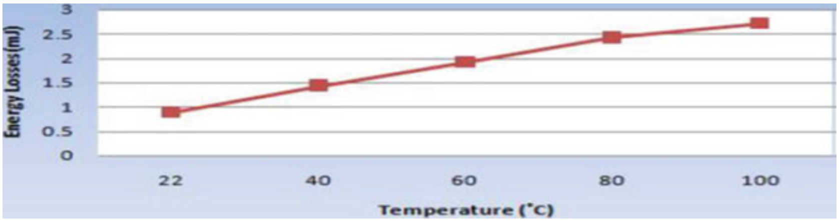


Fig. 8 Energy loss in free-wheeling diode

lost energy loss in this case is reduced to about 0.4 when the transistor is turned on. The energy of the Free-wheeling Diode in the source paper is described in Fig. 8.

As seen from Fig. 8, as the temperature rises, the lost energy is also increased, but it is less energy-efficient than temperature changes in the previous figures. But in our simulations, there are similar and even better examples of source simulations that can be seen in the following illustrations.

As can be seen from Fig. 9, the improvement in lost energy versus the increase in temperature up to 100 °C during switching on IGBT transistors at different temperatures with the FEM method was significantly reduced compared to the method in the base paper. As can be seen, at 20–40°, the amount of energy lost in the proposed method and the method in the base paper is almost the same. At 50–100°, the energy changes in the proposed method were much lower than the base paper, and did not have a strong upward trend contrary to the base method. Therefore, we can trust the proposed method, which uses the FEM algorithm for thermal modeling of IGBT transistors. In Fig. 10, the results of wasted energy during the shutdown of IGBT transistors at different temperatures with the FEM method are shown in comparison with the method in the base paper.

As can be seen from Fig. 10, the improvement in lost energy versus the increase in temperature up to 100 °C during switching on IGBT transistors at different temperatures with the FEM method was significantly reduced compared to the method in the base paper. As can be seen, at 20–40°, the amount of energy lost in the proposed method and the method in the base paper is almost the same. At 50–100°, the energy changes in the proposed method were much lower than the base

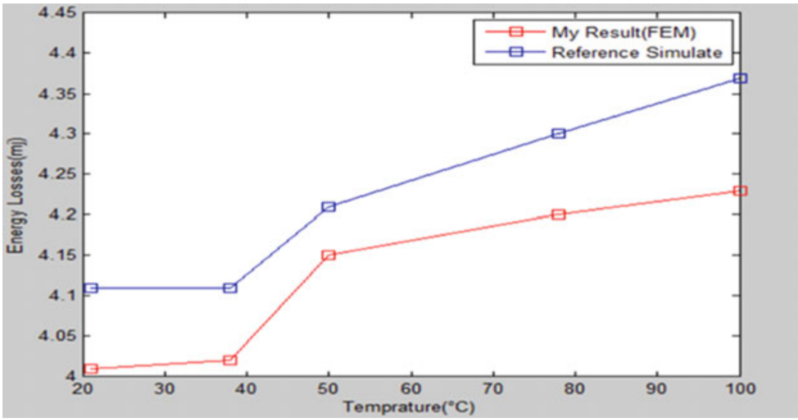


Fig. 9 Waste energies when IGBT transistors are turned on at different temperatures using the FEM method compared to the method in the base paper

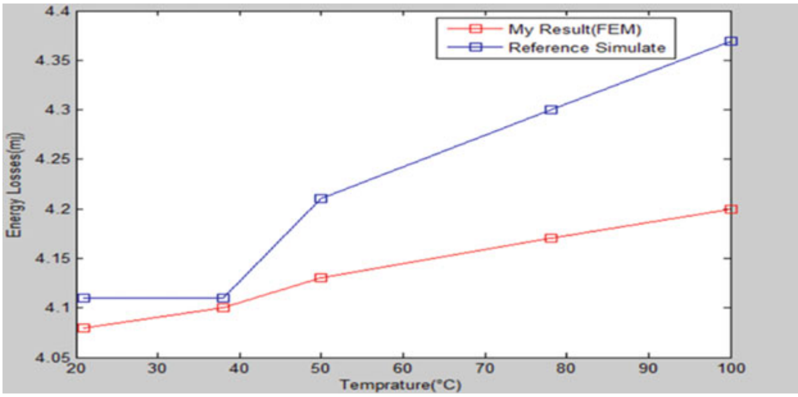


Fig. 10 Waste energies when switching off IGBT transistors at different temperatures using the FEM method compared to the method in the base paper

paper, and did not have a strong upward trend contrary to the base method. Therefore, in general, the proposed method in this study improved the aspect of wasted energy during the shutdown of IGBT transistors at various temperatures using the FEM method with the method in the base paper at about 0.083 mJ. In the case of the constant loss of energy at the time of switching off and illuminating the simulation sample, this study estimates that the maximum energy loss difference is 0.3–0.5 mJ, which is less than 0.1 mJ less than the reference sample, and is generally wasted about 0.15 mJ There is less energy to simulate.

In Fig. 11, the wasted energy in mJ at various temperatures in the Free-wheeling diode shows the IGBT transistor circuit in the FEM method and the method in the base paper.

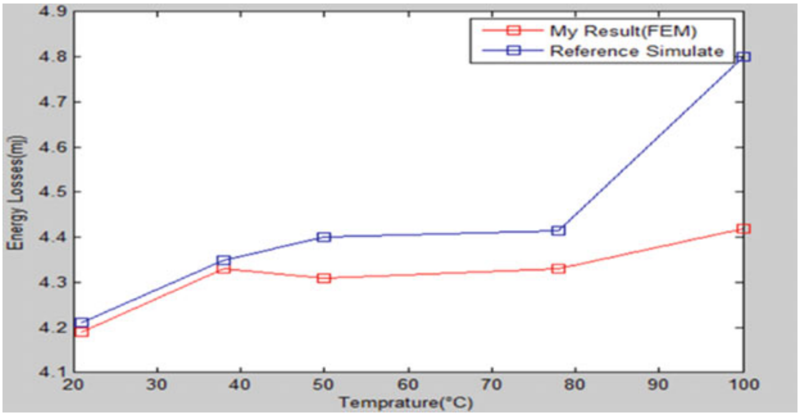
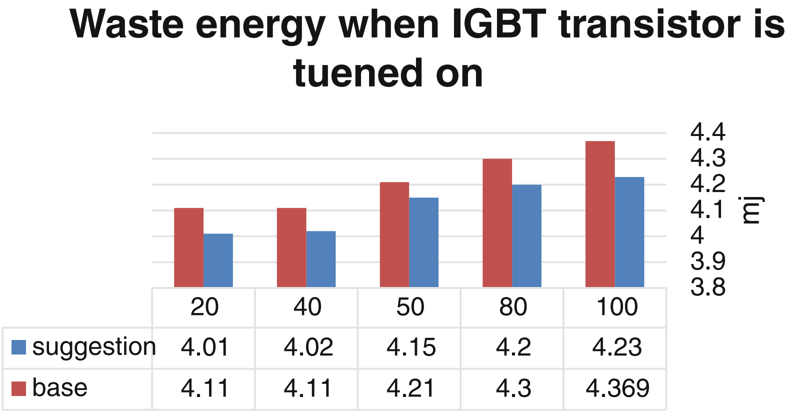


Fig. 11 The wasted energy at different temperatures in the Free-wheeling diode IGBT transistor circuit by FEM method and the method in the base paper



	20	40	50	80	100
🟦 suggestion	4.01	4.02	4.15	4.2	4.23
🟥 base	4.11	4.11	4.21	4.3	4.369

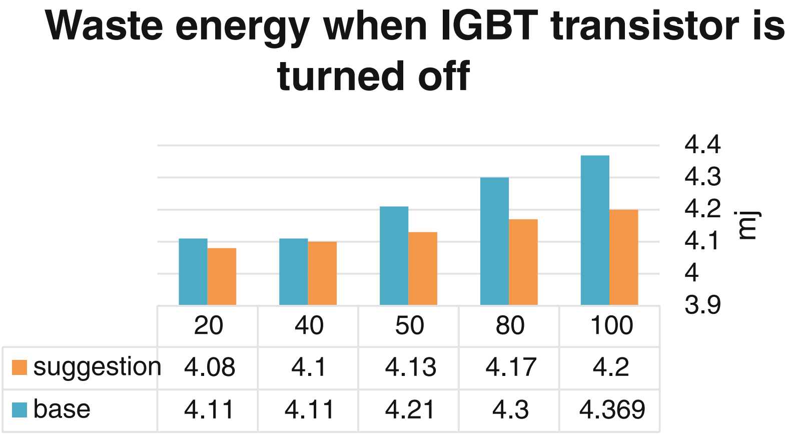
Fig. 12 Comparison of the average of wasted energy when switching on IGBT transistors at different temperatures using the FEM method compared to the method in the base paper

Figure 11 is similar to the results presented in the previous figures, which was obtained from the simulation in the IGBT transistor switching mode and in the FEM method, which has less variation in terms of the energy lost than the base paper. The improvement in the proposed method compared with the method used in the base paper in terms of energy lost at different temperatures in the diode has improved the round IGBT transistor circuit at about 0.119 mJ. Therefore, it can generally be concluded that the use of the FEM method can be very effective in the thermal modeling of IGBT devices. In Fig. 12, the comparison of the average of wasted energy during the activation of IGBT transistors at different temperatures with the FEM method is shown in comparison with the method in the base paper.

As can be seen, the improvement in the proposed method has improved by about 0.097 mJ compared with the method in the base paper. In Fig. 13, the comparison of the average of wasted energy during the shutdown of IGBT transistors at different temperatures with the FEM method is shown in comparison with the method in the base paper.

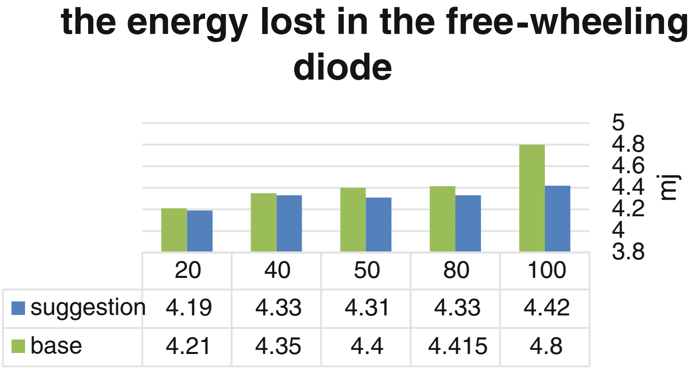
As can be seen, the improvement in the proposed method has improved by about 0.083 mJ compared with the method in the base paper (Fig. 14). Comparison of the average energy lost at various temperatures in the Free-wheeling diode of the IGBT transistor circuit is illustrated by the FEM method and the method found in the base paper.

As can be seen, the improvement in the proposed method has improved by about 0.119 mJ compared with the method in the base paper.



	20	40	50	80	100
🟧 suggestion	4.08	4.1	4.13	4.17	4.2
🟦 base	4.11	4.11	4.21	4.3	4.369

Fig. 13 Comparison of the average of wasted energy during the shutdown of IGBT transistors at different temperatures using the FEM method compared with the method in the base paper



	20	40	50	80	100
🟦 suggestion	4.19	4.33	4.31	4.33	4.42
🟩 base	4.21	4.35	4.4	4.415	4.8

Fig. 14 Comparison of the average wastage energy at different temperatures in the Free-wheeling diode of IGBT transistor circuit by FEM method and method in the base paper

12.2 Overview of Research Findings

In general, the findings of the research are discussed here. Some findings from this research include:

- The improvement in lost energy versus the method found in the base paper significantly reduced the energy lost by up to 100° during the switching of the IGBT transistors at different temperatures using the FEM method. As can be seen, at 20–40°, the amount of energy lost in the proposed method and the method in the base paper is almost the same. At 50–100°, the energy changes in the proposed method were much lower than the base paper, and did not have a strong upward trend contrary to the base method.
- The improvement in lost energy versus the method found in the base paper significantly reduced the energy lost by up to 100° during the switching of the IGBT transistors at different temperatures using the FEM method. As can be seen, at 20–40°, the amount of energy lost in the proposed method and the method in the base paper is almost the same. At 50–100°, the energy changes in the proposed method were much lower than the base paper, and did not have a strong upward trend contrary to the base method. Therefore, in general, the proposed method in this study improved the aspect of wasted energy during the shutdown of IGBT transistors at various temperatures using the FEM method with the method in the base paper at about 0.083 mJ.
- The improvement in the proposed method compared with the method in the base paper in terms of energy lost at various temperatures in the Free-wheeling diode has improved the IGBT transistor circuit at about 0.119 mJ.
- The improvement in the proposed method has improved by about 0.097 mJ compared to the method in the base article.
- The improvement in the proposed method has improved by about 0.083 mJ compared to the method in the base paper.
- The improvement in the proposed method has improved by about 0.119 mJ compared with the method in the base paper.

With the results obtained from the simulation of the proposed method, it was observed that the FEM method could simulate the thermal modeling of IGBT transistor devices to a lesser amount of energy.

13 Conclusion

The FEM method is one of the most popular methods of transistor modeling in electronics and electrotechnics. This method provides facilities for modeling transistor devices such as IGBT. Regarding the purpose of this study, which is the thermal modeling of IGBT transistor devices, we have achieved a series of results simulating the proposed method to improve heat and reduce waste energy, which

Table 2 Comparison of the energy loss ratio of the proposed method with the mean of the base method

	The wasted energy when IGBT transistors are turned on				
	20	40	50	80	100
Suggested method	4.01	4.02	4.15	4.2	4.23
Reference article method	4.11	4.11	4.21	4.3	4.369
	The wasted energy when IGBT transistors are turn off				
	20	40	50	80	100
Suggested method	4.08	4.1	4.13	4.17	4.2
Reference article method	4.11	4.11	4.21	4.3	4.369
	The wasted energy in the Free-wheeling diode				
	20	40	50	80	100
Suggested method	4.19	4.33	4.31	4.33	4.42
Reference article method	4.21	4.35	4.4	4.415	4.8

In the reference paper, the aspects of the amount of energy lost during switching on and off the IGBT transistors and the Free-wheeling Diode

briefly describes the final results in this section. In Table 2, the comparison of the amount of energy lost by the proposed method with an average compared to the base method in the reference paper has been shown in terms of the amount of energy lost during switching on and off the IGBT transistors and the Free-wheeling Diode.

As can be seen from the Table 2, the results of the proposed method, which was performed with the FEM, have improved significantly compared to the results of the method in the article. Therefore, using the FEM method to improve the wasted energy and thermal modeling of IGBT transistor devices can be used and see acceptable responses.

13.1 Future Suggestions

Here are some of the suggestions that can be made as new ideas and ideas to improve the performance of the proposed method in this research, some of which are:

- Using optimization methods such as genetic intelligence algorithms, particle swirling, frog mutations, etc. to reduce the lost energy in IGBT transistors and compare the results with the findings of this research.
- Using machine learning techniques, such as backup vector machine and neural network for modeling IGBT rays devices.
- Improvement of FEM method with the help of biological methods such as Dragonfly Algorithm, Gray Wolf, Cat Algorithm, etc. and compare the results with the results of this research.

Annex

```
% FEM solution of bvp: d^2/dx^2u + 4pi^2u = 0; u(0) = 0; du/dx(1) = 1 using
% quadratic 1D elements
% See also http://www.particleincell.com/blog/2012/finite-element-examples
%==========================================================
===============
clear all
close all
% Store coordinates in
p================================================
p=linspace(0,1,10)';
% Connectivity
t=[1 2
2 3
3 4
4 5
5 6
6 7
7 8
8 9
9 10];
%==========================================================
===============
TotalNumberOfNodes=size(p,1);
NumberOfElements=size(t,1);
% c(x), f(x) en lambda
c=@(x)1*ones(size(x));
f=@(x)0*ones(size(x));
labda=4*pi^2;
%==========================================================
===============
% Quadratic elements contain 3 nodes per segment: add nodes in the middle
% of eacht segment
TotalNumberOfNodes=size(p,1);
NumberOfElements=size(t,1);
S=zeros(TotalNumberOfNodes);
counter=TotalNumberOfNodes+1;
for e=1:NumberOfElements
nodes=t(e,:);
if (S(nodes(1),nodes(2))==0)
S(nodes(1),nodes(2))=counter;
S(nodes(2),nodes(1))=counter;
p(counter,:)=mean(p([nodes(1) nodes(2)],:));
counter=counter+1;
end
```

```
t(e,3)=S(nodes(1),nodes(2));
end
%
Update==========================================================
===========
TotalNumberOfNodes=size(p,1);
NumberOfElements=size(t,1);
% Initialisation of K, M and
F=============================================
K=zeros(TotalNumberOfNodes);
M=zeros(TotalNumberOfNodes);
F=zeros(TotalNumberOfNodes,1);
%============================================================
===============
% lambda, lambda * two outer nodecoordinates yields integrationpoints
lambda=1/2*[
1-1/sqrt(3) 1+1/sqrt(3);
1+1/sqrt(3) 1-1/sqrt(3)];
% weights of Gaussian quadrature
w_Omega(1:2,1)=1;
hold on
% Loop over all elements
for e=1:NumberOfElements
nodes=t(e,:);
% 3 by 3 matrix with rows: [ones;x;x^2]
P=[ones(1,3);p(nodes,:)';p(nodes,1)'.^2];
length=norm(diff(p(nodes([1:2]),:)));
% Determine the two coordinates of the integration points on the
% edge of the element belonging to the Neuman boundary
ip=lambda*p(nodes([1:2]),:);
% plot integration points
% plot(ip,0,'ko','MarkerSize',5,'MarkerFaceColor','y');
% 3 by 2 matrix with rows: [ones;x;x^2] of two integration points
I=[ones(1,2);ip';ip(:,1)'.^2];
Phi=P\\I;
%============================================================
=============
Ix=[zeros(1,2);ones(1,2);ip(:,1)'*2];
diffI=Ix;
diffPhi=P\\diffI;
%============================================================
=============
cvalue=c(ip);
fvalue=f(ip);
Ke=zeros(size(P));
```

```
Me=zeros(size(P));
Fe_Omega=zeros(size(P,2),1);
% Integrate and compute element stiffness matrix (3 by 3, 3 nodes)
for i=1:size(Phi,2)
Ke=Ke+w_Omega(i)*cvalue(i)*diffPhi(:,i)*diffPhi(:,i)'*length/2;
Me=Me+w_Omega(i)*Phi(:,i)*Phi(:,i)'*length/2;
Fe_Omega=Fe_Omega+w_Omega(i)*fvalue(i)*Phi(:,i)*length/2;
end
% Assembly
K(nodes,nodes)=K(nodes,nodes)+Ke;
M(nodes,nodes)=M(nodes,nodes)+Me;
F(nodes)=F(nodes)+Fe_Omega;
end
% Proces Neumann
boundary=================================================
F(10)=F(10)+c(1)*1;
% Proces Dirichlet
boundary===============================================
K(1,:)=0;
K(:,1)=0; % symmetry
K(1,1)=1;
M(1,:)=0;
M(:,1)=0;
M(1,1)=1;
% Set the Dirichlet boundary (r.h.s.: value of fixed displament)===========
F(1)=0;
% Solve
system=======================================================
=====
U=(K-labda*M)\\F;
% plot exact and FEM
solution=============================================
box on
plot(p(1:10),1/2*sin(2*pi*p(1:10))/pi,'r','Linewidth',4)
plot(p(1:10),U(1:10),'o-g','Linewidth',2,'MarkerSize',8,'MarkerFaceColor','g');
figure (7);
plot([21 38 50 78 100],[4.19 4.33 4.31 4.33 4.42],'rs-'),ylabel('Energy
Losses(mj)'),xlabel('Temprature(°C)');
hold on;
plot([21 38 50 78 100],[4.21 4.35 4.4 4.415 4.8],'bs-');

legend('My Result(FEM)','Reference Simulate');
%figure 7.3
% hold on;
% plot([21 38 50 78 100],[0.7 1.2 2 2.3 2.6],'rs-'),ylabel('Energy
Losses(mj)'),xlabel('Temprature(°C)');
% plot([21 38 50 78 100],[0.74 1.23 2 2.4 2.7],'bs-');
% legend('My Result','Reference Simulate');
```

```
figure (7
plot([21 38 50 78 100],[4.19 4.33 4.31 4.33 4.42],'rs-'),ylabel('Energ
Losses(mj)'),xlabel('Temprature(°C)');
hold oı
plot([21 38 50 78 100],[4.21 4.35 4.4 4.415 4.8],'bs-'

legend('My Result(FEM)','Reference Simulate'
%figure 7.
% hold oı
% plot([21 38 50 78 100],[0.7 1.2 2 2.3 2.6],'rs-'),ylabel('Energ
Losses(mj)'),xlabel('Temprature(°C)');
% plot([21 38 50 78 100],[0.74 1.23 2 2.4 2.7],'bs-'
% legend('My Result','Reference Simulate'


figure(5);
plot([21 38 50 78 100],[4.08 4.1 4.13 4.17 4.2],'rs-'),ylabel('Energ
Losses(mj)'),xlabel('Temprature(°C)');
hold oı
plot([21 38 50 78 100],[4.11 4.11 4.21 4.30 4.369],'bs-
legend('My Result(FEM)','Reference Simulate'
figure (5
plot([21 38 50 78 100],[4.01 4.02 4.15 4.2 4.23],'rs-'),ylabel('Energ
Losses(mj)'),xlabel('Temprature(°C)');
hold oı
plot([21 38 50 78 100],[4.11 4.11 4.21 4.30 4.369],'bs-
legend('My Result(FEM)','Reference Simulate'
clf;
hold oı
plot([0 500],[-50 -50]
plot([0 195 210 225 240 255 258 263 270 290 305 325 500],[0 0 5 25 25 400 420 41
360 398 402 390 390],'r'
c=0:1:47;
m=inline(vectorize('(-(x^2))/20+75'),'x');
mm=m(c);
plot([0 195],[75 75]
plot(195:242,mm);
c2=1.23*pi:pi/15:1.5*pi;
plot([242 244 246 247 249],sin(c2)*50
%clf;
%hold oı
%plot([0 1],[-.5 -.5]
%c=sin(-pi/2+.5:pi/10:pi/2)./2;
%d=1:.1:2;
%plot(d,c);
```

```
%figure 5.2
clf;
hold on;
plot([0 1000],[450 450]);
plot([0 1000],[-50 -50]);
plot([50 100 200 300],[400 400 400 400],'r');
plot([300 320],[400 390],'r');
plot([320 350],[390 350],'r');
plot([350 380],[350 340],'r');
plot([380 400],[340 300],'r');
plot([400 455],[300 45],'r');
plot([455 468],[45 20],'r');
plot([468 1000],[20 17],'r');
plot([0 321],[0 0]);
plot([321 395],[0 70]);
plot([395 403],[70 70]);
plot([403 407],[70 69]);
plot([407 1000],[69 69]);

%plot([2 3],[sin(1.4) sin(1.4)]-1);
figure (7);
plot([21 38 50 78 100],[4.19 4.33 4.31 4.33 4.42],'rs-'),ylabel('Energy
Losses(mj)'),xlabel('Temprature(°C)');
hold on;
plot([21 38 50 78 100],[4.21 4.35 4.4 4.415 4.8],'bs-');

legend('My Result(FEM)','Reference Simulate');
```

Created with MATLAB® Mobile™

References

1. Mohan N, Undeland TM, Robbins WP (2003) Power electronics converters, applications and design. Wiley
2. Volakis JL, Chatterjee A, Kempel LC (1998) Finite element method for electromagnetics: with application to antennas, Microwave Circuit, and Scattering. IEEE Press, New York
3. Ishak KA (2012) Matlab Tutorial of fundamental programming. Department of Electrical, Electronic & System Engineering Faculty of Engineering University Kebangsaan Malaysi
4. Lotfi M, Zohir D (2016) Modeling of the New Transient Behavioral Spice Model of IGBTs including temperature effect. Int J Hybrid Inf Technol 9(1):141–152. http://dx.doi.org/10.14257/ijhit.2016.9.1.13

Part IV
Power Engineering

An Overview on the Probabilistic Safety Assessment (PSA), the Loss of External Power Source Connected to the Nuclear Power Plant

Mohsen Ahmadnia and Farshid Kiomarsi

Abstract Loss of Offsite Power source connected to the nuclear power plant (LOOP) as the starter event and consequently the probability Station Blackout (SBO) are among accidents that are considered in the Probabilistic Safety Assessment (PSA) and have a significant impact on the melting frequency of the reactor core. Thus, in this paper, the Probabilistic Safety Assessment (PSA), the loss of the external power source in the nuclear power plant is investigated.

Keywords Probabilistic Safety Assessment · Loss of Offsite Power source
Loss of the Station Blackout

1 Introduction

Power system is a complex system whose main function is the generation, transmission and distribution of the electrical energy. Malfunction in a part of the system will lead to failure in the delivery of electrical energy to load and sometimes to a complete shutdown of the power system. Thus, the investigation of the power system reliability is essential in order to prevent frequent shutdowns. According to the increasing share of the energy production by the nuclear power plants and the potential environmental and human hazards, the assurance of their safety is of special importance [1].

The electrical feed of nuclear power plants is composed of two internal and external parts. The feed by the external and internal power sources (emergency diesel generators) are respectively considered as the main and auxiliary parts [2].

M. Ahmadnia (✉)
Department of Electrical Engineering, Hakim Sabzevari University, Sabzevar, Iran
e-mail: m.ahmadnia@hsu.ac.ir

F. Kiomarsi
Department of Electrical Engineering, Azad University, Neyshabour, Iran

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_33

The loss of external power source occurs when the power transmission is lost completely are partly. Availability to the alternate current (AC) is essential for the execution of safety operations in the nuclear power plants [3].

If the external power source connected to the nuclear power plant cannot supply the nuclear power plant considered power for any reason, then the internal electrical power supply (emergency diesel generator) enters the circuit and takes the responsibility of the electrical energy supply in the nuclear power plant. The Loss of Offsite Power supply (LOOP) as the starter event and consequently the probability of internal Station Blackout (SBO) are among events that are considered in the Probabilistic Safety Assessment (PSA) and have significant influence on the melting frequency of the reactor core. Fukushima accident as SBO is an alarm for other countries to analyze the effects of such an accident. The frequency of this event occurrence in a specific nuclear power plant is dependent on one hand on several factors such as the structure and topology of the power grid connected to the power plant and on the other hand on the internal design of the power plant. Thus in order to increase the reliability and safety factor of the nuclear power plant, it is necessary to investigate the LOOP phenomena.

This article is organized in 6 sections. The first section was the introduction that was presented. In the second section, an overview of the probabilistic safety analysis, in the third section, an overview on the analysis Loss of Offsite Power supply (LOOP), in the fourth section, a brief description about the emergency diesel generators, in the fifth section, conclusion and in the sixth section the references are presented.

2 Probabilistic Safety Assessment (PSA)

The Probabilistic Safety Assessment (PSA) is an important means for the analysis of the nuclear power plant safety against events resulting from random components malfunction, human errors and also internal and external events and risks [4].

Such analysis can be implemented in processes and possible mechanisms between nuclear power plant systems, for available power plants with operational record and also in power plants that are still in design stage. PSA covers both design stages of nuclear power plant and the safety and control management of the power plant on its whole service lifetime and updates according to the changes that is called alive PSA [5].

PSA is classified in three levels:

1. PSA level 1:

In level one of PSA, the events that lead to damage to the reactor heart are investigated. Events such as internal malfunctions, disturbances and errors, fires, floods, severe environmental conditions, Loss of Offsite Power supply (LOOP) and events started by human are investigated in the first level of PSA [6].

2. PSA level 2:

In the second level of PSA, the events that are excluded in level 1 of PSA and cause damage to the reactor core, are investigated and some methods are determined that prevent the dispersal of radioactive material to the environment [7].

3. PSA level 3:

The third stage of PSA starts from the results of level 2. In this level, the general health and other social consequences are calculated. Such as the pollution of earth, food and etc. that is due to the dispersal of radiocolonoid material into the environment [8].

3 Loss of Offsite Power Source in the Nuclear Reactor

The availability of alternate current is essential for safe operations and accident recovery in nuclear power plants, commonly the alternate current of nuclear power plants is supplied from external sources and by national network. LOOP event can have negative and important effect on the power plant capability to achieve and maintain the accident recovery [4].

Investigating the papers, we can find the extreme importance of LOOP event. LOOP event has been analyzed in different NRC reports in years (1996-1988-1988b-2003) [9–12]. In 2009, a research center for the supervision and inspection of the operational conditions of the electrical networks connected to the nuclear power plant has been established by European nuclear organization to improve the safety level of the nuclear power plants [13].

According to the presented statistics, LOOP events are recorded 228 times in the IRSN SAPIDE database, 190 times in GRS VERA database and 120 times in LER database and 52 times in IRS database [14].

LOOP event is composed of 8 sections that are described in the following [15]:

1. Power plant condition:

The events are classified based on the operational condition of the power plant and are as follows:

- On Power
- Hot Shutdown
- Cold Shutdown.

In Fig. 1 the LOOP events for IRSN SAPIDE database are recorded.

As can been seen in Fig. 1, 64% of the LOOP events happens in the On Power condition.

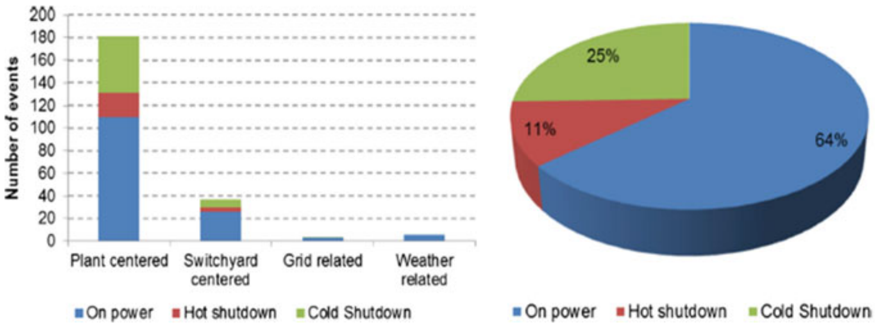


Fig. 1 Power plant condition at LOOP event for IRSN SAPIDE database according to Ref. [15]

2. Event conditions:

The event conditions are classified based on the power plant conditions at the event time:

- Normal Operation
- Shut down and Start up
- Corrective Maintenance
- Maintenance
- Fault Finding
- Inspection Functional Tests
- Moment Finding (Modification)
- Others.

Figure 2 shows the event conditions as a result of LOOP event for IRSN SAPIDE database.

As can be observed in Fig. 2 more that 36% of the events in the IRSN SAPIDE database are related to the normal conditions. Functional tests with 24% and maintenance with 14% are in the next ranks of a nuclear power plant event conditions.

3. Event type:

In this section, the events are classified based on the loss of electrical power.

- Partial loss of the external power
- Total loss of the external power with successful startup of the emergency diesel generators
- Total loss of the electrical energy with failure in the diesel generators start up
- Physical loss of the electrical bus bars
- Loss of power by emergency bus bar
- Loss of power by auxiliary bus bar.

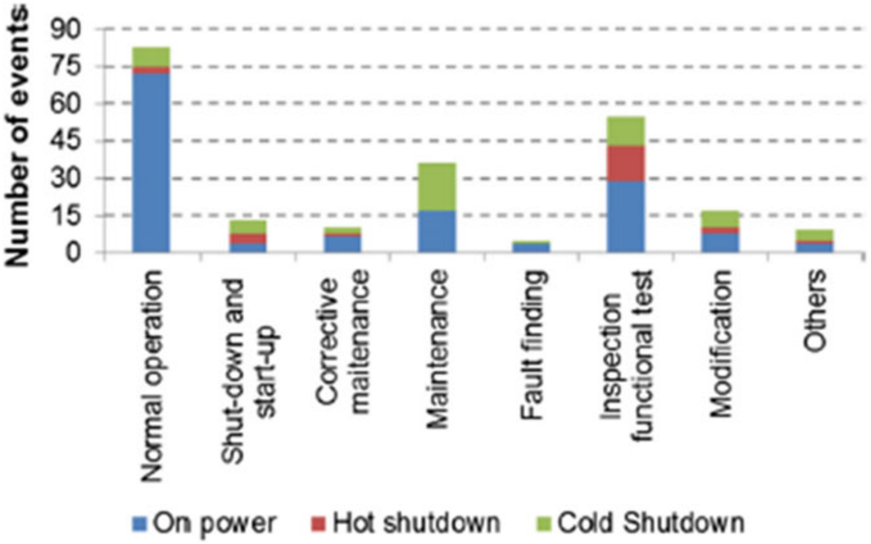


Fig. 2 Power plant event conditions at LOOP accident for IRSN SAPIDE database according to Ref. [15]

Table 1 Event type in LOOP accident for IRSN SAPIDE database according to Ref. [15]

	On Power	Hot Shutdown	Cold Shutdown
Partial loss of external	114	12	23
Total loss of external power	12	3	17
Loss of power supply	1	0	1
Physical loss of electrical bus bar	1	0	3
Loss of power to emergency bus bar	7	7	7
Loss of power auxiliary bus bar	10	3	7

Table 1 determines the event type for IRSN SAPIDE.

As can be observed, the biggest event type recorded for LOOP accident is related to the partial loss of the external power.

4. Type of damaged equipment:

Events are classified based on the equipment types that are subject to failure:

- Main and auxiliary connections (Main or Second Line)
- Switchyard and Breaker
- Transformer
- Emergency Generators
- Bus bar
- Inverter

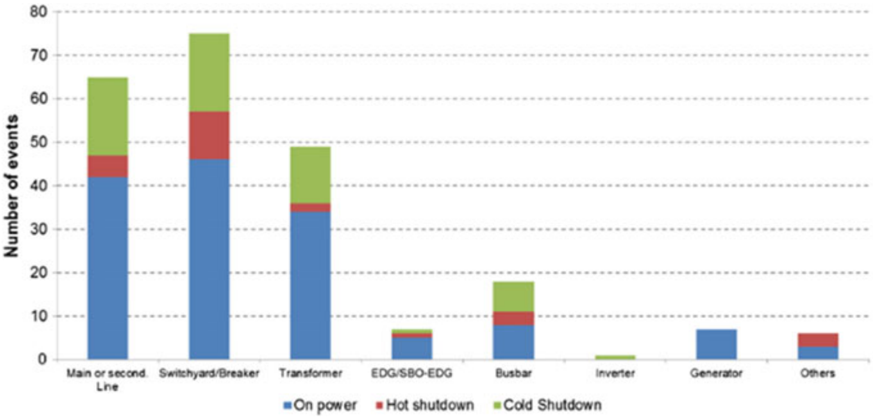


Fig. 3 Type of damaged equipment in at LOOP accident for IRSN SAPIDE database according to Ref. [15]

- Generator
- Others.

Figure 3 shows the failed equipment for IRSN SAPIDE database.

According to Fig. 3, most equipment failures in IRSN SAPIDE database are reported for the failure of transformers and breakers. Main and auxiliary lines and transformers are in the next ranks of failed equipment.

5. Direct cause of events:

The accidents are classified in three groups based on the place of accident.

- Electrical Grid Deficiency
- Switchyard Deficiency
- Plant Related Event.

Each of the above groups is divided into several subgroups such as: mechanical, electrical, instrumentation and control, environmental factors, human factors, unknown factors and etc. [15].

Figure 4 shows the direct causes of events based on the IRSN SAPIDE database.

According to Fig. 4, the most frequent direct cause of events as a result of LOOP accident that is recorded in IRSN SAPIDE database is the plant related event that is caused by human factors, instrumentation and control and electrical section.

6. Main cause or root of the event:

According to the direct causes of the event, the main causes are classified in the following groups:

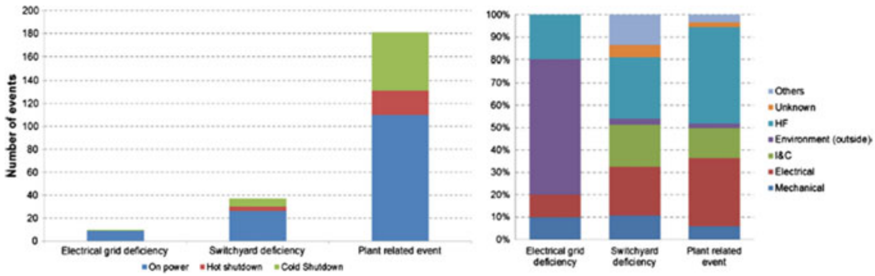


Fig. 4 Direct cause of LOOP accident for IRSN SAPIDE database according to Ref. [15]

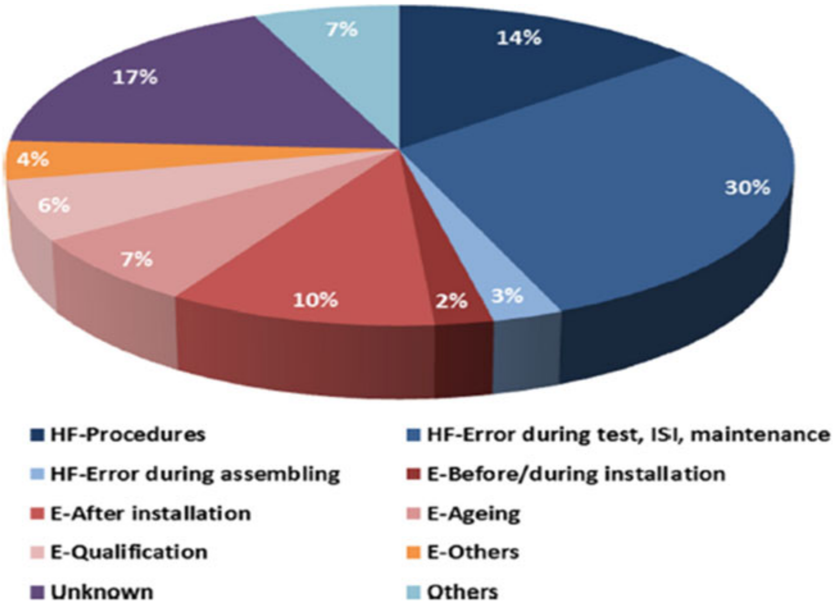


Fig. 5 Main causes of LOOP accident for IRSN SAPIDE database according to Ref. [15]

- Human performance
- Equipment performance
- Others.

Figure 5 shows the main causes of events for IRSN SAPIDE database.

Figure 5 represents that more than half of the main causes of the LOOP event in the IRSN SAPIDE database are due to the wrong performance of humans in the use of equipment and malfunctions removal.

7. Consequences of the event:

- Non-Compliance with Operational Technical Specifications
- House Load Operation
- (Offsite Line Switching/External System Connection Switching)
- Starting EDG Without Connecting
- Starting and Connecting EDG
- Starting SBO EDG
- Reactor Trip
- Material Degradation
- Others.

Figure 6 shows the LOOP event consequences based on IRSN SAPIDE database on the equipment.

As can be observed in Fig. 6, most LOOP accident events lead to reactor damage and in the next level, it affects the external lines/switching and external system connection.

8. Time period:

- Less than 2 min
- More than 2 min
- No data.

Figure 7 shows the time period of LOOP accident occurrence based on information of IRSN SAPIDE database.

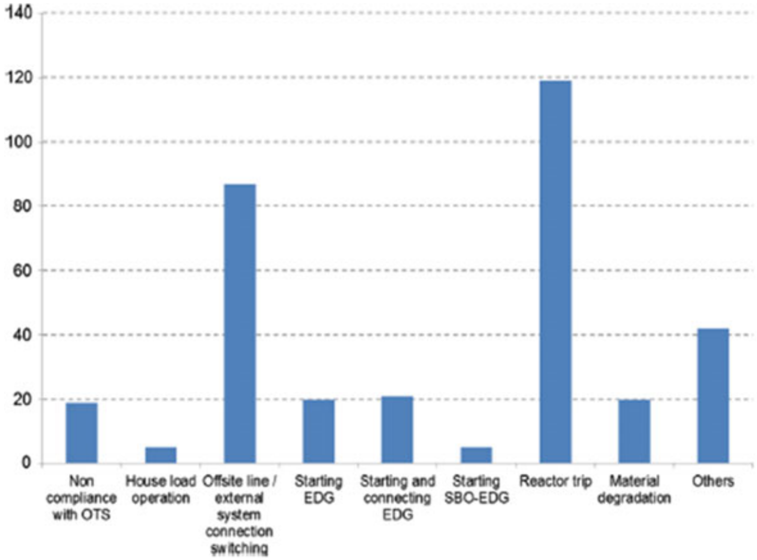


Fig. 6 Consequences of LOOP accident for IRSN SAPIDE database according to Ref. [15]

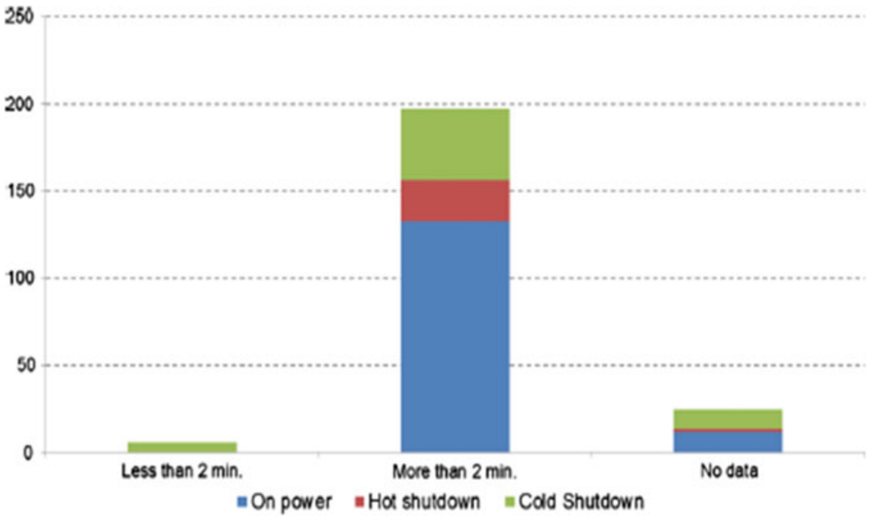


Fig. 7 Time period of LOOP accident for IRSN SAPIDE database according to Ref. [15]

Figure 7 represents that most of LOOP events are recorded for a time period of more than 2 min.

According to the reports presented in [NUREG 6890], the estimation of occurrence frequency of the external power source loss for nuclear power plants by probabilistic safety analysis is essential. Regarding the oldness of nuclear power plants and their high amount in developed countries, most of the presented methods are mainly based on the retrospective methods and not only this is not wrong but also helps to better estimate the occurrence frequency of the event, but for Iran in which few data are available and the life of nuclear power plants is not too long, use of these methods is not suitable and the event occurrence frequency should be estimated always using the failure data of the national grid connected to the nuclear power plant and structure and topology of the power system and creating of event and error trees.

4 Emergency Diesel Generators

One of the most important accidents that may lead to damage to the reactor core, in the case of malfunction in the supporting systems (emergency diesel generators) is the LOOP event (investigated in Sect. 3), that in case of accident expansion along with power plant emergency diesel generators failure can lead to complete AC power outage accident in the power plant that is called SBO accident. Thus upon occurrence of such an accident, the normal operation of the power plant emergency power system that is responsible for the energy provision of many reactor safety systems, is absolutely necessary and is considered as the most important index in the occurrence of such an event. The most important section of this system is composed of emergency diesel generators that are responsible for the provision of AC voltage of the emergency bus bars.

In order to assure the performance of diesel generator on demand, the procedure and quality of inspection and maintenance is of high importance in this system. Since the diesel generators have specific operating conditions and will face problems due to ageing, thus it is important to consider some provisions in their design [16].

5 Conclusion

Regarding to the increasing share of the energy production by the nuclear power plants, the probabilistic safety analysis should be performed continuously in the nuclear power plant to improve reliability and decrease the risks. Loss of external power source of the power plant is an event that enters the probabilistic safety analysis and if it is not recognized and the diesel generator cannot operates on time, SBO phenomenon will happen. Thus, according to the performed investigations, the estimation of occurrence frequency of the external power source loss in the nuclear power plant is essential. This is different for Iran and the difference is that regarding the short lifetime of nuclear power plants and their few amount, the use of retrospective methods is not suitable and it is always necessary to investigate the failures of the national grid connected to the power plant, structure and topology of the power grid connected to the power plant and then create the related error and event tree in order to predict the event occurrence frequency correctly.

References

1. Farahani AZ, Yousefpour F, Rajahsay M, Hosseini SM (1394) Thermohydraulic analysis of the MELCOR Model of the First Circuit of IR-360 Power Plant. In: 23rd Iran nuclear conference, Yazd University
2. IAEA (2012) Electric grid reliability and interface with nuclear power plants, IAEA Safety Standards Series, no. NG-T-3.8. International Atomic Energy Agency, Vienna, p 78
3. Volkanovski A, Čepin M, Mavko B (2007) An application of the fault tree analysis for the power system reliability estimation. In: Proceedings of the international conference nuclear energy for New Europe, Portorož, Slovenia, Sept 10–13
4. NUREG/CR-6890 (2012) Reevaluation of station blackout risk at nuclear power plants analysis of loss of offsite power events: 1986–2004
5. International Atomic Energy Agency (IAEA) Vienna (1992) Probabilistic safety assessment INSAG-6. A report by the International Nuclear Safety Advisory Group
6. Stuk Helsinki (2003) Probabilistic safety analysis in safety (Management of nuclear power plants). ISBN 951-712-788-X (PDF)

7. Organization for Economic Co-operation and Development Paris—59015—level 2 PSA methodologies and severe accident management Prepared by the (CNRA Working Group on Inspection Practices) WGIP
8. Canadian Nuclear Safety Commission April 2005—Probabilistic Safety Assessment (PSA) for Nuclear Power Plants S- 294. ISBN 0-662-40139-5
9. NRC (1996) Rates of initiating events at U.S. nuclear power plants: 1987–1995. U.S. NRC, Washington
10. NRC (1988) Evaluation of station blackout accidents at nuclear power plants: technical findings related to unresolved safety issue A-44: Final report. U.S. Nuclear Regulatory Commission, Washington
11. NRC (1998b) Evaluation of loss of offsite power events at nuclear power plants, 1980–1996 [microform]/prepared by Atwood CL et al. U.S. Nuclear Regulatory Commission, Washington, DC
12. NRC (2003) Operating Experience Assessment—effects of grid events on nuclear power plant performance, U.S. NRC
13. Ballesteros A, Peinador M, Heitsch M (2015) EU Clearinghouse Activities on Operating Experience Feedback. BgNS Trans 20:93–95 (Bulgarian Nuclear Society, Sozopol, Bulgaria)
14. Volkanovski A, Avila AB, Veira MP, Kančev D, Maqua M, Stephan J-L (2016) Analysis of loss of offsite power events reported in nuclear power plants. Nucl Eng Des 307:234–248
15. NRC (2005) Reevaluation of station blackout risk at nuclear power plants, Washington
16. Rastaesh S (1392) Dynamic reliability assessment of diesel generators of Bushehr Nuclear Power Plant. Master's Degree, Azad University, Science Research Branch

Optimization of the Fuel Consumption for the Vehicle by Increasing the Efficiency of the Electrical Transmission System

Mohsen Ahmadnia

Abstract The author studies the method of improving the fuel efficiency parameters of automobile power installation with electromechanical transmission. He also carries out the calculating research of the power installation balance during the movement of an automobile according to the NEDC—new European drive cycle. The fuel efficiency of automotive internal combustion engines remain the main indicators of his work. Among these indices, nominal and effective efficiency of the engine at a particular operating mode (often nominal), the corresponding specific fuel consumption, hourly fuel consumption in this mode, the fuel consumption on 100 km of run of the car or operational fuel consumption. Among these indicators is the most informative, operational fuel consumption, since it reflects the fuel efficiency of the power train of a car taking into account the distribution of the modes of operation of the installation in real operation conditions.

Keywords Vehicle · Gearbox · European driving cycle (NEDC)

1 Introduction

Electromechanical transmission in the vehicle allows optimizing the operation of the reciprocating internal combustion engine by the rotation frequency, ensuring a reduction in the fuel mileage, by eliminating the rigid communication in the transmission between the engine and the propulsion mechanism inherent in the mechanical transmission change gearbox.

The inclusion of the energy storage element in the electromechanical transmission allows, in certain limits, to optimize the work of the engine on load, having as an objective function the minimization of the operational (road) fuel consumption of the car.

M. Ahmadnia (✉)
Department of Electrical Engineering, Hakim Sabzevari University,
Sabzevar, Iran
e-mail: m.ahmadnia@hsu.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_34

To carry out parametric analysis of the scheme of electromechanical transmission with battery, the program was updated to calculate the fuel mileage during vehicle movement in accordance with various standardized driving cycles [1]. The program is a calculation and experimental research tool, since all calculations in it are based on experimental data on a specific ICE model. The investigated engine in the program is represented by its experimental universal characteristic for fuel consumption and external speed characteristic. We investigated the diesel Volkswagen TDI model ALH with a working volume of iV_h = 1.9 L, the rated power is N_e = 66 kW at 3750 min^{-1} [2]. It should be noted that in the program there are no fundamental limitations on the study of the car engine's performance when driving on an arbitrary driving cycle, because data on the cycle is entered into the program as an array of current cycle speed in increments of one second. The most popular of them driving cycles NEDC, FTP-75 and JC08 are implemented in the modified version of the program used in this study.

The European driving cycle NEDC (New European Driving Cycle), adopted for use on January 1, 2000 and consisting of two phases of the movement: from the Urban Driving Cycle, with the vehicle speed limitation of 50 km/h and the EUDC, Extra Urban Driving Cycle speed of 120 km/h. The Russian Federation adheres to this particular test cycle. The federal test procedure USFTP-75 (Federal Test Procedure), consists of three phases of the movement, with the initial and final phases identical, and after the second phase, the engine stalls for 10 min. Compared to the European cycle, the American FTP-75 is more dynamic, and the number of vehicle starts is 22. The Japanese driving cycle JC08 adopted in October 2011 is close to the European, for a longer time of stopping the car, simulating waiting at traffic lights and traffic jams in megacities.

Different cycles are not comparable, i.e. it is not possible to introduce a correction factor, for example, in fuel consumption during the transition from one cycle to another. So Volkswagen experts tested a 1.4-L gasoline engine using the NEDC and FTP-75 methods and determined that in the certification documents for the USA fuel consumption is 10–16% lower than the European data [3]. However, for a motor with a different distribution of specific effective fuel consumption in the field of the operating modes "load speed of the crankshaft" by the universal characteristic, the difference will be different. This is due to the fact that the operating time in specific driving points of the motor characteristics is not the same for different cycles.

2 Methodology

Table 1 shows the calculation of the total work A cycle performed in the driving cycle, the fuel consumption of the engine of the car cycle in the cycle, taking into account the experimental characteristic g_e = f (n, P_e) with a 5-speed manual gearbox (Gearbox) and fuel consumption for cycle a cycle of mines, in the event that the engine has performed the required work while in the mode of maximum economy

Table 1 Possible reduction of fuel consumption due to optimization of the ICE 1

Parameter	Driving cycle		
	NEDC	*JC*08	*FTP*-75
Total cycle work A_{cycle}, kJ	4787	2494	6422
Fuel consumption in cycle G_{cycle}, g	566.3	380.2	763.2
Fuel consumption during operation $g_{e\ min}$ = 197 g/(kWh), g	262.0	136.5	351.4
Reducing the flow rate due to operation in the mode the maximum profitability, %	53.7	64.1	54.0

$$A_{\text{cycle}} = \Delta t \cdot \sum_{i=1}^{\text{K}} N_{e\ \text{required}\ i} \tag{1}$$

$$G_{\text{cycle}} = \sum_{i=1}^{\text{K}} g_{e\,i}(n, P_e) \cdot N_{e\,i} \cdot t_i \tag{2}$$

$$G_{\text{cycle min}} = g_{e\,min} \cdot A_{cycle}/3600 \tag{3}$$

where

Δt is the calculation step in the program, in our case 0.1 s;

K is the number of calculation steps;

$N_{e\,\text{required}\,i}$ the power required from the engine at the *i*-м calculated step, calculated by the power balance of the car;

t_{i} the time of the engine operation in the loading-speed zone (P_e, n).

According to the experimental data, the minimum effective fuel consumption of the diesel engine studied was $g_{e\ min}$ = 197 g/(kWh).

It can be seen from the data in Table 1 that the engine is used least efficiently in the Japanese cycle, for which the low engine load is characterized by an average car speed of only 24.4 km/h, compared to 32.8 km/h in the European cycle and 34.2 km/h in the American cycle. Also in the Japanese cycle, the maximum engine running time is at idle. In this case, even in the European cycle, it is theoretically possible to reduce by more than half the track fuel consumption, only by optimizing the engine operation modes.

As it was shown in our works, in order to optimize the operational fuel consumption of a piston engine of a vehicle with an electromechanical transmission, it is necessary to determine its two operating frequencies n_1 and n_2. The first frequency n_1 corresponds to the region of the minimum specific fuel consumption for the universal characteristic of the engine. At the second n_2, the increased engine speed, the engine passes if it's available power (by the external speed characteristic) at a frequency n_1, less than the required one, for driving in accordance with the driving cycle.

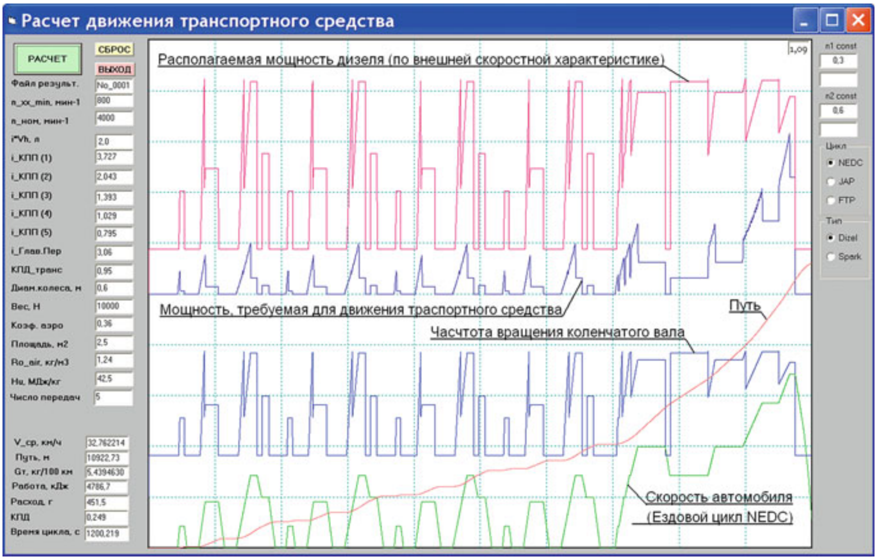
The engine power control can include its short-term stop and then start. So, with prolonged braking and the subsequent stopping of the vehicle, the internal combustion engine is muffled. This mode is known as the “Start-stop”. The subsequent beginning of the movement is carried out on the electric drive, with the selection of energy from the battery.

Figure 1 shows the user interface of the calculation program when calculating the vehicle’s movement from a mechanical 5-speed transmission.

Preliminary calculations of the parameters of the car with a mechanical gearbox were performed. At the same time, fuel consumption was obtained for the test cycle $G_{cycle} = 566.6$ g.

For carrying out of numerical researches of work of the engine with an electromechanical transmission in the program the following algorithm of work ICE is realized. The internal combustion engine when the vehicle is stopped (vehicle speed $V_a = 0$ and vehicle acceleration $J_a = 0$) and its deceleration ($V_a > 0$ and $J_a < 0$) operates at minimum stable idling speed $n_{xx\ min}$. When accelerating ($V_a > 0$ and $J_a > 0$) and moving with constant speed ($V_a > 0$ and $J_a = 0$), the engine runs at a constant speed of rotation n_1 provided that its available power N_e is greater than the required power N_e for driving the vehicle in the current driving point of the driving cycle. If the N_e condition is not satisfied, then the motor moves to another high-speed operating mode with a higher speed. Naturally, the available power increases and the condition N_e cannot be satisfied.

On the form of the program, you can set the share of the nominal speed from the nominal speed for n_1 and n_2 (in Fig. 2, the fraction is $k_1 = 0.3$ for the speed of n_1 and $k_2 = 0.6$ for n_2 for speed n_2). The program for controlling numerical values calculates the rotation speed in $\min^{-1}$ [4–10].


Расчет движения транспортного средства

Файл результ.	No_0001
n_xx_min, мин-1	800
n_ном, мин-1	4000
i*Vh, л	2,0
i_КПП (1)	3,727
i_КПП (2)	2,043
i_КПП (3)	1,393
i_КПП (4)	1,029
i_КПП (5)	0,795
i_Глав.Пер	3,06
КПД_транс	0,95
Диам.колеса, м	0,6
Вес, Н	10000
Коэф. аэро	0,36
Площадь, м2	2,5
Ro_air, кг/м3	1,24
Hu, МДж/кг	42,5
Число передач	5
V_ср, км/ч	32,762214
Путь, м	10922,73
Gт, кг/100 км	5,4394630
Работа, кДж	4786,7
Расход, г	451,5
КПД	0,249
Время цикла, с	1200,219
n1 const	0,3
n2 const	0,6

Цикл
● NEDC
○ JAP
○ FTP

Тип
● Dizel
○ Spark

1,09
Располагаемая мощность дизеля (по внешней скоростной характеристике)
Мощность, требуемая для движения траспортного средства
Путь
Частота вращения коленчатого вала
Скорость автомобиля
(Ездовой цикл NEDC)

Fig. 1 Type of the user interface of the program

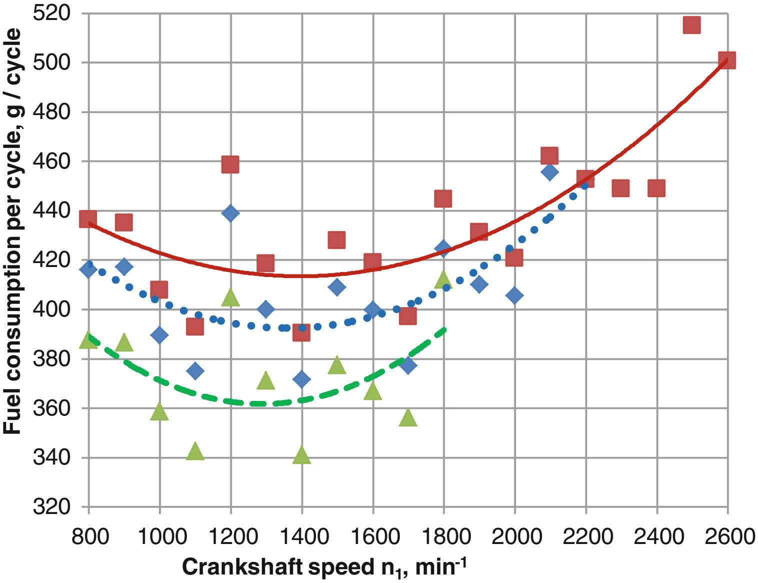


Fig. 2 Fuel consumption by the engine of the car when performing 2600 min^{-1} (k_2 = 0, 65); ——driving cycle NEDC for the values of n_2: 2200 min^{-1} (k_2 = 0.55); ········ ———— 1800 min^{-1} (k_2 = 0.45)

In the area of minimum fuel consumption (Fig. 2), the choice of an increased frequency of 2200 min^{-1}, instead of 2600 min^{-1}, reduces the operating fuel consumption by 5%, and the choice of frequency 1800 min^{-1} reduces consumption by 12.5%.

However, as already noted, at low values of n_2, an increase in the frequency may not give the necessary increase in power. Figure 3 shows the values of the useful work done in the driving cycle, calculated as

$$A_{cycle} = \Delta t \cdot \sum_{i=1}^{\kappa} N_{e\text{Required}\,i} \tag{4}$$

where Δt is the calculation step, in our case 0.1 s, and K is the number of calculation steps.

From the graphs shown in Fig. 3 it is seen that for the frequency n_2 = 1800 min^{-1}, the choice of the speed n_1 of less than 1200 … 1300 min^{-1} leads to a decrease in the cycle operation, which can be explained by the inability to provide the required power in some sections of the driving cycle. The most critical, in this sense, NEDC cycle is the most energy-intensive site for accelerating the car from a speed of 100 km/h to a speed of 120 km/h. The algorithm of the program is such that in each calculation point of the program the conformity of the available engine power is checked (the maximum power developed by the engine at the crankshaft

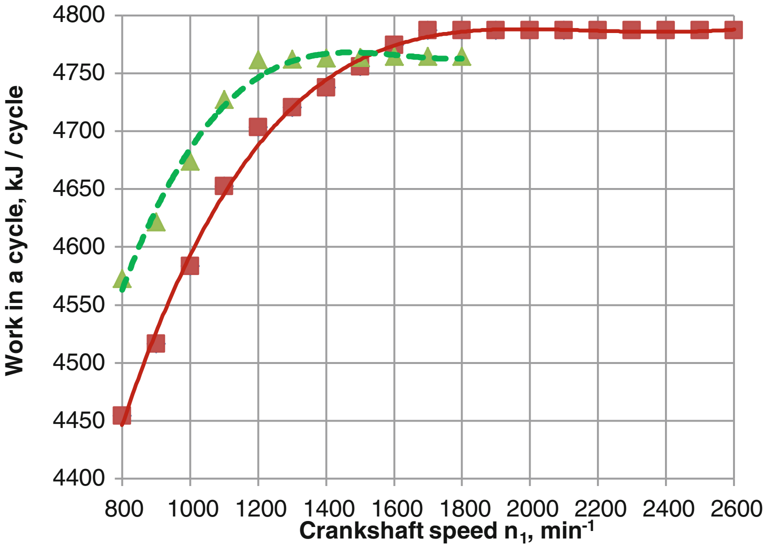


Fig. 3 Work performed by a motor of a vehicle with an electromechanical transmission when driving conditions are specified in accordance with the NEDC cycle when an increased speed n_2 is selected equal to: Δ 1800 min^{-1}; □ 2200 min^{-1} and 2600 min^{-1}

speed specified by the mode) and the required power for movement in accordance with the cycle. If the $N_{e\ required} \geq N_{e\ available}$ condition is not met, the acceleration of the vehicle during acceleration decreases, by recalculating the value of the available engine power, and the total operation of the cycle decreases.

Consequently, as can be seen from Fig. 3, it is impossible to reduce the frequency n_1 by less than 1200 … 1300 min^{-1} at n_2 equal to 1800 min^{-1} and less than 1700 min^{-1} at n_2, = 2200 … 2600 min^{-1}, because otherwise the driving conditions for the driving cycle are not met.

3 Conclusion

Returning to the graph in Fig. 2, we can conclude that, according to the conditions formulated above, the minimum path flow in the cycle will be the values given in Table 2.

Table 2 Fuel consumption for driving on a driving cycle NEDC

Number of the mode	n_1, min^{-1}	n_2, min^{-1}	G_T, g/cycle
1	1300	1800	365
2	1700	2200	400
3	1700	2600	420

Thus, it is necessary to stop on the adjustment number 1, which ensures minimum fuel consumption. Compared with the mechanical 5-speed transmission, this adjustment results in a 19% reduction in fuel consumption.

It should be noted that the calculation was carried out at an electrical transmission efficiency of 0.95. It should be noted that such a transmission involves the use of modern high-efficiency electric generator and electric motor, the communication between which is carried out by a control unit that uses digital frequency and drive power control, which reduces losses in the electrical part of the transmission. At lower values of transmission efficiency, the vehicle's road fuel consumption proportionately increases.

It should be noted that the electro-mechanical transmission has the prospect of further reducing the fuel consumption of the engine. With the use of an accumulating element (electric accumulator or super capacitor), in addition to regenerating the braking energy, it becomes possible to optimize the operation of the internal combustion engine not only by the speed of rotation, but also by the load.

References

1. Boguslawski L (2004) Influence of pressure fluctuations distribution on local heat transfer on flat surface impinged by turbulent free jet. In: Proceedings of international thermal science seminar II, Bled, 13–16 June 2004
2. Muhs D et al (2003) Roloff/Matek mechanical parts, 16th edn. Vieweg Verlag, Wiesbaden, 791 p (In German). ISBN 3-528-07028-5
3. ISO/DIS 16000-6.2 (2002) Indoor Air—Part 6: Determination of volatile organic compounds in indoor and chamber air by active sampling on TENAX TA sorbent, thermal desorption and gas chromatography using MSD/FID. Geneva, International Organization for Standardization
4. Kulchitsky AR (2000) Toxicity of automobile and tractor engines/AR Kulchitsky. Publishing house Vladimir State University, Vladimir, 256 p
5. Bosch (2004) Control systems for diesel engines: trans. With him. - M.: Behind the wheel, 480 p
6. Grekhov LV, Ivaschenko NA, Markov VA (2005) Fuel supply and control systems for diesel engines. Legion-Avtodata Publishing House, Moscow, p 344
7. Kravets VN, Gorynin EV (2002) Legislative and consumer requirements for cars. NSTU Publishing House, Nizhny Novgorod, p 400
8. Gusakov SV (2008) Hybrid propulsion system based ICE. PFUR Publishing House, Moscow, 207
9. Turevsky IS (2005) The theory of the vehicle. Higher School, Moscow, p 121
10. Bosch (2013) Automotive engineering. Traditional and hybrid drives. Publishing House Behind the wheel, Moscow, 332 p

Improve the Reliability and Increased Lifetime of Comb Drive Structure in RF MEMS Switch

Faraz Delijani and Azim Fard

Abstract This paper presents a novel electrostatic comb-drive RF MEMS switches for the purpose of increasing reliability of micro-electromechanical switches, some new designs analyzed for elimination of adhesion, friction and dielectric charging in non-contact switches, targeted to decrease electrostatic force and applied voltage in the comb driven structure. The reason of this choice is simple manufacturing process (1 or 3 masking) compared to the other MEMS structures, ability to build with more material, linear functioning without static spring fingers, wide applications in addition to switches, such as Resonators, micro-filters, gyroscopes, accelerometers etc.

Keywords RF MEMS · Reliability · Life time · MEMS switch
Comb drive

1 Introduction

RF switches are one of the most important components in wireless communication systems and radars [1]. Traditional solid-state switches are used to realize RF signal switching, such as GaAs PIN diode and transistor switches [2]. With the development of semiconductor manufacturing technology, MEMS switch is invented and quickly becomes a hotspot of study [3–7]. The comparison between these semiconductor and MEMS switches shows that the latter one is more suitable for high frequency signal processing, due to their excellent performance, i.e. high isolation, low insertion, low power consumption, no or negligible distortion.

F. Delijani (✉)
Department of Electrical Engineering, Islamic Azad University,
Arak Branch, Arak, Iran
e-mail: delijanifaraz@gmail.com

A. Fard
Communication Regulatory Authority, Tehran, Iran
e-mail: azimfard@cra.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_35

Despite better performance over other competing technologies such as PIN or FET switches, the commercialization of shunt capacitive RF MEMS switches is hindered by the reliability problem of RF MEMS switch. The causes of the switch failure are mainly the electrical failures such as dielectric breakdown and the mechanical failures such as capillary stiction, dielectric charging induced stiction, the buckling or broken off the bridge, and the self-actuation under high power condition [8]. There are two generalized types of switches: Ohmic and capacitive. Ohmic switches make direct contacts between two electrodes while capacitive switches form metal insulator metal contacts. Both types of switches have the ability to operate more than billon of cycle without any reliability issue. Ohmic switches more often fails by stiction, whereas capacitive switches often fails due to charging of their dielectric insulators [9]. The metal to metal contacts are always forming in metal contacting switches to achieve Ohmic contact between two electrodes. The capacitive switches have a thin dielectric film and an air gap between the two metallic contact surfaces. The air gap is electromechanically adjusted to achieve a capacitance change between the 'up' and 'down' state. The capacitance ratio of the downstate value to the upstate value is a key parameter for such a device; a high capacitance ratio is always desirable [10]. Among various reported reliability concerns for electrostatic capacitive MEMS switches, the dielectric charging and its resulting stiction is considered the main failure mechanism of these devices.

To rectify the problem of stiction, we have designed a noncontact- type MEMS switch. In this micro structure, the micro welding and stiction problems in the contact switches are eliminated. The proposed micro structure is designed with variable capacitance structure which does not allow direct contact or indirect contact [11]. Comb drive actuators consist of two interdigitated fingers structures, where one comb is fixed (Stator) and the other is connected to a compliant suspension (Rotor). Applying a voltage difference between the comb structures will result in a deflection of the movable comb structure by electrostatic forces. Electrostatic forces increase with decreasing gap spacing and an increasing number of comb fingers [12].

2 Problem Statement—Stiction Mechanism

Arguably stiction is one of the most important reliability challenges in contact MEMS switches. Stiction is the case in which two normally isolated surfaces that are in operational contact cannot be separated through normal operation. Capacitive switches often depend on electrostatic attraction of parallel plates. The relative motion of these plates is governed by two parameters, the "pull-in" voltage and the "pullout" voltage. By all accounts, the pull-in voltage is a good representation of restoring force. The difference between the pull-in and pull-out forces is a good representation of the adhesion force. For devices to operate properly, the restoring force should always be greater than the adhesion force [13].

A consensus has developed that stiction can be caused by electrostatic attraction between charges trapped in the dielectric and the moving electrode. If the charge density is uniform, then pull down and hold down voltage both shift by the same amount. If the charge is injected from the fixed electrode and penetrates some distance into its attached dielectric, then it increases the electric field at the interface between the dielectric and the movable electrode. Thus the voltages required to operate the switch become lower. If the charge injection is large enough that the hold down voltage crosses zero, then the switch does not open at zero bias; it is stuck. This is the simplest form of stiction in dielectric switches [14].

In our design approach, we use comb drive actuators switch which remove the true cause of stiction i.e. adhesion, friction and dielectric charging. This structure remains the submicron gap between two electrodes when the actuation voltage applied to the comb drive actuator. This structure removes the stiction problem from the MEMS switch.

3 Design and Mechanism of a Comb Drive Switch

Comb drive actuator is one of the most common electrostatic actuator used in MEMS applications. It uses both electrostatic energy from a DC voltage applied between the moving & fixed comb drive structures, and the mechanical restoring force provided by the spring structure [15]. Comb drive actuators have been used as resonators, electromechanical filters, optical shutters, micro grippers and voltmeters. These have also been used as the driving element in vibromotors and micromechanical gears [16]. It is desirable in comb drive to achieve large displacements at low actuation voltages. The well-known electrostatic micro-actuators include side drive silicon micro motor, wobble micro-motor, comb drive micro actuator and out of plane diaphragm micro actuator [17].

Comb Drive actuators consist of two interdigitated finger structures, where one comb is fixed and the other is connected to a compliant suspension. The driving voltage between the comb structures causes the displacement of the movable fingers towards the fixed fingers by an attractive electrostatic force. The position of the movable finger structure is controlled by a balance between the electrostatic force and the mechanical restoring force of the compliant suspension. Mechanical forces are generated through spring structures. So besides electrostatic forces, mechanical forces also play a very important role. Mechanical forces are directly depending upon the stiffness of the flexures. By changing these flexures, mechanical forces also changes. It is very important to find the flexure compatibility for large deflections at low actuation voltage. In the present work, different flexures of electrostatic comb drives are simulated for large displacements.

Typical Comb-Drive considered in this study is shown in Fig. 1. It consists of movable and fixed combs. The fixed comb is anchored to the base and movable plate is supported by a spring through a shuttle mass plate. Different types of spring designs have been applied in comb-drive actuators. Simulation for different 2D

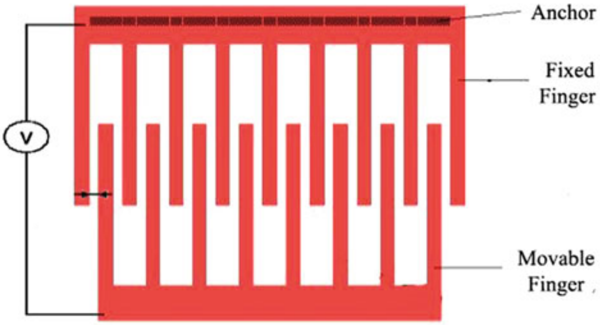


Fig. 1 Typical comb-drive

structures is done by using COMSOL 3.5a as it provides advanced methods of solving moving boundaries with FEM. The basic structure consists of 7 fixed and 6 movable combs attached to different flexure spring having 5 μm distance between the comb fingers.

The capacitance of comb drive is [18]:

$$C = \frac{2n\varepsilon_0 t(y_0 + y)}{g} + NC_p \tag{1}$$

where, t is the thickness of comb finger, y_0 is the initial overlap, y is the comb displacement and g is the gap spacing between movable and fixed combs. The lateral electrostatic force in y direction can be expressed as [18]:

$$F = \frac{1}{2}\frac{\partial C}{\partial y}V^2 = \frac{n\varepsilon_0 tV^2}{g} \tag{2}$$

where, V is the applied voltage between the movable and fixed combs. The displacement, y (along the y-axis), of the movable plate as a function of applied force is [18]:

$$y = \frac{n\varepsilon_0}{2Eg}\left(\frac{L}{W}\right)^3 V^2 \tag{3}$$

Increasing the beam stiffness will require large electrostatic force to cause the deflection and consequently require higher driving voltages. A general formula for calculating the stiffness is [18]:

$$K = 2Et\left(\frac{W}{L}\right)^3 \tag{4}$$

Table 1 Properties of Polysilicon

Dimensions of actuator	
Comb length	40 μm
Comb width	2 and 4 μm (Jagged)
Gap between moving comb and fixed combs	5 μm
Overlapping area	20 μm
Spring length	280 μm
Spring width	2 μm
Gap between spring legs	19 μm
Thickness of actuator	2 μm

where y is the displacement, covered by the movable comb finger from its initial overlap position in positive y-direction when an electrostatic field, F is applied on it and K is the mechanical stiffness of the flexure spring. The suspension spring (beam or cantilever) must be flexible enough in the direction of the actuation.

E is the Young Modulus for Polysilicon, *W* is the width and *L* is the length of cantilever. For choosing a suspension there are usually four main characteristics to watch: the spring constant, the compliance in the other directions (it needs to be low to keep the motion in the desired direction), the tolerance to internal stress (long beam may buckle during fabrication) and its linearity during large deformation.

In this paper we have chosen folded-beam flexure spring design. The design parameters of comb drive are shown as Table 1.

The basic material used for the design are Polysilicon, Single-Crystal Silicon and Polyimide. Tables 2, 3 and 4 provide the properties of these materials which are used for the design.

Table 2 Properties of Polysilicon

Parameters	Value
Young’s Modulus	$160e^{9}$ (Pa)
Poisson’s ratio	0.22
Density	2320 (kb/m^{3})
Thermal expansion	$2.6e^{-9}$ (1/k)
Relative permittivity	4.5

Table 3 Properties Single-Crystal Silicon

Parameters	Value
Young’s Modulus	$150e^{9}$ (Pa)
Poisson’s ratio	0.17
Density	2320 (kb/m^{3})
Thermal expansion	$2.7e^{-9}$ (1/k)
Relative permittivity	11.68

Table 4 Properties Polyimide

Parameters	Value
Young's Modulus	$3.1e^{9}$ (Pa)
Poisson's ratio	0.35
Density	1300 (kb/m^{3})
Thermal expansion	25×10^{-5} (1/k)
Relative permittivity	3.5

4 Results

4.1 *Polysilicon*

In Fig. 2 operating principle of contact-less comb drive MEMS switch, showed and the displacement highlighted compared with off-state. Switch enabled by applying of 100 V DC voltage.

4.2 *Single-Crystal Silicon*

In Fig. 3 operating principle of contact-less comb drive MEMS switch, showed and the displacement highlighted compared with off-state. Switch enabled by applying of 100 V DC voltages.

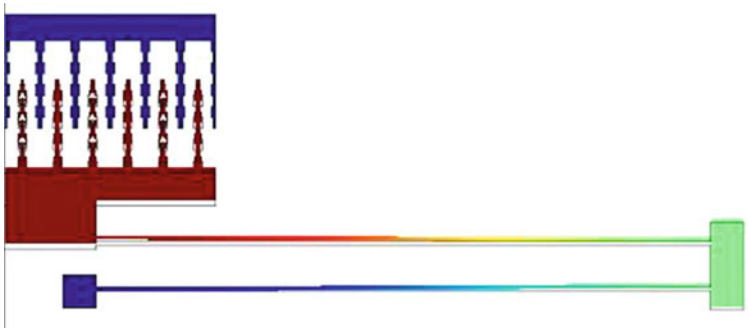


Fig. 2 Switch in On state

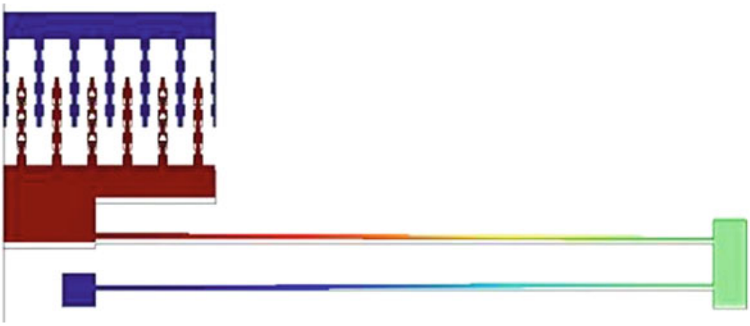


Fig. 3 Switch in On state

4.3 Polyimide

In Fig. 4 operating principle of contact-less comb drive MEMS switch, showed and the displacement highlighted compared with off-state. Switch enabled by applying of 100 V DC voltages (Figs. 5, 6, 7) and (Table 5).

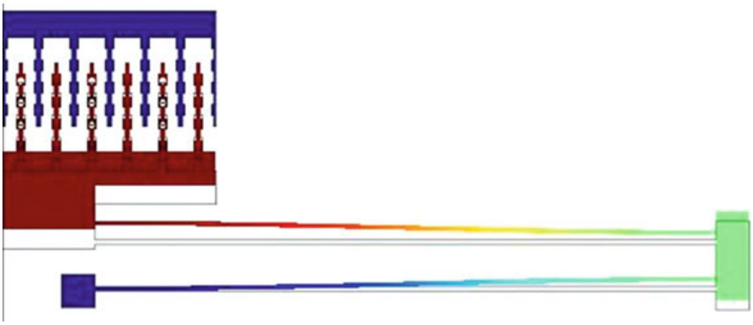


Fig. 4 Switch in On state

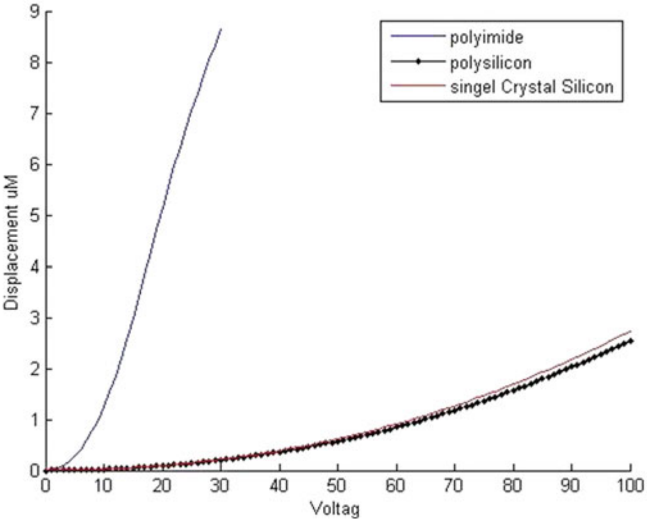
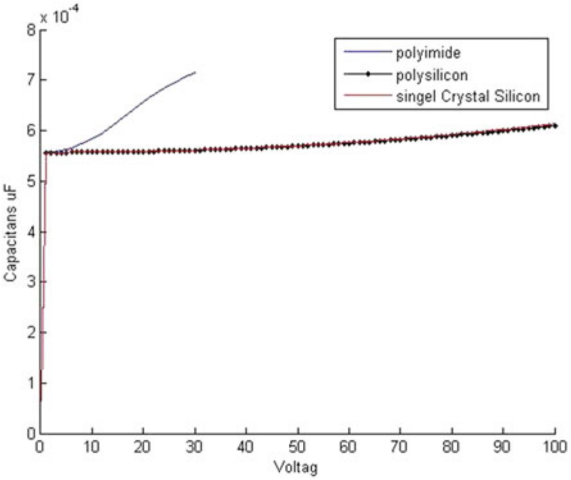


Fig. 5 Compare the displacement

Fig. 6 Compare the capacitance



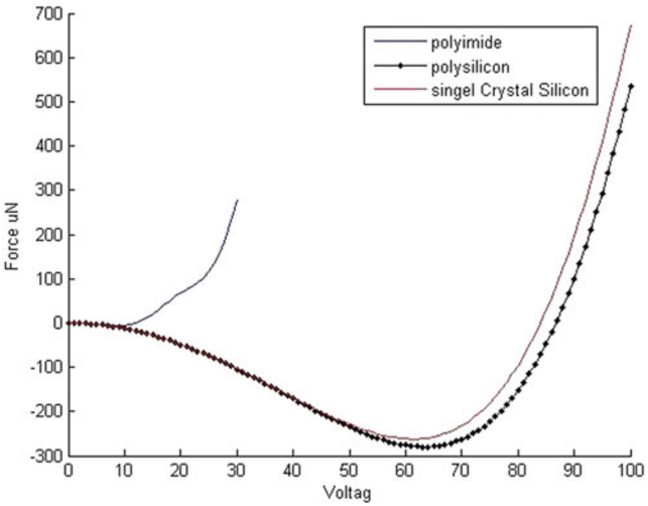


Fig. 7 Compare the electrostatic force

Table 5 Output results

Material in structures	Polysilicon	Single-Crystal Silicon	Polyimide
Max. voltage (V)	100	100	30
Displacement (μm)	2.5433435	2.7239835	8.636105
Capacitance (μF)	609	613	713
Force (μN)	534	672	276

5 Conclusion

In this paper, a new MEMS switch with comb-driven structure proposed and simulated to reduce the necessary electrostatic force for switch functioning, which in turn enables reduction of DC driving voltage. By the reduction of driving voltage the proposed approach removes stiction and extends the life time of RF MEMS switch. Moreover, and the switch provides a high isolation between input and output.

Modeling and simulation were carried out with COMSOL software employing three different construction materials with a joint spring structure. The final results showed a reduction to more than one third in the electrostatic force and the applied voltage is decreased considerably. Therefore the life time expected to be extended to 3 times of initial designs and accordingly, reliability of these switches would be also improved through the reduction of failure statistics to one third.

References

1. Nguyen CT-C, Katehi LPB, Rebeiz GM (1998) Micro machined devices for wireless communications (invited). Proc IEEE 86:1756–1768
2. Rohde UL, Newkirk DP (2000) RF/Microwave circuit design for wireless applications. Wiley, New York
3. Zavracky PM, McGruer NE, Morrison RH, Potter D (1999) Micro switches and micro relays with a view toward microwave applications. Int J RF Microw Comput Aided Eng 9:338–347
4. Hyman D, Lam J, Warneke B, Schmitz A, Hsu TY, Brown J, Schaffner J, Walston A, Loo RY, Mehregany M, Lee J (1999) Surfacemicromachined RF MEMs switches on GaAs substrates. Int J RF Microw Comput Aided Eng 9:338–347
5. Gong S, Shen H, Bark NS (2009) A cryogenic broadband DC contact RF MEMS switch. In: IEEE MTT-S International Microwave Symposium Digest, Boston, 7–12 June 2009, pp 1225–1228
6. Goldsmith CL, Yao Z, Eshelman S, Denniston D (1998) Performance of low-loss RF MEMS capacitive switches. IEEE Microw. Guided Wave Lett 8:269–271
7. Fernández-Bolaños M, Perruisseau-Carrier J, Dainesi P, Ionescu AM (2008) RF MEMS capacitive switch on semi-suspended CPW using low-loss high-resistivity silicon substrate. Microelectron Eng 85:1039–1042
8. Hou Z, Liu Z, Hu G, Liu L, Li Z (2006) Study on the reliability of capacitive shunt RF MEMS switch. In: 2006 7th international conference on electronics packaging technology
9. Hwang JCM (2007) Reliability of electrostatically actuated RF MEMS switches. In: RFIT 2007-IEEE international workshop on radio-frequency integration technology, Singapore, 9–11 Dec 2007
10. Yao JJ (2000) RF MEMS from a device perspective. J Micromech Microeng 10(4):R9–R38
11. PandhariPande VM et al (2011) A non-contact type RF MEMS switch to remove stiction problem. IEEE
12. Legtenberg R, Groeneveld AW, Elwenspoek M (1996) Combdrive actuators for large displacements. J Micromech Microeng 6:320
13. de Groot WA, Webster JR (2009) Review of device and reliability physics of dielectrics in electrostatically driven MEMS devices. IEEE Trans Dev Mater Reliab 9(2):190–202
14. Muldavin J, Bozler C, Wyatt P (2012) Stiction in RFMEMS capacitive switches. IEEE
15. Almeida L (2006) Experimental and theoretical investigation of contact resistance and reliability of lateral contact type ohmic MEMS relays. Thesis Master of Science Auburn, Alabama, Dec 15
16. Legtenberg Rob, Groeneveld AW, Elwenspoek M (1996) Comb-drive actuators for large displacements. Micromech Microeng 96:320–329 (IOP Published)
17. Kapoor S, Kumar D, Prasad B (2011) MEMS Electrostatic comb actuators with different materials by using COMSOL 3.5a. Department of Electronic Science Kurukshetra University, Kurukshetra
18. Tang WC (1990) Electrostatic comb drive for resonant sensor and actuator applications. Doctoral dissertation

Comparing the Efficiency of Proposed Protocol with Leach Protocol, in Terms of Network Lifetime

Javad NikAfshar

Abstract The wireless sensor network is a collection which consists of large number of sensor nodes, with small dimensions, and limit communication and calculation capabilities, that in it, energy consumption is very important, and neural network of self-organizing plan (SOM) is an uncontrolled neural network, and creates by neuronal neurons, in a regular grid structure with a low dimension. Also, clustering is performed by using the K-means algorithm, which divides the data set into k subsets, and uses from three important factors to select the cluster header as follows: The sensor which has the highest energy level, the nearest sensor, to the basis station, and the center of the cluster gravity. Additionally, for estimating the cost, the above three criteria used to select the cluster header, plus a new benchmark, namely the number of cluster nodes, and sending data in this method, is similar to sending data step, in CDDA protocol, which in there, after selecting the cluster head, through the mentioned methods, each cluster node distributes its data to other cluster nodes.

Keywords Optimization · Protocol · Wireless sensor networks
Energy consumption

1 Introduction

Sensor network technology is one of the important technologies for the future, and can be considered as the most important technology for 21st Century. Sensor network technology consists of sensor parts, and computational and telecommunication components, which caused that director can view and regulate observations, and as well as react simply to the events that occur in a particular area. Recent developments, on the one hand, in small-scale integrated circuit technology, and on

J. NikAfshar (✉)
Computer Department, Faculty of Engineering, Science and Research Branch,
Islamic Azad University, Tehran, Iran
e-mail: jnikafshar@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_36

the other hand, in development of technology, provided the necessary field for designing the wireless sensor networks. The first method which chooses for precise investigating is the method based on LEACH clustering, which in fact is the basis of all methods based on proposed cluster for these networks. The reason for choosing this method for investigating is that almost all other methods, which work by clustering of nodes, or hierarchically, improved from this method. Therefore, this method was considered as the basis of work, and presentation of optimal model. In the following subsections, first, an energy model uses for estimating the energy consumption of network. It should be noted that this model applies for performing the examinations, and all simulations, and for measuring energy consumption.

2 The Method Based on LEACH Clustering

Low-energy adaptive clustering hierarchical method is a routing algorithm which designed for collecting data and delivers them to the base station. The main objectives of this method are as follows:

1. Increasing the network lifetime.
2. Reducing the power consumption of each network sensor nodes.
3. Using consensus, and aggregating data, for reducing the number of telecommunication messages.

For arriving to these goals, LEACH chooses a hierarchical method, for organizing the network, as a set of clusters. Each cluster is selected, and managed by a cluster head. The cluster head is responsible for performing various tasks. The first task of cluster heads is collecting the data alternatively from cluster members. After collecting data, the cluster heads compact collected data, by removing redundancy, between the correlated data values. The second main task of the cluster heads is direct transferring of compact data to the base station. Compact data transferring is performed by a single jump. This operation is shown in Fig. 1.

The third main task of cluster heads is creating a time schedule based on sharing time, which in there one time slot is assigned to each cluster nodes, and each node can, use from its own time slot, for transferring data. The cluster head announces a time schedule for its cluster members through public broadcasting, and for reducing

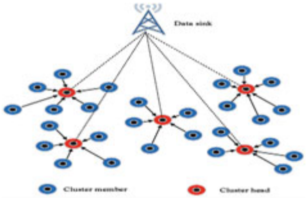


Fig. 1 Performance of Low-energy adaptive clustering hierarchical method

the possibility of an accident, between sensors, inside and between clusters, LEACH nodes, used from a multiple access scheme, and code division, for doing their own communications.

In short, it can be said that the basic operations of this method is carried out in two separate phases. The first phase, which is the beginning phase, consists of two steps, first step selecting the cluster heads, and building the clusters. The second phase which is steady state phase is based on the collecting and aggregating data and delivering them to the data station.

These two phases, are shown in summary form in Fig. 2.

The startup phase length is shorter than the steady state phase, and thus minimized the overhead of the protocol. At the beginning of the startup phase, the selection period of cluster head starts and the process of choosing cluster heads gives us this confidence that this role spins between sensor nodes. In order to determine that whether one node must be selected as a cluster head or not, each nodes generates a random number between zero and one, like V and, and compares it with the selection threshold of cluster, T(n). A node will be cluster head if it's generated random value, V be less than the desired threshold. The cluster selection threshold designing method can ensure, with high probability that the predefined fraction of nodes, such as p, selected in each period as a cluster header. Threshold also ensures that the nodes used in the last I/P of previous periods will not be selected as cluster heads in the current period. In order to obtaining these goals, the T(n) of desired node, such as N, must be determined by the following equation.

$$T(n) = \begin{cases} 0 & if\ n \notin G \\ \frac{p}{1-p\left(r \bmod \left(\frac{1}{p}\right)\right)} & \forall n \in G \end{cases} \tag{1}$$

In this equation, the variable G stands for a collection of the nodes which are not selected in the last 1/p of the rounds to be the head clusters and r refers to the current round. The predetermined parameter of (p) depicts the possibility of the (cluster) head. It is obvious that if a node is used in the last 1/p round as the cluster head, it won't be selected in this round. At the end of the procedure of selecting the cluster head, each node which had been chosen for being the cluster head, announces its new role for the rest of the network. As the announcement for being

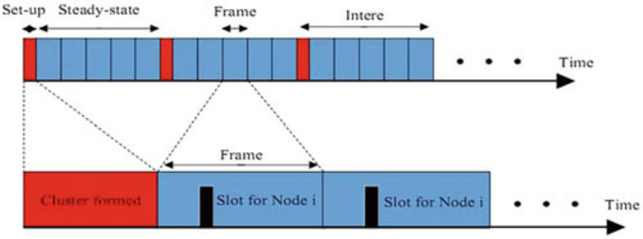


Fig. 2 Two used phases, in a Low-energy adaptive clustering hierarchical method

the cluster head is received, each remaining node selects a cluster to be attached to it. The selection criterion may be due to the power of the received signal among other factors. Afterwards, the nodes inform the selected cluster head from their decision about the membership in that cluster. As the cluster is formed, each cluster head construct and distribute the time sharing. This time planning specifies the time slots which are determined for each member of that cluster. In addition, each cluster head performs a Code Division Multiple Access (CDMA). This code is selected accurately and in a manner that deduct the interference among the clusters. The end of the signals for the start phase, are the beginning the stable mode. At this phase, the nodes gather data and use from the slot specified to them to transfer the gathered data to their heads. This data gathering is performed consecutively.

In comparison to the former methods which use sending high volume of data, it is totally clear that LEACH method is notably useful in saving the energy. The amount of saving the energy is initially related to the proportion of the data density which is obtained from the cluster head. Despite all these profits, this algorithm has some defects. The assumption that all nodes can reach the base station with a jump may not seem logical, since, the capabilities and the energy savings of the nodes may change during time and from one node to the other. In addition, in comparison to the amount of energy decrease for compensating the extra load that is formed during the cluster selection process, the length of the stable mode is variable and critical. A short status period increase the protocol extra load as a long period may cause the energy discharge of the head cluster. Various algorithms are suggested to overcome such problems. Extended X-LEACH protocol was introduced in order to consider the nodes energy level in the cluster head selection process. In this protocol, the resulted threshold of selection of the cluster head meaning $T_{(N)}$, by the node (n) and if that node in the current round will be a cluster head or not, is described by the following equation:

$$T(n) = \frac{p}{1 - p\left(r\,mod\left(\frac{1}{p}\right)\right)}\left[\frac{E_{n,current}}{E_{n,max}} + \left(r_{n,s}div\frac{1}{p}\right)\left(1 - \frac{E_{n,current}}{E_{n,max}}\right)\right] \tag{2}$$

In this equation, $E_{n,current}$ stands for the current energy and $E_{n,max}$, stands for the initial energy of the sensor node. The variable $r_{n,s}$ is the consecutive periods in which a node has not been the cluster head. When the amount of $r_{n,s}$ reaches to 1/p, the threshold T(n) is set to the amount which previously had before entering the remaining energy in the threshold equation and again is set to the equal amount of this quantity. In addition to $r_{n,s}$, when a node becomes the head of the cluster, it will be equal to zero. The protocol X-LEACH has lots of properties that lead to the decrease of the energy use in this protocol. The need for energy in this protocol has been distributed in all sensor nodes, since they consider the role of being the head of cluster in a robin round and according to their remained energy. Actually, this algorithm is totally distributed which does not any controlling information from the base station. This method of managing the clusters is locally performed and thus, eradicates the need to the general information about this network. Furthermore, amassing the data by the clusters will widely help to save the energy, due to the reason that by condensing the data, the length of packets sent in this situation is less than the length of the packets which are not amassed, therefore, there will be less energy for transferring them in comparison to the condition that the packet is not condensed.

The strength of LEACH in the mechanism is rotating the clusters role and gathering the data which could increase the life span of the network, but LEACH has also some weak point.

First, it supposes that all the network nodes have the sufficient power to send the information to the base node and they have enough calculating powers to support the MAC protocols, therefore, it can't be utilized in the extended scale networks. Moreover, the LEACH supposes that the nodes have always some data to send, and the nodes which are close to each other, have dependent data to one another. This protocol supposes that in each selection round, all nodes start with the same energy level, and the cluster approximately consumes the energy which is equal to the amount that other nodes use. The most important defect of LEACH is that it is not clear how the predetermined numbers of the clusters (meaning p) is going to be distributed in the network equally. Actually, it does not offer any guarantee regarding the place or the numbers of the clusters in each round. Thus, it is possible that the selected clusters are centralized in a part of the network. The solution for this problem could be using a centralized Clustering algorithm.

The Centralized LEACH is a clustering algorithm in which the formation of the clusters is done in a centralized way and by the base station. This algorithm has a similar transferring data phase (permanent condition) with LEACH algorithm. Through this algorithm, each node sends some information about its location and the current energy level to the base station. Usually, it is supposed that each node has GPS. The base station should guarantee the uniform distribution of energy among the clusters.

Therefore, it determines a threshold for the energy level and those nodes having more energy than the required amounts are selected as the possible cluster heads. The issue of determining the efficient number of the cluster heads is a NP-Hard issue. LEACH-C uses the simulated fusion algorithm to solve this problem. After determining the clusters of the current round, the base station sends a message containing the cluster ID to each node. If the ID of a node cluster conforms to its own ID, that node is a cluster, otherwise, that is an ordinary node and can be dormant until the phase of transferring its respective data. LEACH-C is more efficient that LEACH and against each unit of energy, approximately transfer 40% extra data, due to the reason that the base station (data sink) has thorough knowledge regarding the location and the energy level of the network nodes. Moreover, unlike LEACH, LEACH-C guarantees the number of the clusters' efficient protocols (K) in each round (Figs. 3 and 4).

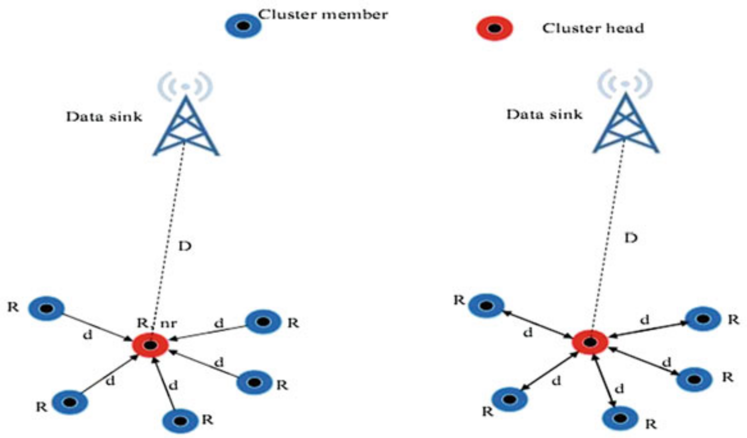


Fig. 3 Sending data in the clusters for LEACH and CDDA protocols

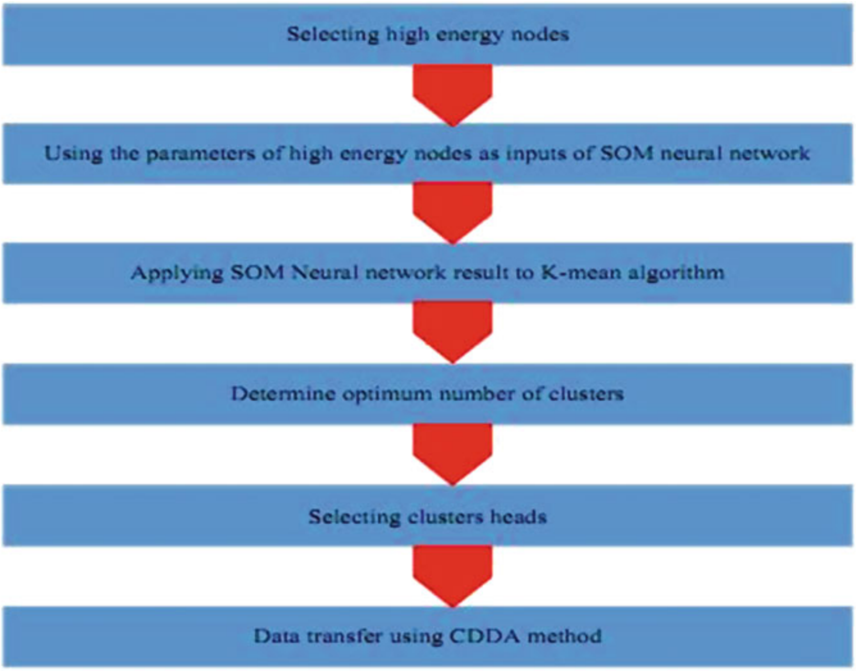

Selecting high energy nodes
Using the parameters of high energy nodes as inputs of SOM neural network
Applying SOM Neural network result to K-mean algorithm
Determine optimum number of clusters
Selecting clusters heads
Data transfer using CDDA method

Fig. 4 The procedures for performing the suggested protocol

3 The Energy Model

Due to the reason that our purpose is the comparing the consumed energy of the entire network for the existing algorithms in various conditions, therefore, it is necessary to have a standard model to calculate the consumable energy in these networks.

In most implementations used for these networks, usually similar models are used in order to measure the total energy of the network, therefore, we also use a similar model as the basis for our experiments [1–4]. Regarding the issue that the energy used for transferring the data packets among the sensor nodes, is more obvious in comparison to other consumable energies, in this model, we assume that the chief consumable energy of the network is related to receiving or sending data and we ignore other consumable energies such as the energy required for sensing or the ones related to processing of the information. In this model, the energy required for transferring L bit of data from the interval d, the following equation is used [1–4]:

$$E_{tx}(L,d) = LE_c + Led^s \tag{3}$$

In this equation, E_C is the base energy that is necessary for performing the sender or receiver circuits. A sample amount for this parameter is *50 ni/bit*, for the receiver-sender of 1 Mbps [1–4]. (e) is the energy unit which is the amplifier of the sender and is obtained for various intervals according to the below equation:

$$\begin{array}{c} e = \{e_1, S_1\} \; \mathit{if} \; dcr \\ e = \{e_2, S_2\} \; \mathit{if} \; d > d_{cr} \end{array} \tag{4}$$

In this relation, d_{cr}, is the threshold interval which is in amount of 86.2 m, the amounts of e_1 and e_2 are 10 pj/bit m^2 and 4 pj/bit m, 0.0013 [1–4].

Furthermore, in this model, the energy required for receiving the L bit of data on d interval, is obtained by the following equation [1–4]:

$$\mathrm{E_{rx} = LE_C} \tag{5}$$

Therefore, the total energy required for transferring L bit of data from the source node to the destination node is obtained by the following equation:

$$\mathrm{E_{I,j}(L,d) = L(a1 + a2d^s)} \tag{6}$$

The variables a2, a1, S are described according to the relations 4–5.

$$\mathrm{a1 = 2E_c} \tag{7}$$

$$\mathrm{a2 = e1 \text{ or } e2} \tag{8}$$

$$\mathrm{S = 2 \text{ or } 4} \tag{9}$$

As it is obvious from the above relations, for the intervals smaller than the threshold, the energy model is considered as a second-handed model, while, for bigger intervals, this model is turned into a four-handed model. Actually, when the traveled interval is increased from a specified amount, by increasing each unit of the interval, the consumable energy is added by power 4.

4 Introducing the Simulation Scenario

Before we accurately mention the results of the simulations, first we introduce the scenario which was used for performing these experiments.

Since, in all the wireless sensor networks applications, all sensors are considered as similar and fixed, we also utilized some similar sensor nodes which were randomly placed in a square-form area. Figure 5 which is drawn by the Matlab software, depicts the random formation of the sensors in the square-form area. The number of nodes varies due to the issue that the respective experiment examines what type of parameters.

5 Comparison of the Efficiency of the Suggested Protocol with LEACH Protocol Concerning the Network Life Span

For the respective comparison, four scenarios are considered. The number of nodes in 1st, 2nd, 3rd and 4th are respectively equivalent to 50, 100, 200 and 400. In all of these scenarios, the area is considered as 100 × 100 m^2, the accurate information regarding the simulation are brought in Table 1.

Another parameter which should be determined in the simulation is the amount of m (the number of nodes having the maximum amount of energy level which are

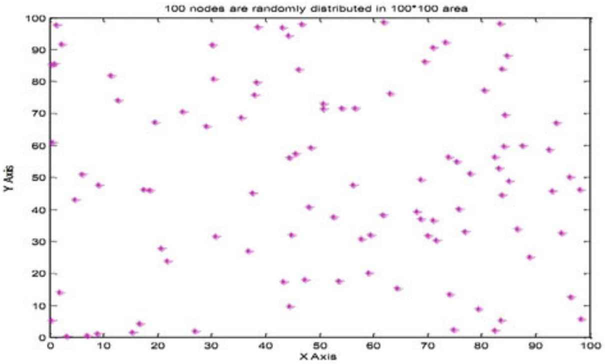


Fig. 5 The wireless sensors are randomly distributed in the area

Table 1 The required data for simulation

Scenario	1st	2nd	3rd	4th
Number of nodes (N)	50	100	200	400
Area	100 * 100	100 * 100	100 * 100	100 * 100
Base station place	(100 & 100)	(100 & 100)	(100 & 100)	(100 & 100)
D_{cr} (Threshold Interval)	86.2	86.2	86.2	86.2
Initial energy	0.5 J	0.5 J	0.5 J	0.5 J
e1	10 pj/bit m^2	10 pj/bit m^2	10 pj/bit m^2	10 pj/bit m^2
e2	0.0013 pj/bit m^4	0.0013 pj/bit m^4	0.0013 pj/bit m^4	0.0013 pj/bit m^4
Packet size	4000 Bits	4000 Bits	4000 Bits	4000 Bits

used as the weights of the self-organized neural networks). This number is experimentally determined and its amount is dependent upon the number of the efficient clusters that we expect to have. In the first scene (with 50 nodes) we suppose that m = 15, in the 2nd scene (with 100 nodes), we suppose that m = 30, in the 3rd scene (with 200 nodes), we suppose that m = 60, in the 4th scene (with 400 nodes), we suppose that m = 120. Furthermore, the amounts of the coefficients related to the expenditure for the cluster head are experimentally determined and based on the importance criterion in decision making.

6 Consequences

The comparison is done by using three standard criteria in the routing algorithms in the sensor network:

First dead time: the number of round which the first node stop working due to ending energy, half dead time: the number of round which half of the network nodes (in the 1st scenario, 25 nodes, in the 2nd scenario, 50 nodes, in the 3rd scenario, 100 nodes, in the 4th scenario 200 nodes) stop working due to ending energy. Last dead time: the number of round which the last node of the network stops working due to ending energy.

As it is clear in Figs. 6, 7, 8 and 9, the suggested protocol is superior than LEACH protocol. These results reveal that the suggested algorithm of all the scenarios which were experimented, guarantee the network life, complete survive of the network (network coverage). This feature is important in applications in which the network coverage has a special significance. For instance, in Monitoring or Tracking applications which need very accurate data, postponing the first dead time is more important than postponing the last dead time, though, in some applications it is possible that increasing the total life time of the network (delaying the last dead time) is more appropriate, for example in the weather periodic monitoring.

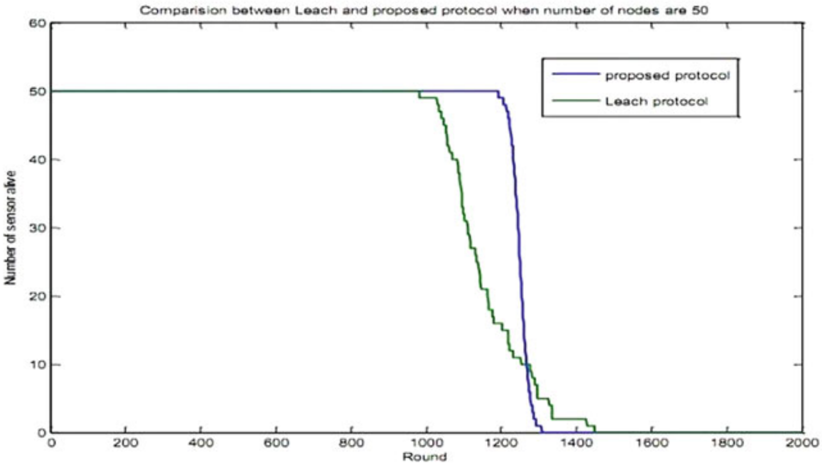


Fig. 6 Comparison between LEACH protocol application and the suggested protocol for the 1st scenario

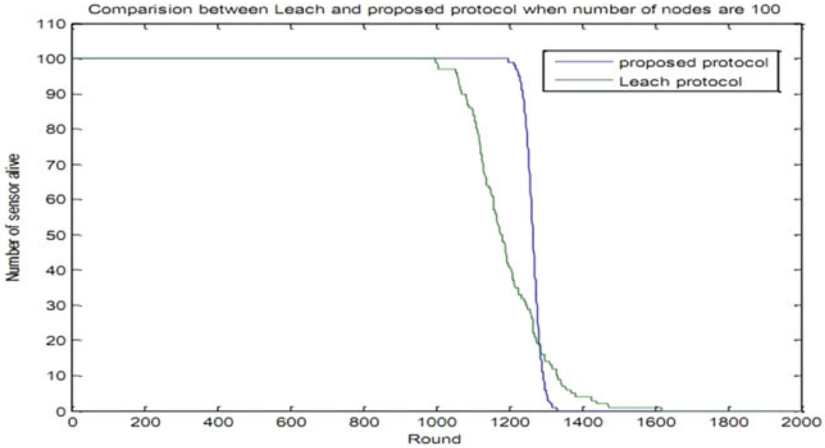


Fig. 7 Comparison between LEACH protocol application and the suggested protocol for the 2nd scenario

If the improvement of the network life time regarding LEACH protocol, is described as follows, we would have:

$$\text{Network life time improvement} \\ = \frac{\text{First dead time in ith scenario} - \text{first dead time in jth scenario}}{\text{First dead time in ith scenario}} \times 100$$

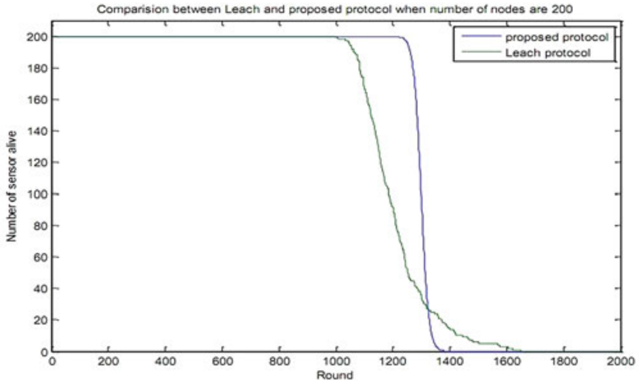


Fig. 8 Comparison between LEACH protocol application and the suggested protocol for the 3rd scenario

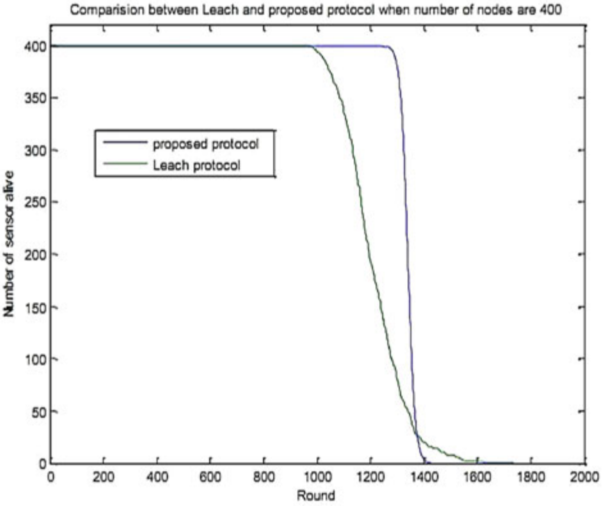


Fig. 9 Comparison between LEACH protocol application and the suggested protocol for the 4th scenario

According to the data obtained from the above figures:

For the 1st scenario, it equals to:

$$1221 - 926 \div 926 \times 100 = 31.85\%$$

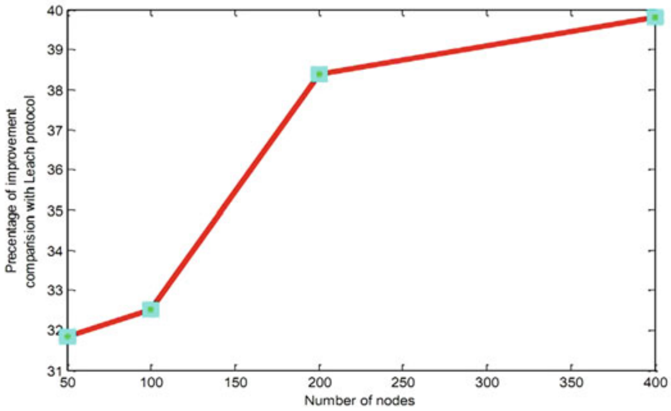


Fig. 10 Improvement percentage of the suggested protocol compared to LEACH protocol against different nodes

For the 2nd scenario, it equals to:

$$1235 - 932 \div 932 \times 100 = 32.51\%$$

For the 3rd scenario, it equals to:

$$1304 - 941 \div 941 \times 100 = 38.38\%$$

And For the 4th scenario, it equals to:

$$1338 - 957 \div 957 \times 100 = 39.81\%$$

As you observed in above calculations, in all scenarios, there are 30% improvement compared to LEACH protocol. Another important result in our suggested algorithm is that by increasing the number of nodes, there would be more improvements. These results are obvious in Fig. 10.

References

1. Schurgers C, Srivastava MB (2001) Energy efficient routing in wireless sensor networks. In: Proceedings of the IEEE Military Communications Conference (MilCom): communications for network-centric operations-creating the information force, McLean, VA
2. Xiangning F, Yulin S (2007) Improvement on LEACH protocol of Wireless Sensor Network. In: Proceedings of IEEE international conference on sensor technologies and applications, pp 260–264

3. Banerjee B, Khuller S (2001) A clustering scheme for hierarchical control in multi-hop wireless networks. In: Proceedings of INFOCOM
4. Chen H, Wu C, Chu Y, Cheng C, Tsai L (2007) Energy residue aware (ERA) clustering algorithm for leach-based wireless sensor networks. In: Proceedings of second international conference on systems and networks communications (ICSNC), IEEE Computer Society

Voltage Stability Enhancement Along with Line Congestion Reduction Using UPFC and Wind Farm Allocation and Sizing by Two Different Evolutionary Algorithms

S. Ehsan Razavi, Mohsen Ghodsi and Hamed Khodadadi

Abstract Due to the growing demand of electricity, adding the new power lines for compensating the demand is not possible. Connecting the renewable energy sources, as a favourable alternative source of energy, to the power grid makes some new challenges. In this paper, allocating and sizing the UPFC along with wind farms is covered. The most important purpose of this study is to mitigate the congestion of the power transmission line along with improving the voltage stability margin. For this objective, two important evolutionary algorithms, namely Genetic Algorithm and Particle Swarm Optimization are utilized for finding the optimum size and place of the UPFC and wind farm. An IEEE 24 bus RTS system is simulated. Load data is selected according to IEEE database and mimic the real load situation. The wind data is collected from Manjil, a city in north of Iran, and the output power of the wind farms is simulated. At the end, the optimum place and size for UPFC and wind farm is presented in order to have the minimum congestion for transmission line along with maximum voltage stability margin, system load ability.

Keywords Voltage stability margin · Power line congestion · UPFC
Wind farm · Genetic algorithm and particle swarm optimization

S. Ehsan Razavi (✉) · M. Ghodsi
Department of Electrical Engineering, Islamic Azad University,
Mashhad Branch, Mashhad, Iran
e-mail: e_razavi_control@yahoo.com

H. Khodadadi
Department of Electrical Engineering, Islamic Azad University,
Khomeinishahr Branch, Isfahan, Iran

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_37

1 Introduction

Application of the renewable energy resources in power system requires considering different aspects such as environmental impact minimization especially on conventional plant [1]. Some of the technical problems in merging the wind energy and existing power system are considering voltage regulation, power quality and stability problems. One of the essential customer-focused measurements is power quality which also is effectively influenced by the transmission operation and distribution network [3]. In recent years, quick development and extensive rising in the employing of the wind energy is considered. Large capacity of individual units can be feed into distribution network [2]. All the fluctuations in the wind speed for fixed-speed wind turbine operation affected both of the electrical power and mechanical torque on the grid. The results can be shown as large value fluctuations on the voltage.

The power quality can be considered as several aspects. Wind, generation, transmission and distribution of electrical power, for example voltage swells, sag, harmonics and flickers are some of this issues. Employing the wind generator makes some disturbances for the distribution network. Using the induction generator with direct connect to the power system is one way to start up a wind generating system. Main benefit of using the induction generator is its robustness and cost effectiveness.

However; reactive power, wind and voltage of the terminal in an induction generator produce some variation in induction generated power [3].

Several approaches have been proposed by researchers for installation of the UPFC in power systems [4–6] and its specifications have been investigated in the related literature [7]. In recent years, there have been many different researches around the UPFC topic and several studies have been published concern to UPFC analysis, control, modelling and application. The effects of converters and the generators dynamics have not been considered on the steady state characteristics of UPFC mathematical models [8, 9]. The conventional controllers are used to design a series converter for UPFC [10, 11]. Topologies related to many power converters are proposed to implement the FACTS devices like multi-level inverters and multi-pulse converter such as 48 and 24 pulses [12–14]. Baskar et al. [15] reported the limitations and merits of high power converters. In [16] analysis of the UPFC dynamic control with six pulse converter is modeled using switching level. In this paper by considering all the mentioned problems in above, feed-back line is considered in the controller to add the ability to be smarter in sudden variation as well as being able to be controlled from distances. This paper is categorized as below, in Sect. 2. Unified Power Flow Controller (UPFC) with its basic principle is explained, Sect. 3. UPFC injection model and Sects. 4 and 5 covers the two evolutionary algorithms that are used in this paper, Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), respectively. Section 6 covers the congestion management as well as Sect. 7 that covers the Voltage stability margin problem. Finally, Sect. 8 presents the simulation and the numerical results.

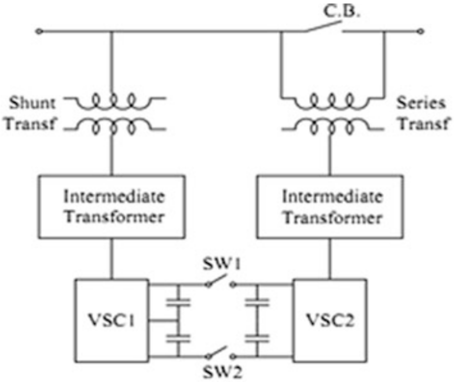
2 Unified Power Flow Controller (UPFC)

UPFC is one of the main adjustable FACTS controller used to regulate the voltage and power in a transmission line. UPFC is composed of one series-connected, one shunt-connected and two Voltage Source Converters (VSC). The configuration of DC capacitors are in parallel form as illustrated in Fig. 1, which are related to two converters. If both 1 and 2 switches operate in open mode, the two converters work as SSSC and STATCOM control the reactive voltage and current injected into the line in series and shunt, respectively. Closing switches 1 and 2 permit both converters for exchanging the active power flow between them. By the series-connected converter, the active power can be either supplied or absorbed [17].

By injecting a series voltage to the transmission line, the series converter acts the main performance of the UPFC, with controllable phase angel and magnitude. UPFC provides simultaneous reactive and active series compensation with no need to any external energy source. If there is no angular constraint, the UPFC can control the voltage impedance, transmission angle or reactive and active power flow through the line by the injected voltage. Similar to the SSSC, the injected voltage via the series converter is produced internally by itself and the active power which is provided by the shunt converter is transported by the DC link. Supplying or absorbing the active power demanded through the series converter is the main function of the shunt converter. By generating or absorbing active power from the bus, the DC capacitor voltage is controlled by the shunt converter. Consequently, shunt converter performs such a parallel synchronous source in the system. It also can support the controllable reactive compensation for the bus similar to the STATCOM. The main functions of the UPFC are as belows:

- Similar to a tap-change transformer, voltage regulation in anti-phase/phase voltage injection in a continue mode.
- The compensation of series reactive through injection of a voltage which is in quadrature to the line current.

Fig. 1 A UPFC schematic



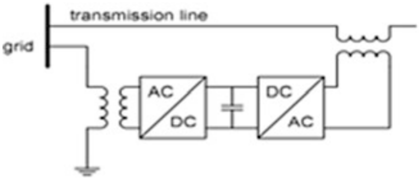


Fig. 2 UPFC configuration

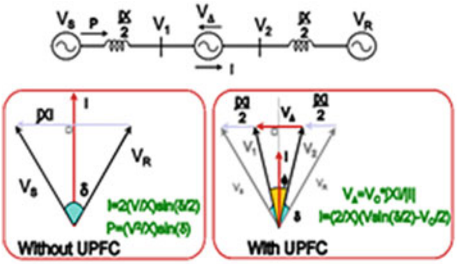


Fig. 3 UPFC operation

- Phase shifting by injecting a voltage in which the voltage angle is related to the bus voltage. By changes in voltage amplitudes, the phase shift can be controlled.

The UPFC functions can be executed simultaneously, which makes the UPFC the most powerful PFCD [18] (Fig. 2).

UPFC has three degrees of freedom by means of controlling three features simultaneously or selectively while the other FACTS elements have only one. By combination of more converters, more flexibility and degrees of freedom can be provided [17, 19] (Fig. 3).

3 UPFC Injection Model

In order to achieve a model for UPFC injection, Fig. 4 is presented for illustration of the considered essential series of voltages.

The indicates the seen reactance between the series transformer terminals and describes as [20, 21]:

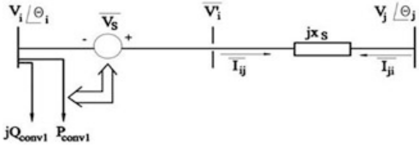


Fig. 4 The electrical circuit of the UPFC

$$x_s = x_k r_{\max}^2 \left(\frac{S_B}{S_s}\right) \tag{1}$$

$$b_s = \frac{1}{x_s} \tag{2}$$

In which the *xk* is the series transformer reactance, *r* max denotes the maximum magnitude of injected voltage, *SB* describes the base power of the system, *SS* = *S* conv2 and are the nominal series converter power and ideal series voltage, respectively. The magnitude of phase of the system can be controlled [20, 21].

$$\begin{gathered}\overline{V_s} = r\overline{V_i^{j\gamma}} \\ 0 \leq r \leq r\text{max2} \\ 0 \leq \gamma \leq 2\pi\end{gathered} \tag{3}$$

In which, *r* and γ are the magnitude and the angle of the injected voltage, respectively. The model of the UPFC (Fig. 5) can be presented as [20, 21]:

$$Psi = -rbsViVj\sin(\theta i - \theta j + \gamma) \tag{4}$$

$$Qsi = -rbsV2i\cos(\gamma) + \text{Qconv1} \tag{5}$$

$$Psj = rbsViVj\sin(\theta i - \theta j + \gamma) \tag{6}$$

$$Qsj = rbsViVj\cos(\theta i - \theta j + \gamma) \tag{7}$$

$$Pi1 = -rbsViVj\sin(\theta i - \theta j + \gamma) - bsViVj\sin(\theta i - \theta j) \tag{8}$$

$$Qi1 = -rbsV2ico(\gamma) + Q\text{conv1} - bsV2i + bsViVj\cos\theta i - \theta j) \tag{9}$$

$$Pj2 = rbsViVj\sin(\theta i - \theta j + \gamma) + bsViVj\sin(\theta i - \theta j) \tag{10}$$

$$Qj2 = rbsViVj\cos(\theta i - \theta j + \gamma) - bsVj2 + bsViVj\cos(\theta i - \theta j) \tag{11}$$

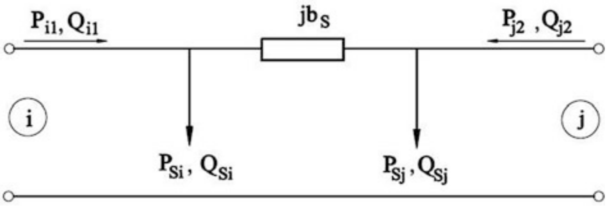


Fig. 5 Injection model of the UPFC [20]

4 Genetic Algorithm

The base of the GA is set according to biological concept. Natural selection mechanism in nature is the basis for this algorithm. A list of binary digit code operates on control parameters and chromosomes [16]. Among other evolutionary algorithms GA robustness is a prominent drawback. Similar to PSO, GA has three main advantages.

To search in a problem space, this approach uses the population method instead of single method. Therefore, several areas will be explored and covered and avoid from local optimum points. For this purpose, a new method is presented in [17, 18].

There is no need to prior knowledge about the function which wants to be optimized. Moreover, is not require to have any space limitation like smoothness, existence of derivatives, convexity or uni-modality.

In space problem each string can be considered as a chromosome which describe one solution for the candidate, in PSO each particle [19]. Furthermore, in [20] the GA, PSO and HBMO is used for improving the margin of voltage stability under different scenarios. The several steps of GA is discussed in some studies such as [20].

5 Particle Swarm Optimization

PSO is another evolutionarily searching algorithm which is utilized for seeking optimal solution in any given problems. Russel Eberhart and James Kennedy in 1995 are developed this method based on the birds flocking behavior imitation. Firstly, without considering any determined destination, all birds fly until one of them find the best possible point and selected as the leader of the others. Direction as well as velocity can be defined by the leader. Once the destination is found, the other birds are conducted to the destination by the leader.

In the PSO approach, by change the acceleration (velocity) at each time step, the pbest and gbest are calculated. The position and speed of each particle are as follows:

$$V_{id}^{t+1} = W \times V_{id}^{t} + c_1 \times \psi_1 \times \left(p_{id}^{t} - x_{id}^{t}\right) + c_2 \times \psi_2 \left(p_{gd}^{t} - x_{id}^{t}\right) \tag{12}$$

$$x_{id}^{t+1} = x_{id}^{t} + x_{id}^{t+1} \tag{13}$$

(1) Where v_{id}^{t} and x_{id}^{t} are the ith particle velocity and position in iteration t, respectively. c_1, c_2 and W denote the weighting factor and inertia weight. p_i and p_g illustrate the best position obtained by ith particle and ith neighbors of particle, respectively. ψ_1, ψ_2 indicate two random factors in the [0, 1] interval. The PSO procedures are discussed in some studies such as [16].

6 Congestion Management

In the electric, the amount of transferable power between two points in the grid can be dictates by the network constraints. In practice, some constraints and network limitation make it impossible to deliver all of the generated power. Violation of operating constraints like over-load (congestion) of lines and voltage limits are some of the main constraints affect on cost related to the pool demand. The limitation on transmission is called congestion. Due to causing some outages in cascade with uncontrolled loss of load, presence of congestion is not be allowed. Cost-free means like congested branches outage (lines or transformers) can mitigate the congestion. The main device to do so is operation of transformer taps and using FACTS devices. Transmission access rationing should be performed to manage the congestion. The philosophy of rationing is based on user-pay system. The objective function of dispatch problem in a system by the multilateral and bilateral dispatches without any pool load [22] can be described as below:

$$\min f(u,x) = \left[u - u^0 A\right] w \left[\left(u - u^0\right)^T A\right] \tag{14}$$

where u is denoted the control variables including the active power may be injected or extracted to the generator and load buses, w as a diagonal matrix in which, the elements are willingness-to-pay price premiums for avoiding the transmission curtailment; x as the dependent variables; u^0, the desired values for the u and A as a fixed matrix used to reflect the curtailment strategies of market. The w as the willingness-to-pay parameters are consequently based on the customer and have a relation to the lost load values. The users which have sensitive loads can employed the w with high value caused to reducing the curtailment. When both of the pool and bilateral transactions exist, the (15) illustrates an optimal 'curtailment' problem.

$$\begin{aligned} \min \sum_{i \in I_G} C_i(P_{pi}) - \sum_{i \in I_D} B_j(D_{pj}) + \sum_{j \in I_D} w_{Dj}\left(D_{pj} - D_{pj}^0\right)^2 \\ + \sum_{k \in I_T} w_{Dk}\left(P_{tk} - P_{tk}^0\right)^2 - \sum_{k \in I_T} C_{tk}(P_{tk}) \end{aligned} \tag{15}$$

where IG is the pool generator buses; ID indicates the pool load buses; Ppi denotes the active power of ith pool generator; C_i describes the bid price of ith pool generator; D_{pj} and B_j illustrate the active power and bid price of jth pool load; P_p and D_p are the pool power injection and extraction vectors; P_t, Q and V describe the bilateral contract, reactive powers, voltage magnitude and angle vectors, respectively; D_{0pj} denotes the target value of pool demand at jth bus; P_{0tk} indicates the target value of kth bilateral contract; C_{tk} is the charge paid for delivering Ptk; IT illustrates the set of bilateral/multilateral transactions P_t, and finally, X_k is the control parameter of FACTS device.

7 Voltage Stability Margin

Finding the weakest voltage buses is one of the best possible solution in wind DGs and FACTS allocation. Loss minimization is another advantage of appropriate sitting and placement of wind DGs and FACTS. Voltage and thermal constraint are two significant limitations for power system capability. The ampacity, limits the conducted current of a conductor, is well known as thermal limitation or maximum current capacity. Moreover, allowable minimum and maximum variation in voltage is known as voltage limitation.

(1) Evaluating the voltage stability based on static technique can be calculated by the P-V curve (the relation between voltage (*V*) and power (*P*)) which can be drawn through the method of Continues Power Flow (CPF).

(2) As indicated in Fig. 6 the $\lambda_{\max}$ is the presentation of the system maximum load. $\lambda_{\max}$ is directly related to the Jacobian of the power flow equation (as a singular point). This is the MW distance between operation and critical points. Voltage stability margin can either be decreased or increased by considering the penetration of the either DG or FACTS devices in the power system, depend on the lead or lag forms of power factor. In Fig. 6, which illustrates maximizing load ability and stability margin influenced by FACTS unit, *X*-axis represents λ as the scale factor of load demand in an operational point. λ can be variated from zero to $\lambda_{\max}$. In a normal operational point, maximum load ability increase from $\lambda_{\max 1}$ to $\lambda_{\max 2}$, and consequently, voltage increase from V_1 to $V_{2.}$ Both dynamic and static perspective for voltage stability analysis should be inspected. Although static approach is preferred for assessment and control, dynamic approached is time consuming. One of the most important factor of the presented method is voltage stability index, computes the voltage collapse approximation. Several method has been proposed for evaluating this index. L-index as one of such methods presented in [23]. The 0 and 1 are obtained for this index in no load and voltage collapse condition. Consequently, the maximum value of L-index in each bus represent the most vulnerable bus among the others.

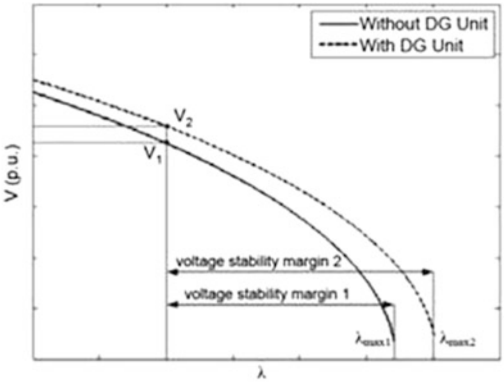


Fig. 6 The effect of DG units on increasing the voltage stability margin

$$L_j = \left| 1 - \sum_{i=1}^{N_g} F_{ij} \frac{V_{gi}}{V_{gj}} \angle\left(\theta_{ij} + \delta_i - \delta_j\right) \right| \tag{16}$$

(2) Where V_{gj} and V_{gi} denote the amplitude of jth and ith generator voltage and the θ_{ij} describes the angle of $F_{ij}\delta_j$. and δ_i are the voltage phase angle of jth and ith generator, respectively. Matrix F_{LG} yields the F_{ij} values.

(3) L-indices can be obtained at load buses in any given state. Finally, $\mathrm{L}^{\max}$ as the largest L-indices can illustrate the voltage collapse proximity. $\mathrm{L}^{\max}$ can be assumed as a measuring tools for actual state distance estimation used to obtain the stability limit.

8 Simulation and Numerical Results

A 24 IEEE RTS system, illustrated in Fig. 7, is simulated and utilized in this paper. The main objectives are reducing the power line congestion as well as improving the voltage stability margin.

The load is simulated hourly as shows in Fig. 8 base on IEEE RTS data and wind output power is simulated using method and the wind variation data is collected from Manjil, a country in north of Iran. Among all of the calculated data for wind and load the 362th day in a year is selected to apply the simulation for 24 h. After determination of the wind speed, the calculation of the wind generator output

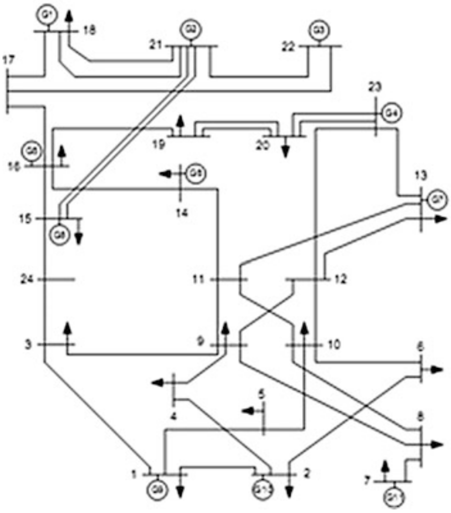


Fig. 7 IEEE 24 reliability test system

Fig. 8 Hourly load data

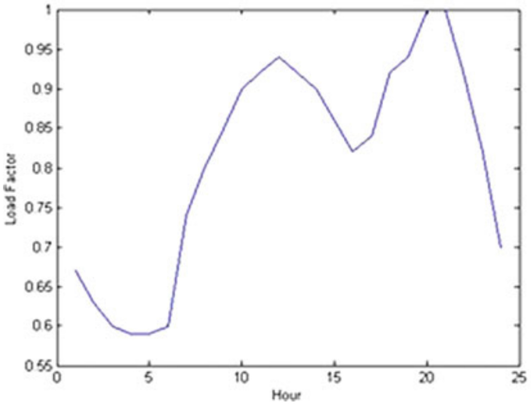
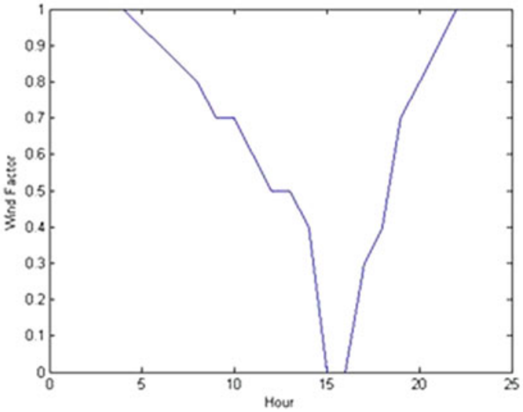


Fig. 9 Wind farm output power



power is by (17). Where, in V_r it reaches the nominal power and continues its power output up to V_{cout}. V_{cin} is the speed in which the wind farm begins to operate.

$$P_{WF} = \begin{cases} 0 & x < V_{cin} \\ P_r.(A + Bx + Cx^2) & V_{cin} \leq x < V_r \\ P_r & V_r \leq x < V_{cout} \\ 0 & x \geq V_{cout} \end{cases} \quad (17)$$

In which, A, B and C parameters have some constant values, depend on the wind turbines. By assuming a constant power factor for a wind turbine, the reactive power generated can be achieved. Output power of a WPG is the sum of wind turbines power outputs (Fig. 9).

In Table 1 the results of this simulation is presented. As it is evident the Lambda, which is a factor that illustrate the voltage stability margin, for the base grid is 0.4798. after apply the proposed method to the power system it enhanced to 0.8935 by using PSO and increased to 0.9677 in using GA while the power line losses also

Table 1 The simulation result for wind farm and UPFC allocation and sizing

	LAMBDA	LOSS	Congestion	R	Line	Q_conve	GAMA	BUS WF	CAP. WF
Base Grid	0.4798	0.6578	315.8573	0	0	0	0	0	**0**
PSO	0.8935	0.3673	98.7513	0.008	38	−68	−2.9762	20	340
GA	0.9677	0.33	95.6898	0.0084	2	−36	−2.9762	23	360

decreased from 0.6578 for the basic grid to 0.3673 and 0.33 for PSO and GA respectively. As mentioned in above congestion of the power line is another objective function for the proposed method which is minimized from 315.85 for the base grid to 98.75 and 95.68 for PSO and GA respectively after calculating the optimum place and sizing for the UPFC and wind farm. Capacity of the wind farm is 340 for PSO and is located at 20th bus, and also for 360 for the GA while it is located at the 23th bus. Injected reactive power from UPFC is 68 for PSO and 36 for GA while the GAMA, the value of injected voltage angle, according to Eq. (3) for both PSO and GA is −2.97.

9 Conclusion

In the current paper, two important evolutionary algorithms, namely GA and PSO are utilized for determination of the optimum size and place of the wind farms along the UPFS. Minimizing the congestion of the transmission line as well as improving the voltage stability are the two important purposes of this study. Furthermore, the power losses in the simulated system, 24 bus IEEE RTS, is also considered as presented in Table 1. Evidently the loss after the siting and sizing increased significantly. In comparison of the two evolutionary algorithms results related to GA are satisfying than PSO. By finding a proper placement for wind farm and UPFC not only power line congestion and loss are minimized but also voltage stability margin is improved, LAMBDA. The main reason for GA operation is due to the ability of cross over and mutation is the algorithm that can prevent the answer from local minimums and lead it to the global one.

References

1. Hook KS, Liu Y, Atcitty S (2006) Mitigation of the wind generation integration related power quality issues by energy storage. EPQU J XII:2
2. Carrasco J (2006) Power electronic system for grid integration of renewable energy sources. IEEE Transl 53:1002–1014

3. Deppa SN, Roger Rozario AP, Kumar M, Yuvaraj V (2011) Improving grid power quality with FACTS device on integration of wind energy system. IN: Fifth Asia modeling symposium, vol 1
4. Hao J, Shi LB, Chen Ch (2004) Optimizing location of unified power flow controllers by means of improved evolutionary programming. IEE Proc Genet Transm Distrib 151(6):705–712
5. Hingorani NG, Gyugyi L (2000) Understanding FACTS: concepts and technology of flexible AC transmission systems. IEEE Press, New York
6. Gyugyi L, Rietman T, Edris A (1995) The UPFC power flow controller: a new approach to power transmission control. IEEE Trans Power Delivery 10(2):1085–1092
7. Papic I (2000) Mathematical analysis of FACTS devices based on a voltage source converter, part II: steady state operational characteristics. Electr Power Syst Res 56:149–157
8. Papic I (2000) Mathematical analysis of FACTS devices based on a voltage source converter, part 1: mathematical models. Electr Power Syst Res 56:139–148
9. Round SD, Yu Q, Norum LE, Undeland TM (1996) Performance of a unified power flow controller using a D-Q control system. In: IEEE AC and DC power transmission conference, Publication IEE No 423, pp 357–362
10. Yu Q, Round SD, Norum LE, Undeland TM (1996) Dynamic control of UPFC. IEEE Trans Power Delivery 9(2):508–514
11. Soto D, Green TC (2002) A comparison of high-power converter topologies for the implementation of FACTS controllers. IEEE Trans Industr Electron 49(5):1072–1080
12. El-Moursi MS, Sharaf AM (2005) Novel controllers for the 48-pulse VSC STATCOM and SSSC for voltage regulation and reactive power compensation. IEEE Trans Power Syst 20 (4):1985–1997
13. Rodriguez J, Lai JS, Peng FZ (2002) Multilevel inverters: a survey of topologies, controls and applications. IEEE Trans Industr Electron 49(4):724–738
14. Lee CK, Leung JSK, Hui SYR, Chung HSH (2003) Circuit level comparison of STATCOM technologies. IEEE Trans Power Electron 18(4):1084–1092
15. Baskar S, Kumarappan N, Gnanadass R (December 2010) Switching level modelling and operation of unified power flow controller. Asian Power Electron J 4(3)
16. Padiyar KR (2007) FACTS controllers in power transmission and distribution. New Age International Publishers, New Delhi, pp 1–4, 240–264
17. Yuan Z (2010) Distributed power flow controller. Delf University of Technology, The Netherlands, pp 26–30
18. Paserba JJ (2004) How FACTS controllers benefit AC transmission system. IEEE 2:1257–1262
19. Singh B (2012) Introduction to FACTS controllers in wind power farms: a technological review. Int J Renew Energy Res 2:47
20. Dizdarevic N (October 2001) Unified power flow controller in alleviation of voltage stability problem. PhD thesis, University of Zagreb, Faculty of Electrical Engineering and Computing, Department Power Systems
21. Izadpanah Tous S, Gorji M (2011) Unified power flow controller and its working modes. In: Presented at 2011 World Congress on engineering and technology, China
22. Fang RS, David AK (1999) Optimal dispatch under transmission contracts. IEEE Trans Power Syst 2(14):732–737
23. Anbarsan A, Sanavullah MY (November 2012) Voltage stability improvement in power system by using STATCOM. Int J Eng Sci Technol 4(11)

Analysis of a Multilevel Inverter Topology

Shahrouz Ebrahimpanah, Qihong Chen and Liyan Zhang

Abstract Nowadays multilevel inverters have become more popular in electric high power application and the connection of renewable energy such as solar and wind energy to the power grid is an interesting subject. Many topologies and control methods have already been suggested. In this paper, first of all, gives a brief survey of the three prevalently used multilevel inverter systems. Then, in continue there will be examined and compared the most usual multilevel topologies and finally, a short review is presented to verify the control methods for power converters.

Keywords Multilevel inverter · Diode-clamped inverter · Flying capacitor inverter · Cascaded H-Bridge converter · Power converter

1 Introduction

Output in multilevel power electronic converters can be created by joining some DC sources. Batteries, Solar panels, fuel cells, and ultra-capacitors can be named as the most usual independent sources. These converters can be found in two kinds of single phase and three phases [1]. Overall three different major multilevel converters are arranged in groups of flying capacitor, diode-clamped and cascaded H-bridge. The voltage of three levels is known as the smallest number among multilevel converter topologies. If we want the output THD comes near to zero, we need to increase the number of levels to infinity [2]. Multiple voltage levels are provided in the diode-clamped inverter by connection of the phases to a series bank of capacitors. This can be used in high-power ac motor drives in conveyors, pump, fans, and mills. Series connection of capacitor clamped switching cells is used for flying capacitor. Applications of flying capacitor multilevel converters include

S. Ebrahimpanah (✉) · Q. Chen · L. Zhang
School of Automation, Wuhan University of Technology,
P.O. Box No. 205 Luoshi Road, Wuhan, China
e-mail: shahrooz6485@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_38

high-bandwidth, high-switching frequency systems as an example of medium voltage traction drives. At last cascaded H-bridge inverter will be made up of series power conversion cells. So we can use Cascaded H-bridge multilevel converter for high-power and high-quality systems such as static volt-ampere reactive generation, active filters, reactive power compensators, photovoltaic power conversion, uninterruptible power supplies, and magnetic resonance imaging [1].

2 Diode-Clamped Inverter (DCC)

Diode Clamped Multilevel Inverter is also called Neutral-Point Clamped Inverter (NPC). The use of voltage clamping diodes in the Diode Clamped Inverter topology is necessary [3]. Figure 1 shows Diode-clamped multilevel inverter topologies. In A three-level diode clamped inverter, one phase leg consists of two pairs of switches and two diodes. Every pair of switches works in consummating mode with the other pair working properly to prevent short-circuiting the DC source, and the diodes are used to set approach to the mid-point voltage. In a three-level inverter a common dc bus, which has been subdivided by two capacitors into three levels, is shared by each of the three phases. By using two series link of DC capacitors (C1 and C2), the DC bus voltage is divided into three voltage levels. Through the clamping diodes Dc1 and Dc2, for every switching device the voltage stress is limited to V_{dc}. It is assumed that the total dc link voltage is V_{dc} and midpoint is regulated at half of the dc link voltage, the voltage across each capacitor is $\frac{V_{dc}}{2}\left(V_{C1} = V_{C2} = \frac{V_{dc}}{2}\right)$. We can find three different possible switching stages for three level diode clamped inverter,

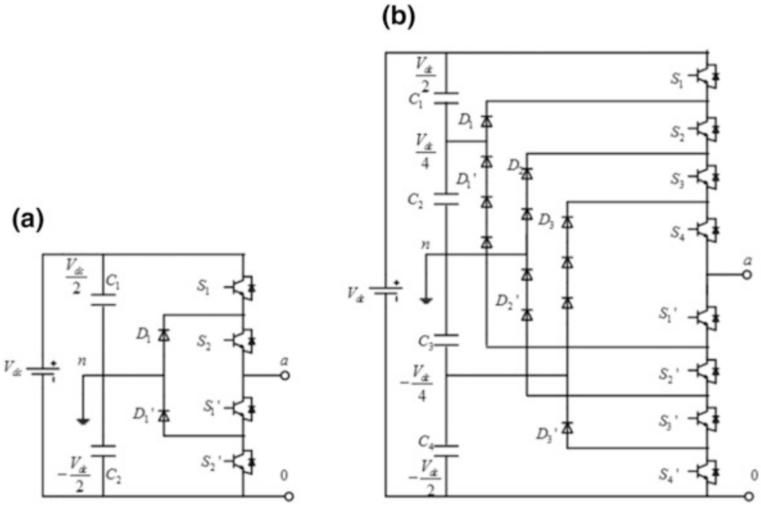


Fig. 1 Diode-clamped multilevel inverter topologies. **a** Three-level. **b** Five-level

Table 1 Switching states in one leg of the three-level DCC

Voltage level	Complementary pair no. 1		Complementary pair no. 2		Leg voltage (v_{an})
	S_1	S'_1	S_2	S'_2	
1	1	0	1	0	$\frac{V_{dc}}{2}$
2	0	1	1	0	0
3	0	1	0	1	$-\frac{V_{dc}}{2}$

which apply the set of steps voltage on output voltage according to DC link capacitor voltage rate. The switching states of the three-level converter are summarized in Table 1 to study the effect of the number of output voltage levels in a diode clamped topology. In a three-level inverter and a five-level inverter in order, a set of two switches and a set of four switches are on at any given time and this process is going on for other levels.

3 Flying Capacitor Inverter (FCC)

The structure of this inverter is similar to that of the diode-clamped inverter except capacitors are used in the Flying Capacitor (FC) topology, instead of diodes to clamp the voltage beyond the devices to a segment of the whole DC voltage [4]. Figure 2 shows Capacitor-clamped multilevel inverter circuit topologies. We need more capacitors by increasing the number of levels. If the flying capacitor works in situation of stability mode and v_{DC} is the input DC-link voltage, then in order to have same step voltages as the output, the clamped capacitor should be adjust to a specific level at $v_{C1} = \frac{V_{DC}}{2}$ in the three-level inverter, and at $v_{C2} = 2v_{C1} = \frac{2V_{DC}}{3}$ in the four-level inverter. For an m-level converter the voltage rating of the capacitors in an FCC is $\frac{V_{DC}}{m-1}$. An m-level FCC will consist of $(m-1)$ DC-link capacitors, with $\frac{(m-1)\times(m-2)}{2}$ flying capacitors in each phase leg [5]. Table 2 is showed the switching states of the three-level converter to study the effect of the number of output voltage levels in a Flying Capacitor topology. The demand of a difficult and complex control strategy in order to enable regulation of the floating capacitor voltages is one of the main problems at FCC.

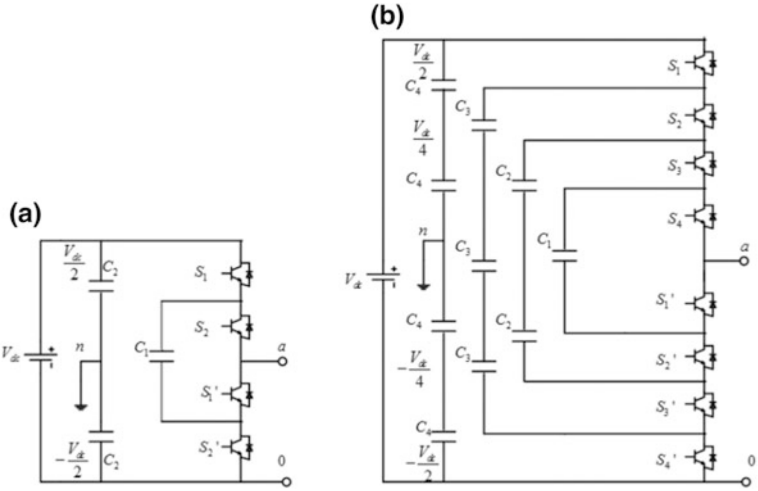


Fig. 2 Capacitor-clamped multilevel inverter circuit topologies. **a** Three level. **b** Five-level

Table 2 Switching states in one leg of the three-level FCC

Voltage level	Complementary pair no. 1		Complementary pair no. 2		Leg voltage (v_{an})
	S_1	S'_1	S_2	S'_2	
1	1	0	1	0	$\frac{V_{dc}}{2}$
2	1	0	1	0	0
	0	1	0	1	
3	0	1	0	1	$-\frac{V_{dc}}{2}$

4 Cascaded H-Bridge Converter (CHBC)

The third topology for the Multilevel Inverter is the cascaded H-bridge inverter, which can be created a compound by a series of single-phase, full-bridge converters. There are some advantages in the cascaded inverter topology that have made it interesting among different applications. First of all, it is modular. Without changing the dimension of the system it is easy to plug into more separate DC sources because each DC source is provided into an individual full bridge inverter. In addition, inasmuch the switch and diode need only withstand one separate DC voltage; the switching stress would be less than the regular two levels topology for each switch device. At last, as mentioned above, the cost of the filter would be decreased since the output voltage waveform is almost sinusoidal [6]. You can see the structure of a cascaded H-bridges converter in Fig. 3. A single H-bridge is a three-level converter. In a three-level converter, the four switches S_1, S_2, S_3 and S_4 are controlled to generate three discrete outputs V_{out} with levels V_{dc}, 0 and $-V_{dc}$.

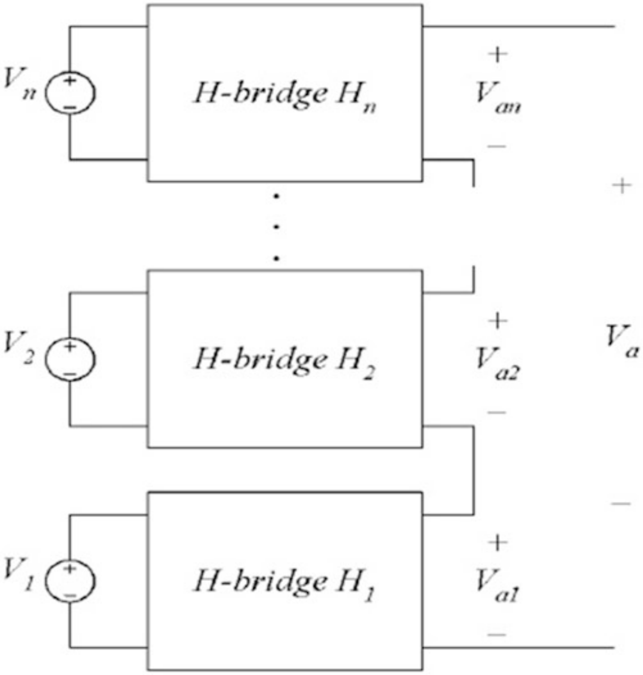


Fig. 3 The structure of a cascaded H-bridges converter

Table 3 Relationship between the number of voltage levels and the number of sources

Voltage levels (i = 1, 2, 3,…, n)	Number of levels S	Redundancy
Independent	3^n	0
$V_{a(i-1)} = V_{ai}$	2n + 1	3^n – (2n + 1)
$V_{a(i-1)} = 2V_{a(i-1)}$	$2^{n+1} - 1$	3^n – (2^{n+1} – 1)

When S_1 and S_4 are on, the output is V_{dc}; when S_2 and S_3 are on, the output is $-V_{dc}$; when either pair S_1 and S_2 or S_3 and S4 are on, the output is 0 [7]. Effective number of output voltage levels S depends on the ratio between the dc sources V_1 and V_2 as shown in Table 3. For example for a two-level converter (n = 2), by opening and closing the switches of H1 appropriately, the output voltage Vat can be made equal to $-V_1$, 0, or V_1 while the output voltage of H2 can be made equal to $-V_2$, 0, or V_2. Accordingly, the output voltage of the converter in different cases is shown in Table 4 [8].

Table 4 Output voltage of a multilevel converter with two cells

V_{a1}	V_{a2}	V_{dc}	V_{an} $(V_1 = V_2 = V_{dc})$	V_{an} $(V_1 and\ V_2 are\ Independent)$
V_1	V_2	V_{dc}	$2V_{dc}$	$V_1 + V_2$
V_1	0	V_{dc}	V_{dc}	V_1
V_1	$-V_2$	$V_{dc}/2$	0	$V_1 - V_2$
0	V_2	$V_{dc}/2$	V_{dc}	V_2
0	0	0	0	0
0	$-V_2$	$-V_{dc}/2$	$-V_{dc}$	$-V_2$
$-V_1$	V_2	$-V_{dc}/2$	0	$-V_1 + V_2$
$-V_1$	0	$-V_{dc}$	$-V_{dc}$	$-V_1$
$-V_1$	$-V_2$	$-3V_{dc}/2$	$-2V_{dc}$	$-V_1 - V_2$

5 Comparison of Multilevel Topologies

Usually to reduce power losses, unit size and costs we need to have a right choice of a multilevel topology for our application, and act of diminishing in the number of components can play the most important role in this circumstance. Hence, in advance it needs to prepare some guidelines to help us to selecting the applicable multilevel topology, Table 5 is showed the number of semiconductors and passive components demanded by the most favorable topologies. For instance, for a three-level approach, according to this analysis, all DCC, FCC and CHBC are needed the 12 switches; but, they are various in the number of clamping components and DC sources required. Because the cascaded H-bridge system needs a complicated transformer to prepare the different independent DC sources, therefore the DCC and FCC topologies have benefits over the CHBC system for applications where only one DC source is ready for use. On the other side when multiple DC sources are attainable, hence the CHBC topology can have advantages over the DCC and FCC system since it demands the smallest number of elements [9].

6 Power Converter Control

Nowadays as you will see in Fig. 4 there are some different control methods for the control of inverters and drives, Some of them are very well based and uncomplicated, as an example of the nonlinear hysteresis control, while on the other hand, newer control methods are usually more complicated or even need to have much more calculation power from the control part.

Hysteresis control Strategy mostly uses in uncomplicated applications such as current control, but also it can be used for more complex schemes like direct torque control (DTC) [10] and direct power control (DPC) [11]. Any linear controller

Table 5 Comparison of the multilevel topologies

Topology	Levels	Swithces[a]	Clamping diodes[b]	Floating capacitors	DC-link capacitors	Isolated DC sources
Diode clamped converter (DCC)	3	12	6	0	2	1
	5	24	36	0	4	1
	m	6(m-1)	3(m-1)(m-2)	0	m-1	1
Flying capacitor converter (FCC)	3	12	0	3	2	1
	5	24	0	18	4	1
	m	6(m-1)	0	$\frac{3}{2}$ (m-1)(m-2)	m-1	1
Symmetric cascaded hbridge converter (CHBC)	3	12	0	0	3	3
	5	24	0	0	6	6
	m	6(m-1)	0	0	$\frac{3}{2}$ (m-1)	$\frac{3}{2}$ (m-1)

[a] Switches with anti-parallel diodes
[b] Series connection for same blocking voltage stress

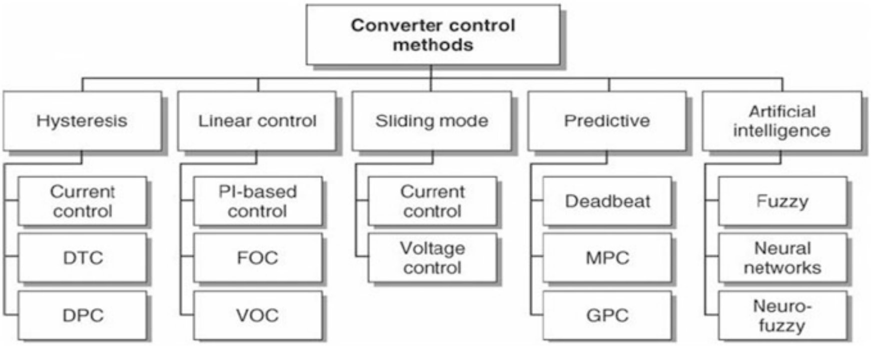


Fig. 4 Different types of converter control systems for power converters and drives (GPC = Generalized Predictive Control)

strategy can be used with the power converters, and proportional–integral (PI) controllers are the most ordinary choice for this purpose.

Power converter systems are major to several system restrictions and specialized requirements (total harmonic distortion (THD), maximum current, maximum switching frequency, etc.), which straight cannot be included into the linear controller scheme. In conclusion, to use classical control theory in modern digitally controlled converters, it has to been made suitable again and again in order [10, 12].

After evolvement of more powerful microprocessors, new control systems have been submitted. Fuzzy logic control, neural networks, sliding mode control, and predictive control are some of the most important ones. Among of them, predictive control strategy seems to be a very interesting option for the control of power converters and drives.

7 Conclusion

The aim of this paper has been to show and illustrate the multilevel converter topologies through examples or physical demonstrations. At the start, an introduction to power electronic converters was presented. And then the fundamentals and concepts of MLC structures were surveyed. Next, in connection with the number of components and isolated DC sources, there was shown a comparison of the most promising multilevel topologies. At the end, there was a brief explanation of different control strategies for power converters. Always one topology will be more suitable and proper than the others for any one special application; just we need to know which one is more appropriate for our PURPOSE. Often, topologies are picked dependent upon what has gone before, however, if that topology might not be the best option for the application.

References

1. Sowjanya T, Veerendranath K (2014) Cascaded H-Bridge with single DC source and regulated capacitor voltage. Int J Adv Sci Technol 73:89–102
2. Choi NS, Cho JG, Cho GH (1991) A general circuit topology of multilevel inverter. In: Proceeding IEEE PESC’91, pp 96–103
3. Franquelo LG, Rodríguez J (2008) The age of multilevel converters arrives. IEEE Trans Ind Electron June:28–39
4. Rodríguez J, Lai JS, Peng FZ (2002) Multilevel inverters: a survey of topologies, controls, and applications. IEEE Trans Ind Electron 49(4):724–738
5. Lai JS, Peng FZ (1996) Multilevel converters-a new breed of power converters. IEEE Trans Ind Appl 32(3):509–517
6. Tolbert L, Peng FZ, Habetler T (1999) Multilevel converters for large electric drives. IEEE Trans Ind Appl 35:36–44
7. KR Chakravarthy, Basha SG (2014) Int J Sci Eng Adv Technol, IJSEAT 2(1)
8. Liao J et al (2007) Cascaded H-bridge multilevel inverters—a reexamination. In: IEEE vehicle power and propulsion conference, institute of electrical and electronics engineers (IEEE)
9. Bai Z, Zhang Z, Zhang Y (2007) A Generalized three-phase multilevel current source inverter with carrier phase-shifted SPWM. In: Power electronics specialists conference, 2007. PESC 2007. IEEE, pp 2055–2060, 17–21 https://doi.org/10.1109/pesc.2007.4342322

10. Ohnishi T (1991) Three phase PWM converter/inverter by means of instantaneous active and reactive power control. In: Proceedings of the international conference on industrial electronics, control and instrumentation, 1991. Proceedings. IECON ‘91. vol 1, pp 819–824, October–November 1991
11. Blaschke F (1972) The principle of field-orientation applied to transvector closed-loop control system for rotating field machines. Siemens Rev XXXIX(5):217–219
12. Kazmierkowski M, Malesani L (1998) Current control techniques for threephase voltage-source pwm converters: a survey. IEEE Trans Industr Electron 45(5):691–703

Control Scheme of Micro Grid for Intentional Islanding Operation

Ronak Jahanshahi Bavandpour and Mohammad Masoudi

Abstract The cluster of multiple distributed generators (DGs) such as renewable energy sources that supply electrical energy are defined as micro grid. The DG is a voltage source inverter with an output low pass filter supplying the load. The connection of micro grid is in parallel with the main grid. When micro grid is isolated from the remainder of the utility system, it is said to be in intentional islanding mode. In this mode, DG inverter system operates in voltage control mode to provide constant voltage to the local load. During grid connected mode, the micro grid operates in constant current control mode to supply preset power to the main grid. An intentional islanding detection algorithm responsible for switching between current control and voltage control is developed using logical operations. The satisfactory performance of the micro grid with the proposed controllers and algorithms is analyzed by conducting simulation on dynamic model using MATLAB.

Keywords Distributed generation · Intentional islanding operation
Islanding detection · Micro grid · Electrochemical storage

1 Introduction

The Renewable Energy Sources-based DG systems are normally interfaced to the grid through power electronics (Inverter) and energy storage (Battery) systems [1]. Most critical section of the control system for a distributed generation (DG) unit's interconnection to the utility grid lies within the grid-connected converter's control

R. J. Bavandpour (✉)
Department of Electrical Engineering, Darolfonoon Higher
Education Institute, Qazvin, Iran
e-mail: lili66630@Gmail.com

M. Masoudi
Department Electrical, Biomedical and Mechatronics Engineering,
Islamic Azad University, Qazvin Branch, Qazvin, Iran

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_39

and protection system; specifically islanding detection algorithms. Through this controller subsection, the system is able to determine whether or not it is safe to remain connected to the grid. These islanding detection algorithms, which are integrated into the control system, are mainly present to prevent the undesirable feeding of loads during fault conditions and disconnections from the grid, whether or not the disconnection is intentional [2]. This is required by standards since the creation of such "power islands" is forbidden. Thus, in effect, standards require DG control systems to sense islanding events and disconnect themselves from the grid. Islanding is a condition in which a micro grid or a portion of the power grid, which contains both load and distributed generation (DG), is isolated from the remainder of the utility system and continues to operate. Some distinctions of islanding are: non-intentional islanding occurs if after the fault it is not possible to disconnect the DG; non-intentional islands must then be detected and eliminated as fast as possible; intentional islanding refers to the formation of islands of predetermined or variable extension; these islands have to be supplied from suitable sources able to guarantee acceptable voltage support and frequency, controllability and quality of the supply, and may play a significant role in assisting the service restoration process. Micro grids, seen as particular types of intentional islands, are basically operated in autonomous mode, not connected to the supply system; the whole micro grid can be seen from the distribution system as a single load and has to be designed to satisfy the local reliability requirements, in addition to other technical characteristics concerning frequency, voltage control and quality of supply [2].

In order to transfer from current control to voltage control mode, detection of transition from grid connected to intentional islanding mode is necessary. This is achieved by using an intentional islanding detection algorithm [3]. After islanding operation, the DGs are connected back to grid. At this instant of grid reconnection, re-closure algorithm has to be established to achieve synchronization [3, 4].

Parallel operation of distributed generation is an issue of high importance to a micro grid, which can provide a highly reliable electric supply service and good power quality to end customers when the utility is unavailable. However, there is a well-known limitation: the power sharing accuracy between distributed generators in a parallel operation.

The main contribution of this paper is summarized as:

(1) Design of a network-based control scheme for inverter-based sources, which provides proper current control during grid connected mode and voltage control during islanding mode.
(2) Development of an algorithm for intentional islanding detection and synchronization controller required during grid reconnection.
(3) Dynamic modeling and simulation are conducted to show system behavior under proposed method using SIMULINK.

The remainder of this paper is organized as follows. In Sect. 2, control techniques used in this paper including PLL structure, resynchronization algorithm, current and voltage controllers are described. In Sect. 3, the design process of current and voltage PI controllers and their block diagrams is provided. In Sect. 4, dynamic modeling and simulation results are presented. Finally, the paper is concluded in Sect. 5.

1.1 Control Techniques for Inverter

The voltage and current control loop has been implemented by using PI controllers working on the D-Q synchronous reference frame. AC quantities are converted into DC synchronous reference frame by Parks Transformation. Correspondingly, all reference quantities become DC in nature, so that simple PI controllers would be sufficient to yield zero steady state error.

2 DQ-PLL Structure

The phase angle and frequency at point of common coupling (PCC) is determined by using a DQ-PLL structure shown in Fig. 1. The PLL structure analogy comprises of a voltage controlled oscillator, an integrator and a phase detector [5, 6].

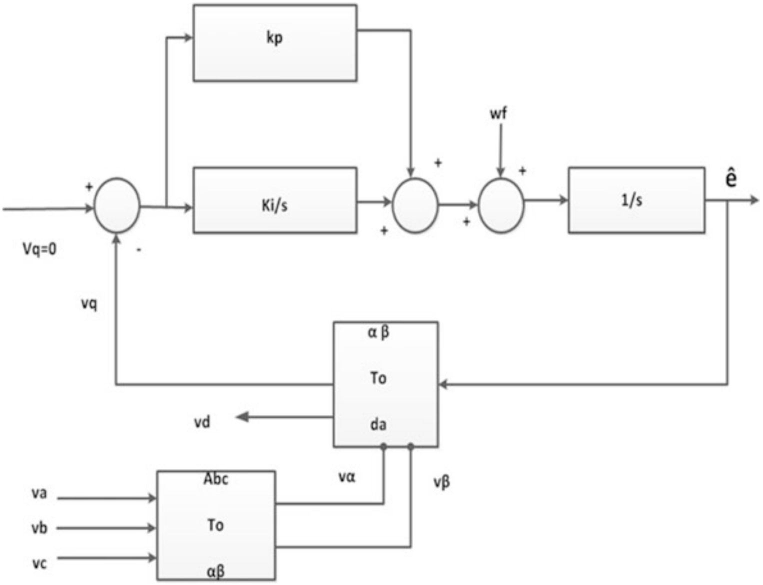


Fig. 1 DQ-PLL structure

Phase estimation is achieved by synchronization of the oscillating waveform generated by oscillator with the measured waveform.

The DQ-PLL structure consists of a Clarke's transformation, Park's transformation, PI regulator and an integrator. The realization of lock in PLL relies on regulating the quadrature component of rotating reference frame to zero using the PI controller.

3 Current Control

Current controller is designed to provide constant current output during grid connected operation [7, 8]. Control System shown in Fig. 2 is used to accomplish current control. In the strategy proposed here, the VSC line current is made controllable by a dedicated scheme and through the control of VSC terminal voltage. The inverter AC output current is transformed into DC quantity in synchronous rotating frame by Park's transformation. The direct and quadrature components are compared with the reference quantities and the error signal is passed to the PI controller to generate the voltage references [9]. The inverter terminal voltage is considered as a disturbance and hence fed forward to compensate it [1]. Finally, the DC reference quantities added with terminal voltages are transformed back to stationary frame by Inverse Park's Transformation. Thereafter it is used to generate the gate pulses by SPWM technique [3].

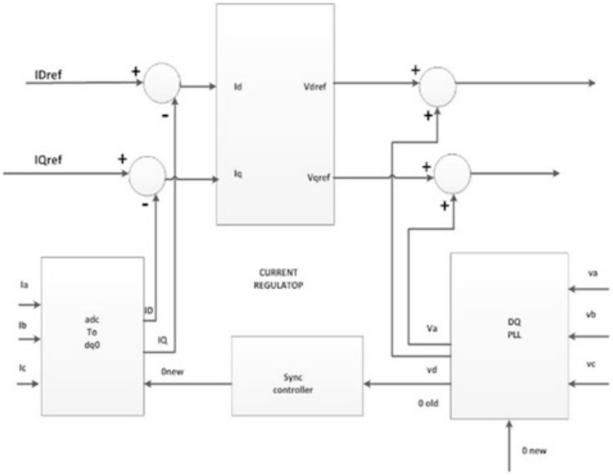


Fig. 2 Current controller

4 Voltage Control

This control scheme makes use of both current regulator as well as voltage regulator. The control works on the principle of voltage regulation through current compensation [7, 8]. Figure 3 represents the voltage control scheme. The converter output voltage is controlled by a synchronous reference frame closed loop voltage controller. Its output is transferred into a closed loop current regulator which is further transformed into stationary frame, and then the space vector PWM generates the gating signals of the IGBTs.

5 Intentional Islanding Detection Algorithm

The control works as voltage regulation through current compensation. The controller uses voltage compensators to generate current references for current regulation. As shown, the load voltages (Vd and Vq) are forced to track its reference by using a PI compensator (voltage regulator). The outputs of this compensator are compared with the load current, and the error is fed to a current regulator (PI controller). The output of the current compensator acts as the voltage reference signal that is fed to the pulse width modulator to generate the high frequency gating signals for driving the three-phase voltage source inverter. The current loop is included to stabilize the system and to improve the system dynamic response by rapidly compensating for near-future variations in the load voltages. In order to get a good dynamic response, VDQ is fed forward. This is done because the terminal voltage of the inverter is treated as a disturbance, and the feed forward is used to compensate for it.

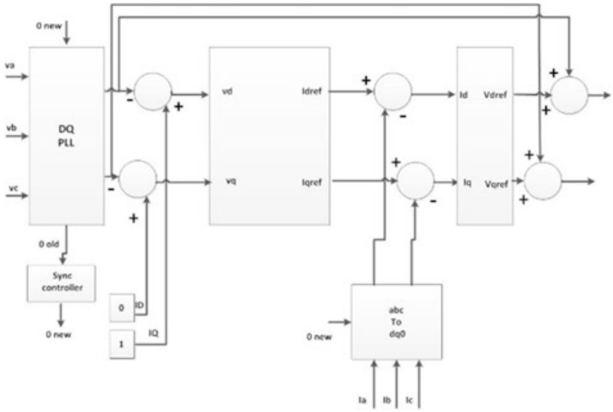


Fig. 3 Block diagram of voltage controller

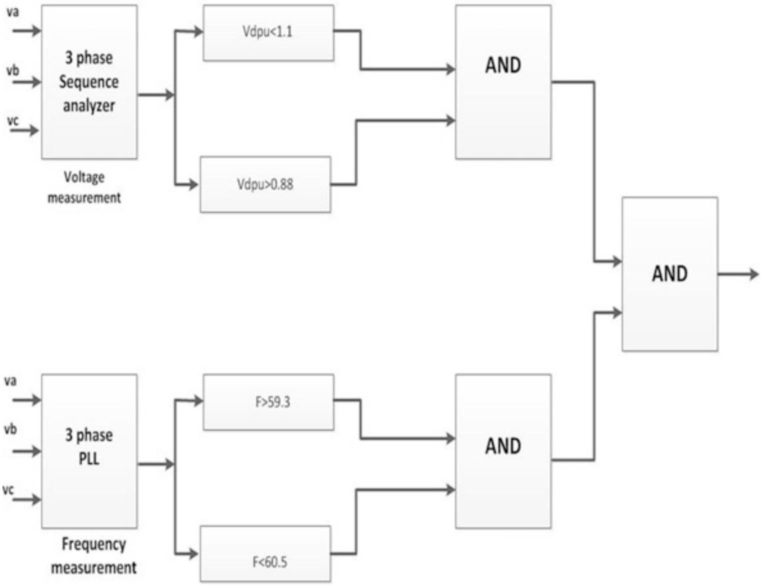


Fig. 4 Proposed algorithm for intentional islanding detection

Figure 4 indicates the algorithm developed to accomplish the detection of intentional islanding. Under deficient grid voltage conditions, the main switch is turned off and disconnects the main grid from the utility. This switching causes transients in voltage and frequency [10]. Therefore, tracking of system voltage magnitude and frequency indicates the transition (switching) between grid-connected and islanding mode, and vice versa. Voltage magnitude and frequency measurement is achieved with the help of three phase sequence analyzer and DQ-PLL [11]. According to this algorithm, values of frequency and voltage magnitude are constrained to a particular limit.

6 Resynchronization Controller Algorithm

When the grid-disconnection cause disappears, the transition from islanded to grid-connected mode can be started. To avoid hard transients in the reconnection, the DG has to be synchronized with the grid voltage. The DG is operated in the synchronous island mode until both systems are synchronized. Once the voltage in the DG is synchronized with the utility voltage, the DG is reconnected to the grid, and the controller will pass from the voltage to the current control mode. This synchronization is achieved by implementing the following algorithm [12].

When paralleling micro grid with utility grid, it is necessary to have the same phase angle for both of them. By closing the breaker at PCC, the two individual systems begin to have parallel operation. In order to achieve stiff synchronization with the utility grid during grid reconnection, synchronization controller is used. The algorithm used here [3] is to determine the new phase angle at which both the micro grid and utility grid have to operate.

Using the variables k and g, $\sin(\theta)$ can be found as:

$$\sin(\theta) = \frac{4/3g + 2/3k}{\sqrt{3}} \tag{1}$$

7 Design of PI Controller

7.1 Grey Wolf Optimization Algorithm

Grey wolf optimization algorithm is a new population-based algorithm which was introduced in 2014 [13]. Grey wolf (Canis lupus) belongs to Candia family. Of particular interest is that they have a very strict social dominant hierarchy as shown in Fig. 5. The leaders are called alphas. The alpha is mostly responsible for making decisions about hunting, sleeping place, time to wake, and so on. The second level in the hierarchy of grey wolves is beta. The betas are subordinate wolves that help the alpha in decision-making or other pack activities. The beta wolf is probably the best candidate to be the alpha in case one of the alpha wolves passes away or becomes very old. Omega is the lowest ranking grey wolf. The omega wolves play the role of scapegoat. Omega wolves always have to submit to all the other dominant wolves. They are the last wolves that are allowed to eat. If a wolf is not an alpha, beta, or omega, he/she is called subordinate (or delta in some references). Delta wolves have to submit to alphas and betas, but they dominate the omega. Scouts, sentinels, elders, hunters, and caretakers belong to this category. Figure 5 shows the caption for a figure that must follow the figure.

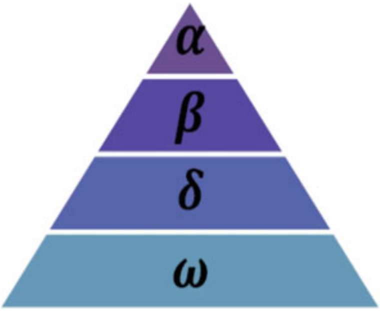


Fig. 5 The caption for a figure must follow the figure

According to [14] the main phases of grey wolf hunting are as follows:

- Tracking, chasing, and approaching the prey.
- Pursuing, encircling, and harassing the prey until it stops moving.
- Attack towards the prey.

In order to mathematically model the social hierarchy of wolves when designing GWO, we consider the fittest solution as the alpha (α). Consequently, the second and third best solutions are named beta (β) and delta (δ) respectively. The rest of the candidate solutions are assumed to be omega (ω).

Grey wolves encircle prey during the hunt. In order to mathematically model encircling behavior the following equations are proposed:

$$\vec{D} = \left|\vec{C}.\vec{X}_P(t) - \vec{X}(t)\right| \tag{2}$$

$$\vec{X}(t+1) = \vec{X}_P(t) - \vec{A}\vec{D} \tag{3}$$

where t indicates the current iteration, $\vec{A}$ and $\vec{C}$ are coefficient vectors, $\vec{X}_P$ is the position vector of the prey, and $\vec{X}$ indicates the position vector of a grey wolf.

The vectors $\vec{A}$ and $\vec{C}$ are calculated as follows:

$$\vec{A} = 2\vec{a}\vec{r}_1 - \vec{a} \tag{4}$$

$$\vec{C} = 2\vec{r}_2 \tag{5}$$

where components of $\vec{a}$ are linearly decreased from 2 to 0 over the course of iterations and r_1, r_2 are random vectors in [0, 1]. Grey wolves have the ability to recognize the location of prey and encircle them. In order to mathematically simulate the hunting behavior of grey wolves, we suppose that the alpha (best candidate solution), beta, and delta have better knowledge about the potential location of prey. Therefore, we save the first three best solutions obtained so far and oblige the other search agents (including the omegas) to update their positions according to the position of the best search agents. The following formulas are proposed in this regard.

$$\vec{D}_\alpha = \left|\vec{C}_1\vec{X}_\alpha - \vec{X}\right|,\ \vec{D}_\beta = \left|\vec{C}_2\vec{X}_\phi - \vec{X}\right|,\ \vec{D}_\delta = \left|\vec{C}_3\vec{X}_\delta - \vec{X}\right| \tag{6}$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{a}_1\vec{D}_\alpha\,,\ \vec{X}_2 = \vec{X}_\beta - \vec{a}_2\vec{D}_\beta\,,\ \vec{X}_3 = \vec{X}_\delta - \vec{a}_3\vec{D}_\delta \tag{7}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{8}$$

The final position would be in a random position within a circle which is defined by the positions of alpha, beta, and delta in the search space. In other words alpha, beta, and delta estimate the victim position and other wolves update their positions

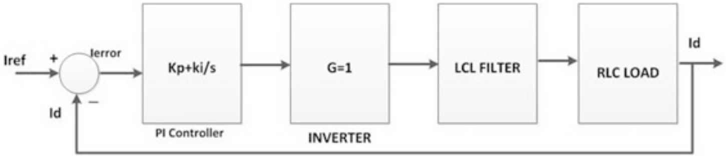


Fig. 6 Block diagram of current controlled inverter

randomly around the victim [15]. Pseudo code of the GWO algorithm is shown in its simplest form in Fig. 6.

8 Current Control Transfer Function

Current control scheme is needed during grid connected operation. The block diagram of current controlled operation is shown in Fig. 6. The inverter is modeled with an ideal gain *GI* = 1. In order to obtain the transfer function of the filter and the load block, the designed LCL filter and RLC load circuit must be taken into consideration. Figure 6 depicts the circuit diagram of LCL filter and RLC load. The transfer function of the LCL filter and the RLC load also remains the same for voltage control transfer function.

Correspondingly transfer function of PI controller block is given as

$$C(S) = K_p + \frac{K_I}{S} \tag{9}$$

9 Voltage Control Transfer Function

Voltage control scheme is required during islanding mode of operation. Figure 8 represents the block diagram of voltage controlled operation.

This scheme consists of an inner current control loop and an outer voltage control loop. The transfer function of individual blocks, such as PI controller, inverter, filter and load remains the same as in current control transfer function.

10 Current Control and Voltage Control Stability

The control method used here has two operating modes, current control and voltage control corresponding to grid-connected and intentional islanding operations of micro grid. The stability of the current and voltage controller can be determined by using their transfer functions [16]. Stability analysis is carried out by using the conventional control theory. According to it, the bode plot of the controller transfer function is plotted using the SISO design tool in MATLAB (Fig. 9). The positive Gain margin in the bode plots of both the transfer functions of corresponding controllers indicates that the system is stable.

11 Dynamic Modeling and Simulation Results

11.1 Simulation Results

To investigate microgrid operational modes, the effect of designed current controller, voltage controller, proposed intentional islanding detection algorithm and re-closure algorithm, the MATLAB/SIMULINK is used to develop a time domain simulation model of the study system. Electrical power system components are simulated with a physical modeling product called simpower systems supported by MATLAB. The simulations have been run with the dynamic model shown in Figure 10 to investigate the behavior of grid-connected and intentional islanding mode of operation. Inside the current and voltage regulator blocks, there exist the schemes shown in Figs. 7 and 8. Similarly, inside islanding detection algorithm and resynchronization controller block, there exists the circuit corresponding to Fig. 9.

For both cases the parameters used for simulation are given in the Table 1 [3]. In the dynamic model depicted here, the inverter is connected through a filter and a circuit breaker to the local load. The simulations are conducted here with the assumption that the irradiation variations are completely absent in this system. Two case studies based on the influence of synchronization controller are conducted to examine the system performance during grid-connected and intentional islanding mode.

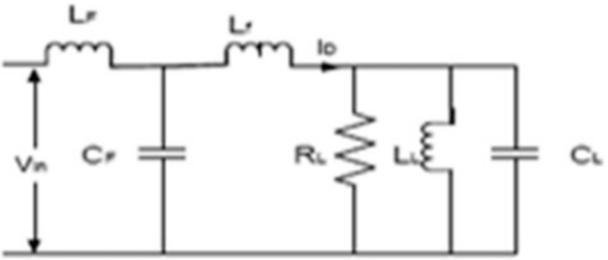


Fig. 7 LCL filter and parallel RLC load circuit

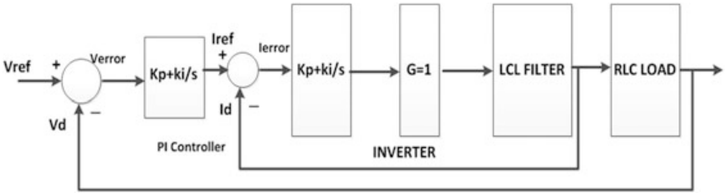


Fig. 8 Block diagram of voltage controlled inverter

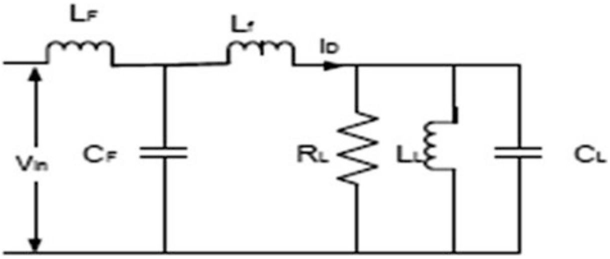


Fig. 9 The bode plot of the controller transfer function

Table 1 Designed values of parameters used for simulation

Parameters	Values
L_f	1 mH
C_F	31 μF
L_F	0.5 mH
C_L	1.535 mF
L_L	4.58 mH
R_L	4.33 Ώ
V_{dc}	400 V
R_{grid}	0.1 Ώ
L_{grid}	0.1 mH
F	60 Hz
f_s	1 kHz
V_o	120 V
k_p, k_i	1.24, 0.02
k_p, k_i	0.8, 50

11.2 Without and with Synchronization Controller

At first, simulation was conducted when the micro grid is connected to the utility grid without any synchronization controller. The grid was disconnected by setting the status in three phase circuit breaker as closed initially and transition time at 0.3 s. The occurrence of transients at 0.3 s is seen clearly. Occurrence of large transient during the instant of grid reconnection is undesirable. The grid was then reconnected at the second of 0.6. The micro grid continues to operate in synchronous islanding mode until both the utility system and the micro grid system is resynchronized. From the beginning of intentional islanding mode, the system achieves synchronization in much less time comparing with that system without synchronization controller and with the implementation of resynchronization controller. Algorithm, the DG voltage is forced to track the voltage at the grid. When the synchronization is completed, the DG is reconnected to the grid, and the controller will be switched from the voltage to the current control mode.

11.3 Conclusion

In this paper current and voltage control techniques have been developed for grid-connected and intentional islanding modes of operation using PI controllers. An intentional islanding detection algorithm responsible for switching between current control and voltage control is developed using logical operations and proved to be effective. The reconnection algorithm coupled with the synchronization controller enable the DG to synchronize itself with the grid during grid reconnection. The performance of the micro grid with the proposed controllers and algorithms is analyzed by conducting simulation on dynamic model using SIMULINK. The simulation results show the effectiveness of the control scheme. Moreover, we also develop a model for optimizing the daily operational planning of an interconnected micro grid considering electrochemical storage.

References

1. Hassanzadeh I, Alizadeh G, Shirjoposht NP, Hashemzadeh F (2010) A new optimal nonlinear approach to half car active suspension control. IACSIT Int J Eng Technol 2:78–84
2. Ata AB, Kunya AA (2015) Half car suspension system integrated with PID controller (Presented conference style). In: presented at the 29th European conference on modelling and simulation, May. 2015
3. Brezas P, Smith MC (2014) Linear quadratic optimal and risk-sensitive control for vehicle active suspensions. IEEE Trans Control Sys Technology 22:543–556
4. Gupta S, Ginoya D, Shendge P, Phadke SB (2015) An inertial delay observer-based sliding mode control for active suspension systems. Proc Inst Mechanical Eng Part D: J Automobile Eng 230:352–370
5. Li H (2012) Robust control design for vehicle active suspension systems with uncertainty. PhD dissertation, University of Portsmouth
6. Wang R, Jing H, Karimi HR, Chen N (2015) Robust fault-tolerant H∞ control of active suspension systems with finite-frequency constraint. Mech Syst Signal Process 62:341–355
7. Feng J, Matthews C, Zheng S, Yu F, Gao D (2015) Hierarchical control strategy for active hydro pneumatic suspension vehicles based on genetic algorithms. Adv Mechanical Eng 7 (2):951050

8. Naim MS, Basir M (2012) Modeling and controller design for an active car suspension system using quater car model
9. Mouleeswaran S (2012) Design and development of PID controller-based active suspension system for automobiles. INTECH Open Access Publisher, pp 71–98
10. Zhao F, Dong M, Qin Y, GU L, Guan J (2015) Adaptive neural networks control for camera stabilization with active suspension system. Adv Mechanical Eng 7(8):1–11
11. Sam YM, Osman JHSB (2005) Modeling and control of the active suspension system using proportional integral sliding mode approach. Asian J Control 7:91–98
12. Sun W, Gao H, Kaynak O (2013) Adaptive backstopping control for active suspension systems with hard constraints. IEEE/ASME Trans Mechatronics 18:1072–1079
13. Aboud WS, Haris SM, Yaacob Y (2014) Advances in the control of mechatronic suspension systems. J Zhejiang Univ Sci C 15:848–860
14. Ang KH, Chong G, Li Y (2005) PID control system analysis, design, and technology. IEEE Trans Control Systems Technol 13(4):559–576
15. Gerla G (2005) Fuzzy logic programming and fuzzy control. Studia Logica 79(2):231–254
16. Talib A, Hussin M, Darns M, Zaurah I (2013) Self-tuning PID controller for active suspension system with hydraulic actuator. In: IEEE Symposium on, IEEE, Computers and Informatics (ISCI), pp 86–91, April 7–9, 2013

Quasi-3D Analytical Prediction for Open Circuit Magnetic Field of Axial Flux Permanent-Magnet Machine

Amir Hossein Sharifi, Seyed Mehdi Seyedi and Amin Saeidi Mobarakeh

Abstract Precise simulation of magnetic field in the air gap is necessary to Predict of electromagnetic performance of permanent magnet machines. In this paper an analytical approach is proposed to calculate the magnetic field of slot-less Axial flux permanent magnet machine. Finite element method is the most precise method for magnetic calculation in the air gap although because of its low speed calculation is not appropriate for parametric and optimization studies. In this paper, an efficient analytical method is used to parametric studies on magnetic field in the air gap in case of no-load operation. By applying above mentioned method with high analytical speed, flux density in the middle of air gap is predicted and then induced voltage is calculated. Result of proposed analytical method is compared with 3-D FEM simulations that are reached from Ansys-Electronics software. These results are in the form of Quasi-3D.

Keywords Axial flux permanent magnet machine · Analytical method
Magnetic flux density · Induced voltage

1 Introduction

AFPM is an appropriate alternative for radial-flux permanent magnet machine Because of some features such as high flux density, compressed structure and their physical shape (like a disc). These kinds of machines are used in some industrial

A. H. Sharifi (✉)
Department of Electrical Engineering, Shahed University, Tehran, Iran
e-mail: amir_822796@yahoo.com

S. M. Seyedi · A. S. Mobarakeh
Department of Electrical Engineering, Iran University of Science and Technology (IUST), Tehran, Iran
e-mail: mehdiseyedi69@gmail.com

A. S. Mobarakeh
e-mail: amin.saeidi72@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_40

applications such as electrical vehicles, electrical pumps and fans. Also, axial-flux machines are used in low (and middle) power generator. Because of large diameter of AFPMs, it can be placed many poles on rotor, so they are appropriate for low-speed applications such as transportation drives, lifters and wind turbines [1]. AFPMs contain varied structures because of their disc shape of stator and rotor so they can be designed with an air gap or multi air gaps, slot-less or with slots and also with armature without ferromagnetic material and some else forms. They are described in [1–4] in detail.

Precise simulation of magnetic field in the air gap is necessary to Predict of electromagnetic performance of permanent magnet machines. The related models are based on two different methods: numerical and analytical methods. In numerical one, detailed structure of machine and some non-linear effects(such as saturation) can be considered although it needs so many times for optimization studies. Against of that, analytical methods are very fast and the results are in the form of Fourier series. These are some benefits of analytical methods that make them appropriate to apply in some studies.

One-dimensional analytical model can only calculate vertical component of flux density in air gap and the other components are ignored. This model is used to analyze electromagnetic field in induction motors, synchronous machine with electric excitation and variable reluctance machines [5]. This model is applied for prediction of cogging-torque [6], distribution of radial-forces [7], unbalanced magnetic forces [8] and some other main feature of machine electromagnetic performances.

Since the permittivity of the air and magnet are the same approximately in surface mounted permanent magnet machines, the air gap is larger than the other machines. So the amount of inter-pole leakage flux is high. In this case, radial flux cannot be ignored so two-dimensional analytical method is necessary to apply. In [9] flux density of the air gap is calculated with direct solving of Maxwell equations in Cartesian coordinate system. In [10], this model is generalized to cylindrical coordinate system. Another group of analytical methods are based on direct solving of equations that are related to different areas with satisfying boundary conditions to reach magnetic field of slot-less machines. An analytical method based on this model is used in [9]. In [11] an analytical method is applied for permanent magnet machine with partial magnet. In [12] an analytical method in case of open-circuit and armature reaction field calculation considering permittivity of magnet is proposed. This model is generalized to permanent magnet machines with parallel magnetization in [13]. On the basis of this model, prediction of Eddy Current loss in magnet is calculated in [14], also in [15, 16] harmonics of radial force is considered and different designs of brushless permanent magnet machines are proposed in [16]. In [15] an analytical method considering finite magnet permittivity is proposed by solving equations in stator, air gap, magnet and rotor. A different two-dimensional analytical method on the basis of subdomains approach for internal magnet machines is proposed in [17]. In order to implement slot-effect there are some proposed approach in [18–20].

Because of the structure of AFPMs, magnetic field intensity is 3-D in nature. So, we have also radial-flux in addition to axial-flux and tangential-flux in the air gap. Three-dimensional simulation in axial flux machines is developed in different approaches. In [20] AFPM is modeled by combined multi linear machines (Quasi-3D approach). This model is used to calculate induced voltage in the case of no-load in [21]. Another approach is solving Maxwell equations with separation of variables method in 3-D space [22].

2 Formulation and Proposed Method

Precise simulation of magnetic field in the air gap is necessary In order to predict AFPM electromagnetic performance such as electric motive force (EMF) waveform, torque waveform, effect of cogging torque, rotor and stator iron losses, Eddy current loss and etc. in this way there are analytical and numerical models. In this paper, an efficient analytical method based on two-dimensional Maxwell equation in case of no-load operation is proposed. Applying Quasi-3D approach can model 3-D nature of electromagnetic field. Also, this model can calculate induced voltage in armature coil.

As mentioned before, analytical methods are preferred in parametric studies because of their closed form results. In this section, the proposed analytical approach is introduced to predict flux density in AFPM machine air gap. This model is based on Maxwell differential equations that are solved by separation of variables method.

2.1 Assumptions

In order to simplification of solving of Maxwell equations by separation of variables method in case of no-load, some assumptions are considered:

- Relative permeability of stator and rotor are assumed to be infinite.
- The stator is slot-less and magnet is surface mounted.
- Magnetization curve is linear and is placed in second area
- Implementing Quasi-3D approach, so, radial flux and end winding effect is ignored.
- Surface magnets permeability is invariable, in the direction of z axis. Also they are in rectangle shape in each ring.
- Coil permeability is assumed to be 1.
- Coil cross section in each ring is in rectangle form and current distribution in coil is constant as it has been shown in Eq. 1.

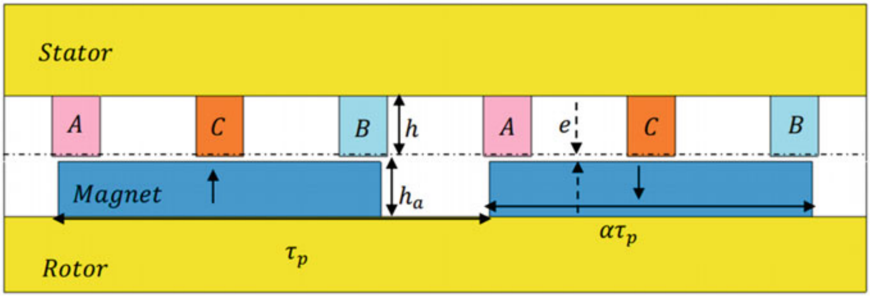


Fig. 1 Two-dimensional model of AFPM machine

$$J(\varphi, z) = J_0 a_r \tag{1}$$

The graphical structure of machine and simulation parameters is shown in Fig. 1.

2.2 *Calculation of Field Caused by Magnet*

The electromagnetic field caused by magnet is calculated in two regions according to Fig. 2. These regions can be denoted as Eq. 2. This will be reached by ignoring coils.

$$\begin{array}{l}\text{Region I (PM): } 0 < \text{z} < h_a \\ \text{Region II (air gap): } h_a < z < \gamma\end{array} \tag{2}$$

Corresponding differential equations in each region is shown in Eq. 3

$$\begin{cases}\nabla^2 A_{\text{II}} = \mu_0 \nabla \times M & \text{in magnet region} \\ \nabla^2 A_{\text{I}} = 0 & \text{in airgap region}\end{cases} \tag{3}$$

In Eq. 3, A represents magnetic potential vector and M represents magnet permeability vector. Magnet permeability is assumed to be invariable in direction of z-axis and it is shown Fig. 3. The Fourier series can be altered as Eq. 4.

$$\begin{aligned} M_z(z, \phi, \theta_r) &= \sum_{n=1,3,\ldots} M_n \cos(np(\phi - \theta_r)) \\ M_n &= \frac{4B_r}{\mu_0 n\pi} \sin(\frac{n\pi\alpha}{2}) \end{aligned} \tag{4}$$

In above equations B_r represents PM residual flux density and α is the proportion of magnet arc length to pole pitch(τ_p).

Boundary conditions in the air gap and magnet can be defined as follow:

$$\begin{aligned}
&\frac{\partial A_{\mathrm{I}}}{\partial z}\bigg|_{z=0} = 0\\
&\frac{\partial A_{\mathrm{II}}}{\partial z}\bigg|_{z=\gamma} = 0\\
&\frac{\partial A_{\mathrm{I}}}{\partial \varphi}\bigg|_{z=h_a} = \frac{\partial A_{\mathrm{II}}}{\partial \varphi}\bigg|_{z=h_a}
\end{aligned} \tag{5}$$

γ Is defined as follow (Eq. 6).

$$\gamma = e + h_a + h \tag{6}$$

General form of magnetic potential vector in two regions is obtained by solving differential equation with separation of variable method. The answer can be denoted as Eq. 7.

$$\begin{aligned}
A_{\mathrm{I}}(\varphi, z) &= \sum_{n=1,3} \frac{r}{nK}[A_{n\mathrm{I}}\sinh(nk\phi) + B_{n\mathrm{I}}\cosh(nk\phi)]\left[C_{n\mathrm{I}}\cos(\frac{nkz}{r}) + D_{n\mathrm{I}}\sin(\frac{nkz}{r})\right] + A_{p\mathrm{I}}(\varphi, z)\\
A_{\mathrm{II}}(\varphi, z) &= \sum_{n=1,3} \frac{r}{nk}[A_{n\mathrm{II}}\sinh(nk\phi) + B_{n\mathrm{II}}\cosh(nk\varphi)]\left[C_{n\mathrm{II}}\cos(\frac{nkz}{r}) + D_{n\mathrm{II}}\sin(\frac{nkz}{r})\right]
\end{aligned} \tag{7}$$

In above equations, $A_p(\varphi, z)$ is private answer in the first region and it can be calculated as Eq. 8.

$$A_p(\varphi, z) = -\frac{\mu_0 r m_n}{np}\sin(np(\varphi - \theta_r)) \tag{8}$$

In these equations, P and θ_r represent number of pole pairs and rotor movement, respectively. r is the radius of the calculating field point. Unknown coefficients corresponding to hyperbolic and trigonometry terms is calculated by applying boundary conditions. Magnetic potential vector is obtained as Eqs. 9 and 10. β_n is defined in Eq. 11.

$$A_{\mathrm{I}}(\varphi, z, \theta_r) = -\sum \beta_n \cosh(\frac{np(z-\gamma)}{r})\sin(np(\phi - \theta_r)) - \frac{\mu_0 r m_n}{np}\sin(np(\varphi - \theta_r)) \tag{9}$$

$$A_{\mathrm{II}}(\varphi, z, \theta_r) = -\sum_{n=1,3} \beta_n \frac{\sinh(\frac{nph_a}{r})}{\sinh(\frac{np(h+e)}{r})}\cosh(\frac{np(z-\gamma)}{r})\sin(np(\phi - \theta_r)) \tag{10}$$

$$\beta_n = \frac{\mu_0 M_n r}{(\cos(\frac{nph_a}{r}) + \sinh(\frac{nph_a}{r})\coth(\frac{np(h+e)}{r}))np} \tag{11}$$

Magnetic field in air gap and magnet can be altered as Eqs. 12–15. α_n is defined as Eq. 16.

$$B_{\phi PM\mathrm{I}}(\varphi, z, \theta_r) = -\sum \alpha_n \sinh(\frac{npz}{r})\sin(np(\phi - \theta_r)) \tag{12}$$

$$B_{zPM\mathrm{I}}(\varphi, z, \theta_r) = -\sum \alpha_n \cosh(\frac{npz}{r})\cos(np(\phi - \theta_r)) - \mu_0 M_n \cos(np(\phi - \theta_r)) \tag{13}$$

$$B_{\phi PM\mathrm{II}}(\varphi, z, \theta_r) = -\sum \alpha_n \frac{\sinh(\frac{nph_a}{r})}{\sinh(\frac{np(h+e)}{r})}\sinh(\frac{np(z-\gamma)}{r})\sin(np(\phi - \theta_r)) \tag{14}$$

$$B_{zPM\mathrm{II}}(\varphi, z, \theta_r) = -\sum_{n=1,3} \alpha_n \frac{\sinh(\frac{nph_a}{r})}{\sinh(\frac{np(h+e)}{r})}\cosh(\frac{np(z-\gamma)}{r})\cos(np(\phi - \theta_r)) \tag{15}$$

$$\alpha_n = \frac{np}{r}\beta_n \tag{16}$$

2.2.1 Calculation of No-Load Voltage

In this paper, Coil distribution function method is used In order to calculate no-load voltage. Coil distribution function Fourier series of phase A is shown in Eq. 17. $E_k^s(n)$ is defined in Eq. 18. It is necessary to note that E_0^s has no effect on amount of flux linkage. Also, coil distribution is full pitch and one-layer.

$$F_{Dc}^s(\varphi) = E_0^s + \sum_{n=1,3} E_k^s(n)\cos(n\,p\,\varphi) \tag{17}$$

$$E_k^s(n) = \left[\begin{array}{c} \frac{2}{n\pi}\sin\left(\frac{n\pi}{2} - \frac{n\pi\gamma_{so}\rho}{2\pi}\right) + 2\left(1 + \left(\frac{\pi - \gamma_{so}\rho}{2\gamma_{so}P}\right)\right)\sin(np\varphi) \\ -\frac{2}{n\pi\gamma_{so}}\varphi\,\sin(np\varphi) - \frac{2}{n^2 p\pi\gamma_{so}}\cos(np\varphi) \end{array}\right]_{\varphi=\frac{\pi-\gamma_{so}P}{2p}}^{\varphi=\frac{\pi+\gamma_{so}P}{2p}} \tag{18}$$

The magnitude of flux linkage of phase A by substitution of Eq. 15 in total flux linkage of c coils is obtained. No-load induced voltage in phase A is expressed in Eq. 20.

$$\phi_c = N_t \sum_{s=1}^{N_c} \frac{R_{os}^2 - R_{is}^2}{2} \int_0^{2\pi} F_{Dc}^s(\varphi) B_{zPM\mathrm{II}}(\varphi, \gamma - h/2, \theta_r) \Bigg|_{r=\frac{R_{os}-R_{is}}{2}} d\varphi \tag{19}$$

$$E = N_t \omega \sum_{s=1}^{N_f} \frac{R_{os}^2 - R_{is}^2}{4} \left[\sum_{n=1,3} np\alpha(n) E_k^s(n) \frac{\sinh\left(\frac{2nph_a}{R_{os}+R_{is}}\right)}{\sinh\left(\frac{2np(h+e)}{R_{os}+R_{is}}\right)} \cosh\left(\frac{nph}{R_{os}+R_{is}}\right) \sin(np\omega t) \right] \tag{20}$$

In above equations, N_t is number of coil series of phase A. and ω is angular velocity (rad/sec). R_{os} and R_{is} are external and internal radius of layer S.

3 Simulation and Results

The machine in last section calculation had one-stator and one-rotor. Although it can be generalized to different kinds of AFPM machines. In this section, the accuracy of analytical proposed method is evaluated by comparing with the results of numerical method. As mentioned before, the proposed model is two-dimensional but it converts into 3-D space with usage of Quasi-3D method. In order to evaluation of proposed model, magnetic field different components that are reached base on analytical method, are compared with the results of numerical method reached from Maxwell software in 3-D space.

3.1 Parameters and Machine Structure

Simulated machine in Maxwell software is AFPM machine with surface magnet. It has one stator and two external rotors. Three dimensional structures of stator and rotor are shown in Fig. 4. Cross sections of all coils are in the form of rectangle. Two sides of each magnet are in radial direction so that proportion of magnet arc length to pole pitch is constant. Assume that magnet arc length (α) is 60°. All parameters of machine are shown in Table 1.

3.2 Magnetic Field in Case of No-Load Operation

In order to calculate the air gap field caused by magnet, current of the coils are assumed to be zero. Figures 5 and 6 show axial-flux and tangential flux of the middle of air gap in case of r = 60 mm, respectively. It is necessary to note that

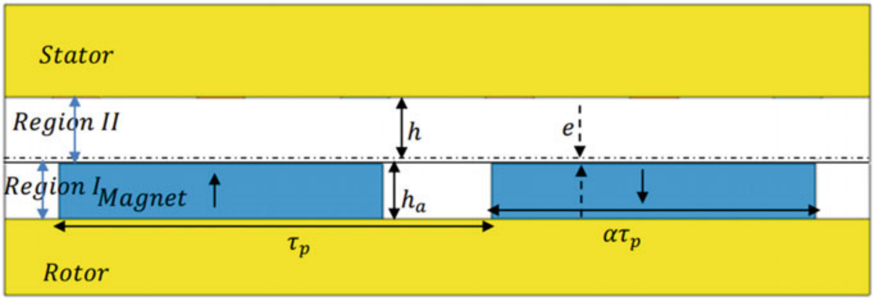


Fig. 2 Two dimensional model of AFPM in order to no-load simulation

Table 1 Parameters of simulation

Electrical parameter	Quantity	Electrical parameter	Quantity	Electrical parameter	Quantity
$\boldsymbol{N}$	30	R_o	80 mm	p	3
h	3 mm	e	1 mm	Q_s	18
γ_{so}	9 mm	B_r	1.21T	R_i	40 mm

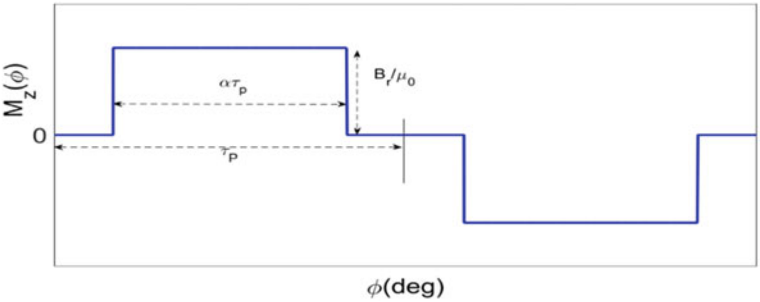


Fig. 3 Magnet permeability

maximum number of considered harmonic in previous analytical equations is 200. Distribution of axial-flux in the magnet where z = 2 mm is shown in Fig. 5a. as the figure shows there are a good adaption in results of analytical and numerical methods. Although in such upper points in curve that are related to edges of magnets, there are some incompatibilities in vertical component of air gap field. The reason is neglecting of high-order harmonics in aforementioned analytical equations. Also, there are some errors in the results of FEM software because of sudden change of the field and flux density is affected by number of meshes in these areas. Since maximum order of harmonics effect on converging above formula, so it causes such these small errors. Figure 6a shows tangential flux density in case of z = 2 mm. They are adapted with numerical results, too. Axial-flux and

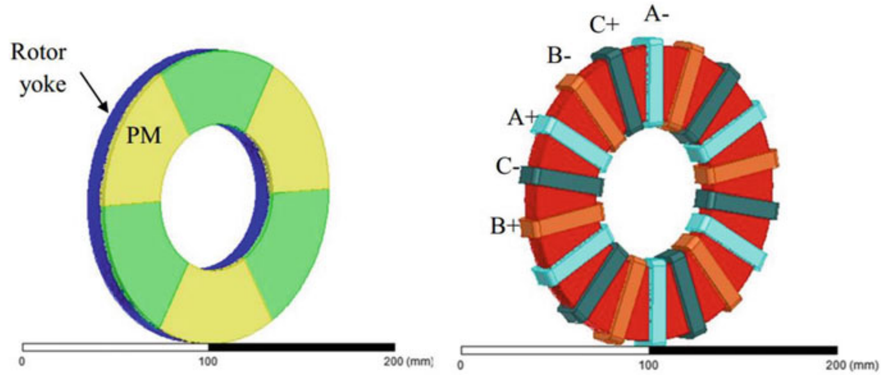
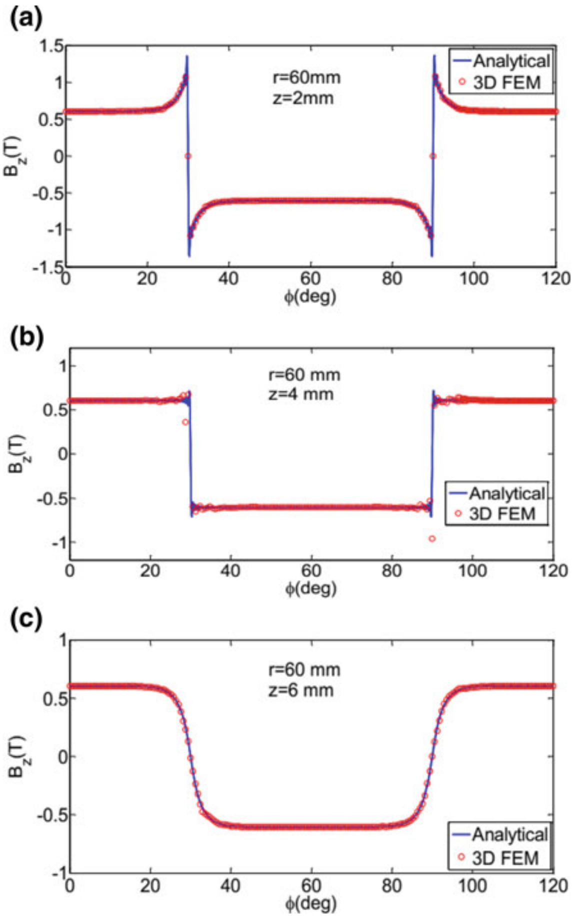


Fig. 4 Three dimensional structure of simulated stator and rotor

tangential-flux in the air gap in case of z = 4 mm are shown in Figs. 5b and 6b, respectively. As it's clear from the figure, the amount of overshoot in the curve in vertical component of field has been decreased as it comes close to stator. Although the amount of overshoot points corresponding to 30° and 90° in axial-flux curve that is related to inter-pole leakage flux, has increased. Figures 5c and 6c represent axial-flux and tangential-flux in coil region in case of z = 6 mm, respectively. In order to more analysis, axial and tangential calculated flux in case of r = 45 mm are shown in Figs. 7 and 8. Predicted flux that are reached from analytical method and numerical results that are reached from FEM software, are the same mostly. Because of leakage-flux in internal and external radius and ignoring of radial-flux in analytical approach, there is a little incompatibility in results that increases with approaching to rotor. Although, It is small and ignorable. Figure 7c shows them clearly.

Ignoring radial-flux causes some errors in analytical proposed approach, in order to analysis of these errors, axial-flux distribution in radius direction in the middle of magnets has been shown by Fig. 9 in cases of two different heights. Assume that axial-flux distribution in radius direction is in form of a rectangle function. Actually, we can assume that axial-flux in radial direction between internal radius to external one is constant (at a certain angle and height) by ignoring radial-flux and leakage-flux and its zero out of this interval. Distribution of axial-flux in internal and external radius is different from the middle of air gap and it depends on considered height. As it shows in Fig. 9 at this interval (between internal radius and external radius of machine) axial–flux is constant and it can be like a rectangle function, approximately. Figure 10 indicates radial flux in radial-direction of two different heights. This is equal to zero in the most regions of mentioned interval but in internal (or external) radius is non-zero.

Fig. 5 Axial flux in different points of air gap, r = 60 mm **a** z = 2 mm **b** z = 4 mm **c** z = 6 mm



3.3 *Induced Voltage in Armature Coil in Case of No Load Operation*

In this section, in order to accuracy evaluation of proposed analytical method, induced voltage is calculated and it's compared with the results of FEM 3-D simulation. It is necessary to note that the equations corresponding to induced voltage in last section was related to one-stator, one-rotor AFPM machine. So they must be multiplied by two. Also, the quantity of machine layers is assumed to be 10 layers in all quasi-3D method and simulations and maximum order of harmonics is 200. Figure 11 indicates no-load induced voltage in armature coil. The bold-line curves are related to analytical method and dash-line curves are reached from 3-D FEM simulation. These are the same mostly although results of analytical method are a little bigger than numerical method results (about 4%). The main reason of these small differences is that Quasi-3D method cannot model 3-D nature of

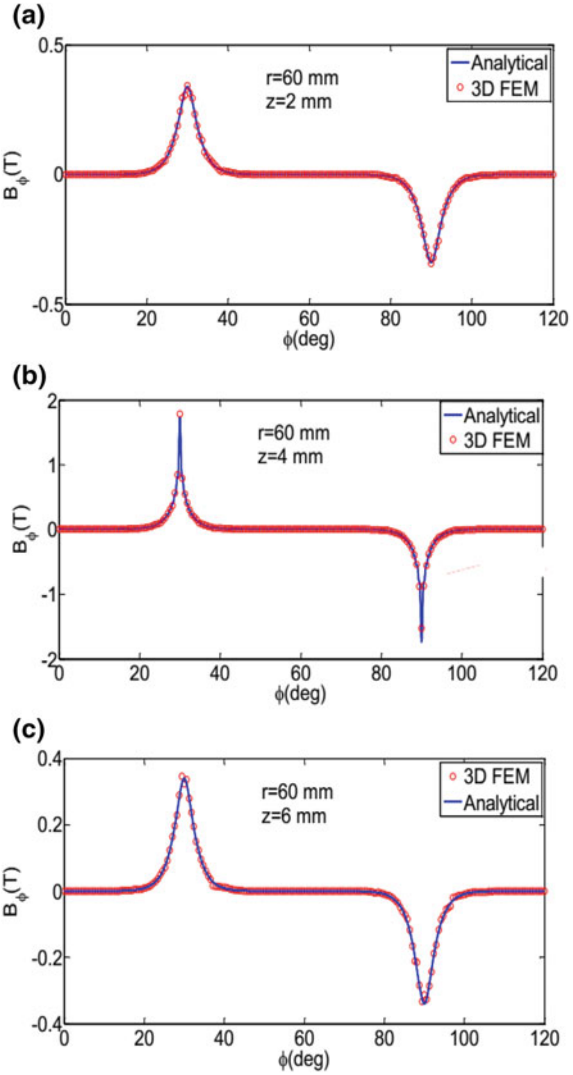
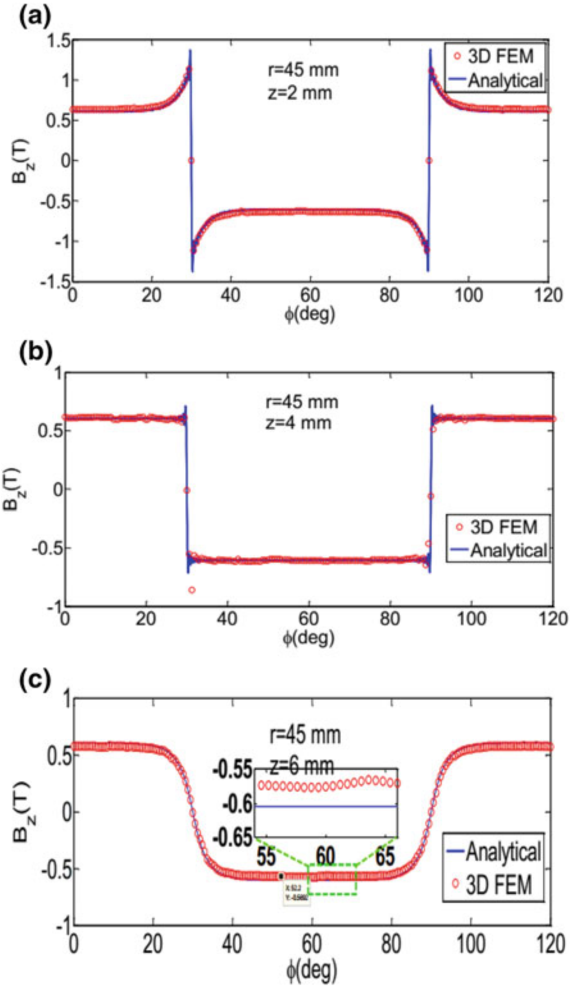


Fig. 6 Tangential flux in different points of air gap, r = 60 mm **a** z = 2 mm **b** z = 4 mm **c** z = 6 mm

magnetic field, exactly (radial leakage flux is neglected in Quasi-3D method). This little amount of differences is shown in Fig. 7c clearly. This error makes analytical induced voltage grows bigger than real numerical induced voltage. Various shapes of magnet can be considered in Quasi-3D method. In last simulation, proportion of magnet arc length to pole pitch was 1. Figure 12 shows no-load induced voltage of armature coil where internal arc length proportion is $\alpha_i = 0.3$ and external one is assumed to be $\alpha_o = 1$. As its clear in Fig. 12, numerical and analytical results have a good adaption.

Fig. 7 Axial flux in different points of air gap, r = 45 mm **a** z = 2 mm **b** z = 4 mm **c** z = 6 mm



Analytical simulation time in order to calculate induced voltage is 2.66 s, while numerical simulation time without considering non-linear effects is about 1 h and 8 s. So, analytical methods are more efficient than numerical ones in design and optimization studies.

Fig. 8 Tangential flux in different points of air gap, r = 45 mm **a** z = 2 mm **b** z = 4 mm **c** z = 6 mm

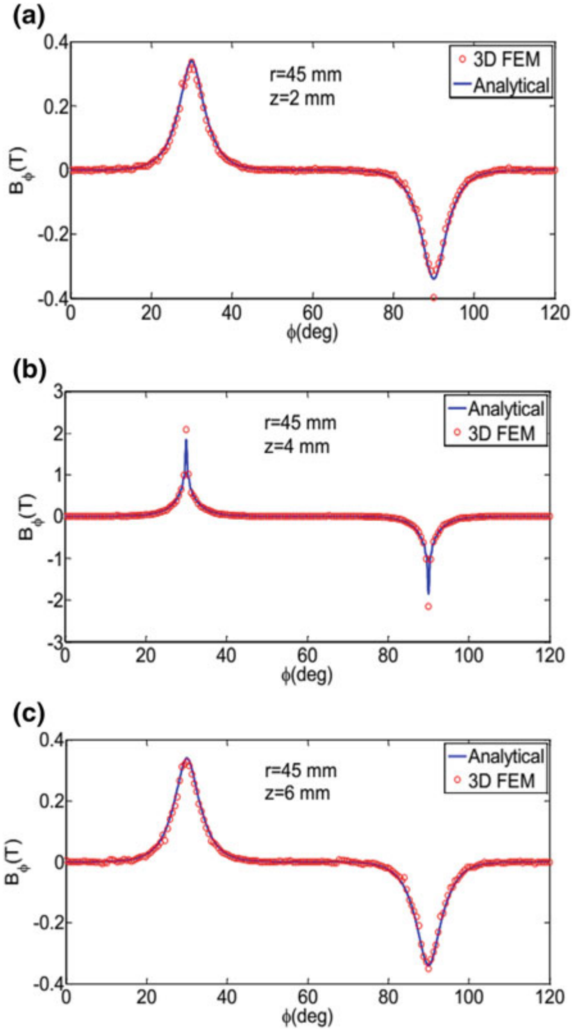
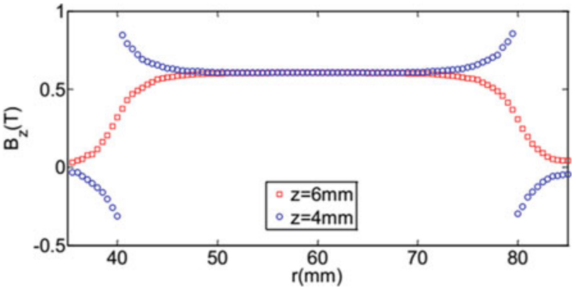


Fig. 9 Distribution of axial flux in direction of radius in center of magnet



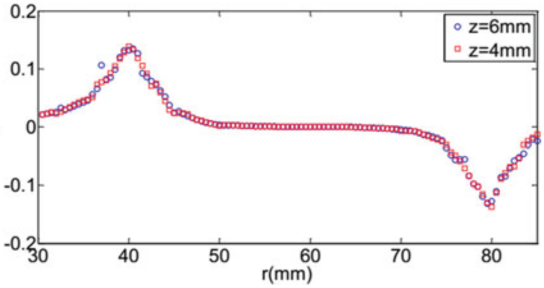


Fig. 10 Distribution of radial flux in direction of radius

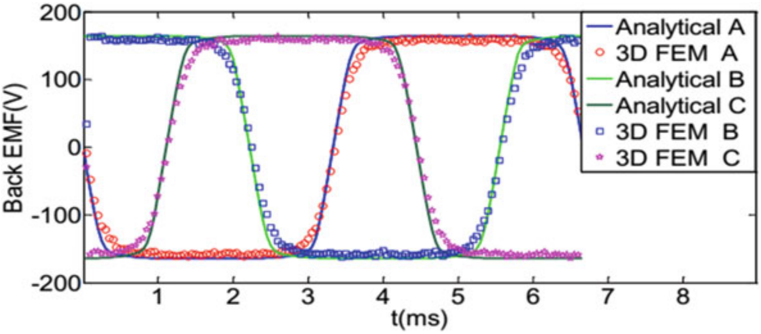


Fig. 11 Induced voltage in armature phases

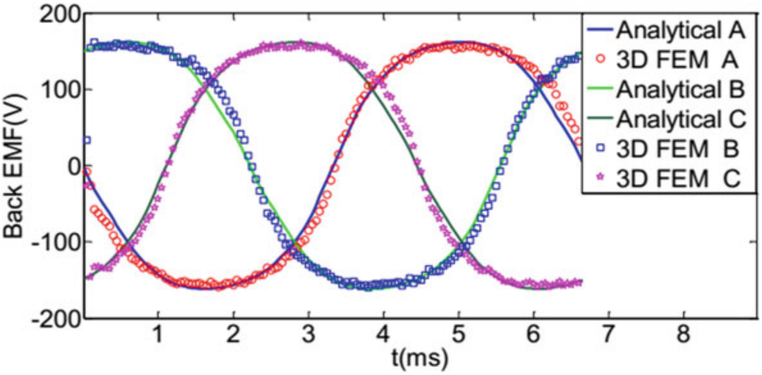


Fig. 12 Induced voltage in armature phases

4 Conclusion

In this paper, a general analytical model for predicting no load magnetic flux density in the air gap of slot-less axial flux machine with a surface mounted permanent magnet was introduced. Although this analytical model was used to investigate the field distribution in axial flux machine with two external rotors and an interior stator, it can be easily extended to other structures of AFPM machines. The only limitation is the uniformity of the air gap with slot-less stator surface mounted PM. The proposed analytical method, in comparison with the numerical method of finite element used in electrical machine software, is very fast, and in a few seconds. No- load flux density of the air gap and induced voltage of stator are calculated. Therefore, this method is suitable for the initial stages of designing and optimizing. Finally, the validity of the analytical method and the results were confirmed by their adjustment to the results of the FEM.

References

1. Gieras JF, Wang RJ, Kamper M (2004) Axial flux permanent magnet brushless machines, 2nd edn, Kluwer Academic Publishers, Dordrecht
2. Aydin M, Huang S, Lipo TA (2004) Axial flux permanent magnet disc: machines a review. Proc Int Symp Power Electron Electric Drives Autom Motion SPEEDAM 2004:61–71
3. Kahourzade S, Gandomkar A (2014) A comprehensive review of axial-flux permanent-magnet machines. Can J Electric Comput Eng 37(1)
4. Giulii Capponi F, De Donato G, Caricchi F (2012) Recent advances in axial-flux permanent magnet machines technology. IEEE Trans Ind Appl 48(6):2190–2205
5. Fitzgerald AE, Kingsley JC, Umans SD (1992) Electric machinery, 5th edn, McGraw-Hill, New York
6. Zhu L, Jiang SZ, Zhu ZQ, Chan CC (2009) Analytical methods for minimizing cogging torque in permanent-magnet machines. IEEE Trans Magn 45(4):2023–2031
7. Carmeli MS, Dezza FC, Mauri M (2006) Electromagnetic vibration and noise analysis of an external rotor permanent magnet motor. In: International symposium power electronics, electrical drives, automation and motion, pp 1028–1033
8. Bi C, Jiang Q, Lin S (2005) Unbalanced-magnetic-pull induced by the EM structure of PM spindle motor. In: Proceedings of eighth international conference on electrical machines and systems, pp 183–187
9. Boules N (1984) Two-dimensional field analysis of cylindrical machines with permanent magnet excitation. IEEE Trans Ind Appl IA-20(5):1267–1277
10. Boules N (1985) Prediction of no-load flux density distribution in permanent magnet machines. IEEE Trans Ind Appl IA-21(3):633–643
11. Zhu ZQ, Wu LJ, Xia ZP (2010) An accurate subdomain model for magnetic field computation in slotted surface-mounted permanent-magnet machines. IEEE Trans Magn 46(4):1100–1115
12. Zhu ZQ, Howe D, Bolte E, Ackermann B (1993) Instantaneous magnetic field distribution in brushless permanent magnet dc motors. Part I: open-circuit field. IEEE Trans Magn 29 (1):124–135
13. Zhu ZQ, Howe D, Chan CC (2002) Improved analytical model for predicting the magnetic field distribution in brushless permanent-magnet machines. IEEE Trans Magn 38(1):229–238

14. Zhu ZQ, Ng K, Schofield N, Howe D (2001) Analytical prediction of rotor eddy current loss in brushless machines equipped with surface-mounted permanent magnets. Part I: magnetostatic field model. In: Proceedings Fifth International Conference Electrical Machines and Systems, pp 806–809
15. Zhu ZQ, Xia ZP, Wu LJ, Jewell GW (2009) Analytical modelling and finite element computation of radial vibration force in fractional-slot permanent magnet brushless machines. In: International Electric Machines and Drives Conference, Miami, FL, pp 157–164
16. Wang J, Xia ZP, Howe D (2005) Three-phase modular permanent magnet brushless machine for torque boosting on a downsized ICE vehicle. IEEE Trans Veh Technol 54(3):809–816
17. Kumar P, Bauer P (2008) Improved analytical model of a permanent-magnet brushless DC motor. IEEE Trans Magn 44(10):2299–2309
18. Zhu ZQ, Howe D, Xia ZP (1994) Prediction of open-circuit airgap field distribution in brushless machines having an inset permanent magnet rotor topology. IEEE Trans Magn 30 (1):98–107
19. Markovic M, Jufer M, Perriard Y (2004) Reducing the cogging torque in brushless DC motors by using conformal mappings. IEEE Trans Magn 40(2):451–455
20. Zarko D, Ban D, Lipo TA (2008) Analytical solution for cogging torque in surface permanent-magnet motors using conformal mapping. IEEE Trans Magn 44(1):52–65
21. Parviainen A, Niemelä M, Pyrhönen J (2004) Modeling of axial flux permanent-magnet machines. IEEE Trans Ind Appl 40(5):1333–1340
22. Jin P, Yuan1 Y, Minyi J, Shuhua F, Heyun L, Yang H, Ho SL (2014) 3-D analytical magnetic field analysis of axial flux permanent-magnet machine. IEEE Trans Magn 50(11)

The Improvement of Voltage Reference Below 1 V with Low Temperature Dependence and Resistant to Variations of Power Supply in CMOS Technology

Amirreza Piri

Abstract In this article, the objective is designing a linear voltage reference based on CMOS technology and a structure which is insensitive to variations of temperature and supply power. In such a case, accuracy of circuit output will be optimal under different conditions. Among such sensitivities, one could point to variation of output in relation to temperature, variation due to output performance of the structure and currents, noises and turbulence. First, different voltage references, their structures and advantages and disadvantages will be reviewed individually. Then, output startup method will be explained through bulk transistor and parallel combination of transistors in output for control of output leakage current. This is followed by elaboration of reference building designed based on this method. Consequently, intended structure will designed by taking above-mentioned objectives into account. The circuit simulation and circuit layout will be done through H-Spice Software and Cadence applications respectively. The pre-layout and post-layout results signify improved results and resistance of suggested circuit against substrate noise and noise of power supply. Simulations will be done through 0.18 um CMOS technology.

Keywords Low voltage · Band-gap reference · Low noise · CMOS
Voltage reference

1 Introduction

Rapid development of CMOS manufacturing process led to reduced top limit of supply voltage and allowable power consumption in analogue, digital and RF integrated circuits. At the moment, is the most significant challenge of design of Nano-Electronic chips, Further limitation of linear performance area of dynamic

A. Piri (✉)
Department of Electrical Engineering, South Tehran Branch,
Islamic Azad University, Tehran, Iran
e-mail: Amirreza_piri@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_41

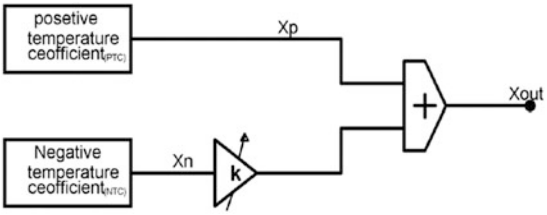
range and added sensitivity of outputs to noise of power supply are among such challenges. The band gap voltage generator circuit is one of the common circuits the output voltage of which should vary insignificantly with temperature. This is while common band gap circuits are not suitable for low voltage applications due to relatively high output voltage (which increases required supply voltage) and non-linear variation of output with higher orders than a specific temperature (Tn). Therefore, adoption of certain methods should be accompanied by lower supply voltage requirement and minimized dependency of circuit output on sentences with higher orders than T. These problems increase when design of integrated circuits for portable systems is concerned since such systems supply their required energy from weak energy sources such as micro-batteries or energy sources that are available in nature. In order to increase the lifetime of batteries and minimize supply noise, designs of such circuits should be characterized by high-efficiency reference voltage supplies and low occupational and output noise levels. In micro-electronics, these supplies are essential parts of analogue and digital systems. The primary role of these supplies is generating precise and stable voltage against variations of temperature. They also should generate linear voltage for other parts such as operational amplifiers, comparators, analogue to digital convertors and digital to analogue convertors. Since introduction of silicon band gap voltage reference by Widlar in 1970s, who suggested that total bass-emitter voltage with positive temperature coefficient could be generated by a stable voltage reference, these blocks and their combinations have been widely used in bipolar manufacturing processes and CMOS. During previous years, significant efforts were made to improve the performance of a band gap reference. Most of these efforts were targeted at adding to independence of size of output voltage from variation of temperature, power supply and manufacturing process. Designs of these references are intended to reduce number of system batteries, areal of chip and power consumption of the system [1].

In this paper, some of these methods will be introduced and a new method is introduced which uses two temperature independent currents (of first order). The two currents were generated by relatively simple low voltage band gap circuits.

The objective of design of voltage reference circuit is generating a voltage which is independent of power supply, process and temperature. There are different solutions for developing a fixed supply which could be grouped into 3 categories:

1. Use of a Zener diode which breaks down at reverse bias of certain voltage. Major disadvantages of this method is that the method does not generate a continuous value and it cannot be used in CMOS technology. In addition, breakdown voltage of Zener diode is usually higher than the supply used in current circuits [2].
2. Use of difference between threshold voltages of an incremental transistor and a discharge transistor. The disadvantage of this method is that most of CMOS circuits cannot access discharge transistors. In addition, if access to such transistors is enabled determination and stabilization of threshold voltage of transistors are difficult [3].

Fig. 1 Generatio of band gap reference voltage with positive and negative temperature coefficient


posetive temperature ceofficient(PTC)
Xp
Negative temperature ceofficient(NTC)
Xn
k
+
Xout

3. Use of band gap circuits in which current of a PTAT element eliminates thermal dependence. Today, this method is widely used for design of integrated circuits. Usually, a PN bond is used as CTAT element. In this case, diagram of a band gap circuit could be represented in the following manner [4] (Fig. 1).

If two quantities with different temperature coefficients and proper weights are summed, zero temperature coefficients will result. For instance, for two voltages of *V1* and *V2* which change opposite to each other in relation to temperature we select $\propto_1$ and $\propto_2$ in a way that we have:

$$\propto_1 \frac{\partial V_1}{\partial T} + \propto_2 \frac{\partial V_2}{\partial T} = 0 \tag{1}$$

In order to obtain the reference voltage $V_{REF} = \propto_1 V_1 + \propto_2 V_2$ with zero temperature coefficient TC = 0, two voltages with positive and negative temperature coefficients should be obtained. Among different parameters of a transistor made based on semiconductor technologies, bipolar transistors have generable quantities through which negative and positive temperature coefficients could be obtained. Although CMOS parts are candidates for generation of reference, the core of such circuits is made up of bipolar transistors.

Development of voltage references with low volume, low supply voltage, low power and high performance contributed to their extensive use in analogue and mixed mode circuits (e.g. DC–DC convertors, PLI, A/D, and D/A).

Voltage references, temperature independent DC voltage, develop power supply for manufacturing process. Typical voltage references are usually based on band gap voltages which limit minimum supply voltage of the whole circuit. In addition, band gap voltage circuits are impractical without bipolar transistor.

In this case, techniques of body biasing and body effect approximation techniques were used. A critically significant CMOS voltage reference without certain resistance and parts is introduced here. The suggested circuit is completely insensitive to temperature and supply voltage. With at supply voltage of less than 1 V and input current of less than 235 nA, the reference could operate in all temperature ranges. In addition, the circuit is characterized by low output resistance and ability to eliminate variations of power supply and noise.

Linear Matching of Threshold Voltage (V_{TH}) and Thermal Voltage (V_T).

As primary approximation, threshold voltage could be regarded as linearly reducing with temperature [1, 2]. Here, K is coefficient of temperature dependent model.

$$\mathrm{V_{TH}} = \mathrm{V_{TH}(T0)} - \mathrm{k(T - T0)} \tag{2}$$

Since $\mathrm{V_{TH}}$ has negative thermal coefficient, equation of output current will be:

$$I_{DS} = \frac{1}{2}\mu C_{ox} K (V_{GS} - V_{TH})^2 (1 + \lambda V_{DS}) \tag{3}$$

If modulation effect of channel length is exclude (λ = 0), we have:

$$V_{GS} = V_{TH} + \sqrt{\frac{2I_{DS}}{K\mu C_{ox}}} \tag{4}$$

A simple solution for determination of zero thermal coefficient, is generation of current as described in the following. The bias current should be linearly dependent on carrier mobility and it should be a coefficient of reference $\mathrm{V_T}$. In this case, we have:

$$I = a\mu C_{ox} \tag{5}$$

$$\mathrm{V_{GS}} = \mathrm{V_{TH}} + \alpha \mathrm{V_T} \tag{6}$$

Since $\mathrm{V_{TH}}$ has negative thermal coefficient and $\mathrm{V_T}$ has positive thermal coefficient, a reference voltage with zero thermal coefficient could be generated. This means that reference voltage is independent of temperature.

A reference voltage, generated by linear combination of VTH and VT, could obtained by a current supply and MOSFET diode connection. Figure 2 shows reference voltage at output A.

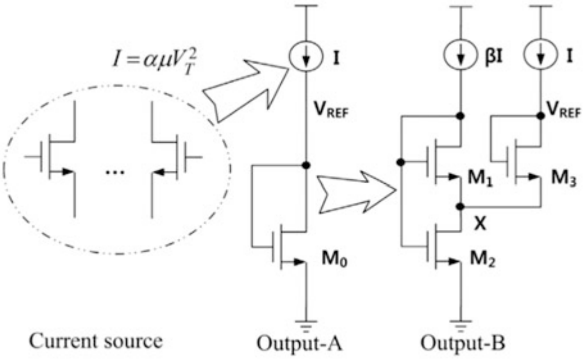


Fig. 2 Schematic representation of suggested reference voltage

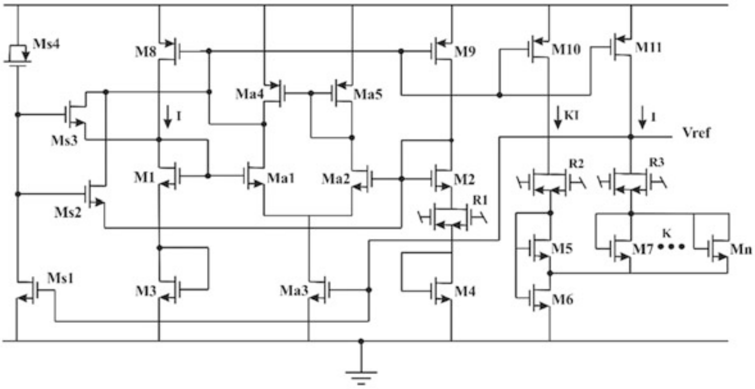


Fig. 3 Schematic representation of suggested circuit

The structure of transistor circuit is represented in the following Fig. 3.

In general, the circuit is made up of three parts: current supply, startup circuit and output each of which will be detailed individually in the following.

- Current Supply

The current supply of the circuit is as shown in the following Fig. 4.

The current supply used to generate current I is the current source connected to output circuit. In this circuit, M8 and M9 are below the threshold and of identical dimensional ratios. The branch made up of M8, M1 and M3 has similar structure to the branch made up of M4, M2 and M9 but M1 and M2 have different dimensional ratios and the same is the case for dimensional ratios of M3 and M4. The design guarantees that M1 and M2 are at saturation zone and M3 and M4 are below bias threshold. This is due to generation of higher current by CMOS transistors when they are in saturation zone. Therefore, bias was located in saturation zone.

A differential input amplifier was used to maintain M1 and M2 at identical gate voltage and to keep drain-source voltage of M8 and M9.

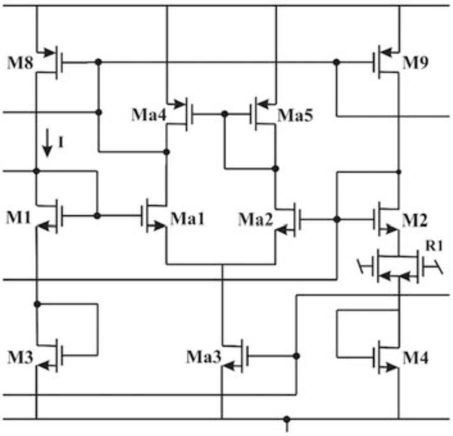
In order to achieve identical bias current, M1 offers higher conductivity as it has larger dimensions. Therefore, negative input of amplifier is connected to B node since M1 has quicker variation of current than voltage variations.

$$\mathrm{V_{GS1}} + \mathrm{V_{GS3}} = \mathrm{V_{GS2}} + \mathrm{V_{GS4}} \tag{7}$$

Since length of M1–M2 channel is not sufficiently large, modulation of length of channel was excluded and gate-source voltage was determined through following equation:

$$V_{GS} = V_{TH} + \sqrt{\frac{2I_{DS}}{K\mu C_{ox}}} \tag{8}$$

Fig. 4 Current supply of circuit



Supposing that currents of two branches are identical, we have:

$$V_{GS3} + V_{TH1} + (\eta - 1)V_{GS3} + \sqrt{\frac{2I}{\mu C_{ox} K_1}} = V_{GS4} + V_{TH2} + (\eta - 1)V_{GS4} + \sqrt{\frac{2I}{\mu C_{ox} K_2}} \tag{9}$$

Since M3 and M4 are below threshold, currents of these transistors are obtained approximately through following equation:

$$I_{DS} = \mu C_{ox} K(\eta - 1)V_T^2 exp\left(\frac{V_{GS} - V_{TH}}{\eta V_T}\right) \times \left[1 - exp\left(\frac{-V_{DS}}{V_T}\right)\right] \tag{10}$$

$$I_{DS} = \mu C_{ox} K(\eta - 1)V_T^2 exp\left(\frac{V_{GS} - V_{TH}}{\eta V_T}\right) \times \left[1 - exp\left(\frac{-V_{DS}}{V_T}\right)\right] \tag{11}$$

Based on suppositions of design of this part of reference voltages [13, 16], we have:

$$\mathrm{V_{DS}} > 4\mathrm{V_T}.$$

In this circuit, we have:

$$\mathrm{V_{TH1}} = \mathrm{V_{TH2}}$$

$$\eta\left[V_{TH3} - V_{TH4} + \eta V_T In\frac{K_4}{K_3}\right] = \sqrt{\frac{21}{\mu C_{ox}}}\left(\sqrt{\frac{1}{K_2}} - \sqrt{\frac{1}{K_1}}\right) \tag{12}$$

$$I = \frac{1}{2}\mu C_{ox}\left[\eta^2 V_T \frac{\sqrt{K_1 K_2}}{\sqrt{K_1}\sqrt{K_2}} In \frac{K_4}{K_3}\right]^2 \tag{13}$$

Above equation could be rewritten in the following manner:

$$I = a\mu C_{ox} V_T^2 \tag{14}$$

$$I = \frac{1}{2}\mu C_{ox}\left[\eta^2 V_T \frac{\sqrt{K_1 K_2}}{\sqrt{K_1}\sqrt{K_2}} In \frac{K_4}{K_3}\right]^2 \tag{15}$$

In regard to current supply, approximate difference or deviation of body effect and approximation of I–V characteristic should be considered. In the suggested circuit, transistors M3 and M4 operate blow threshold zone, namely:

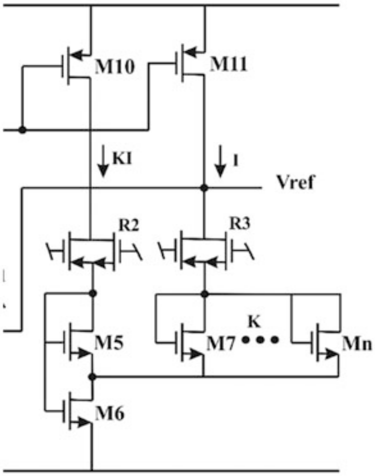
$$V_{SB1,2} = V_{GS3,4} = V_{DS3,4} \tag{16}$$

The output part of the circuit is as shown in the following Fig. 5.

The output circuit used for generating reference voltage with temperature compensation is made up of three branches. In a branch with current source biased by I1 current, identical dimensional ratio (M1 transistor) is identical while second branch is biased by I2.

$$V_{REF} = V_{th}(T0) - K(T - T0) + V_T Ln \frac{(1+\beta)K_5}{\beta(\eta - 1)K_6 K_7} \tag{17}$$

Fig. 5 Output part of circuit



Substituting Eq. 14 into Eq. 16, we have:

$$V_{REF} = V_{th}(T0) - K(T - T0) + V_T Ln\left[\left(\eta^2 \frac{\sqrt{K_1 K_2}}{\sqrt{K_1 -} \sqrt{K_2}} In \frac{K_4}{K_3}\right)^2\right]^{1/2} \frac{(1+\beta)K_5}{\beta(\eta - 1)K_6 K_7} \tag{18}$$

Since we have $\partial V_{ref}/\partial T = 0$, the following temperature-independent reference voltage output will be obtained:

$$\frac{1}{2}\left(\eta^2 \frac{\sqrt{K_1 K_2}}{\sqrt{K_1 -} \sqrt{K_2}} In \frac{K_4}{K_3}\right)^2 \frac{(1+\beta)K_5}{\beta(\eta - 1)K_6 K_7} = exp \frac{kp}{\eta k_B} \tag{19}$$

Since TC of reference voltage (V_{REF}) is equal with zero, drain current of M7 (I) at room temperature is determined through following equation:

$$I = \mu C_{ox} V_T^2 \frac{K_6 K_7 \beta(\eta - 1)}{K_5(1+\beta)} exp \frac{kT_0}{\eta V_T} \tag{20}$$

The quiescent current of the whole current is determined by current I since currents of other branches is sourced by I. In the case of fixed β, quiescent current represented by I in Eq. 19 deceases by increase of dimensional ratio of M5 and decrease of dimensions of M6, M7. Basically, both circuits of supply source and output are sources of CMOS current. In order to reduce modulation effect of channel length which causes mismatch between currents of branches, lengths of channels of transistors M8–M11 were presumed to be significant.

Startup Circuit:

Reference voltage requires a startup circuit. In this case, transistors MS1–MS4 act as startup circuit (Fig. 6).

Since we have $V_{GS8} = V_{GS9} = V_{GS10} = V_{GS11}$, selection of different W/Ls makes transistors identical or a factor of each other. Because a current is a factor of V_T^2 and transistors M8 and M9 operate blow threshold zone, output currents are a factor of V_T^2 too. Mathematical and manual calculations of sizes of transistors are represented in the following.

Simulation and Results

In this paper, different threshold voltages are used by drawing on body bias techniques and using similar part. In this case, structural steps of certain parts are excluded and variations of previous works were ignored. Therefore, body effect was excluded altogether.

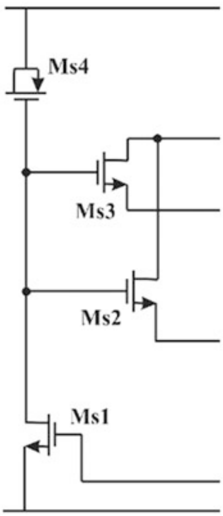
Figure 9 shows dependence of threshold voltage on V_{BS} in room temperature [16]. In this case, SMICA 18 mm technology was used (Fig. 7).

Therefore, V_{TH} has an approximate linear correlation with V_{BS}.

Current Source:

The current source used for current I is current source input to output circuit. In this circuit, M8 and M9 are blow threshold and of identical dimensional ratio.

Fig. 6 Startup circuit



The branch made up of M3, M1 and M8 has similar structure to the branch made up of M4, M2 and M9. However, dimensions of M1 and M2 are different and the same is the case for dimensions of M3 and M4. Design should be done carefully so as to guarantee that M1 and M2 are in saturation zone and M3 and M4 are below bias threshold.

A differential input amplifier was used for maintaining M1 and M2 at identical gate voltage and to maintain drain-source voltage of M8 and M9 (i.e. transductor).

In the case of identical bias current, m1 with larger dimension has higher conductivity. Therefore, negative input of amplifier is connected to node B since M1 has quicker current variation than voltage variation. In addition, amplifier improved the performance of PSRR.

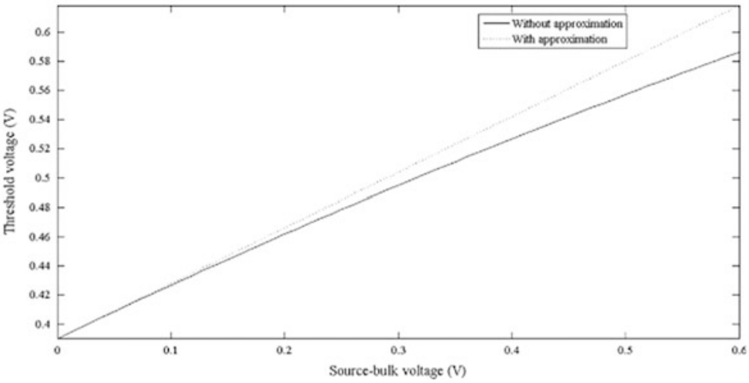


Fig. 7 Dependence of threshold voltage on V_{BS} [16]

Voltage of gate M1–M2 is explained in terms of sum of gate-source voltage of M1 and M3 and sum of gate-source voltage of M2 and M4.

Because of clamp action of amplifier, gate voltages of M1 and M2 are identical. Since lengths of channel M1, M2 is sufficiently large, modulation of length of channel was excluded and gate-source voltage was obtained through its relevant equation.

In regard to current source, approximate difference (deviation) of body effect and approximation of I–V characteristic should be considered. In case of suggested circuit, both transistors M3 and M4 operate below threshold. In this design, $V_{GS4} = 192$ mV and $V_{GS3} = 224$ mV are presumed as they offered better performance.

The output circuit used for generation of reference voltage with temperature compensation is made up of two branches. A branch is biased by current source which is characterized by current I and identical dimensional ratio (transistor M11). This is while second branch is biased at ratio B. Basically, current source circuit and output circuit are sources of CMOS current.

In order to reduce modulation effect of channel length which creates a mismatch between currents of branches, lengths of channels of transistors M8–M11 were opted to be quite significant. Reference voltage requires a startup circuit and transistors MS1–MS6 act as startup circuit.

Variation of Voltage and Output Current per Variations of Supply for 18 um Technology.

In this case, output voltage of supply ranges between 0.85 and 2.5 V and it is equal with 0.6 V which is identical with results of this paper. In order to achieve this result, voltage of supply should be increased from 0 to 2.5 V and the voltage generated by the supply is higher than 0.6 V. The output does not change. Figure 8 represents similar analysis of variation of current. Similarly, when voltage of supply exceeds 0.6 V the current will be fixed at 26 nA (Fig. 9).

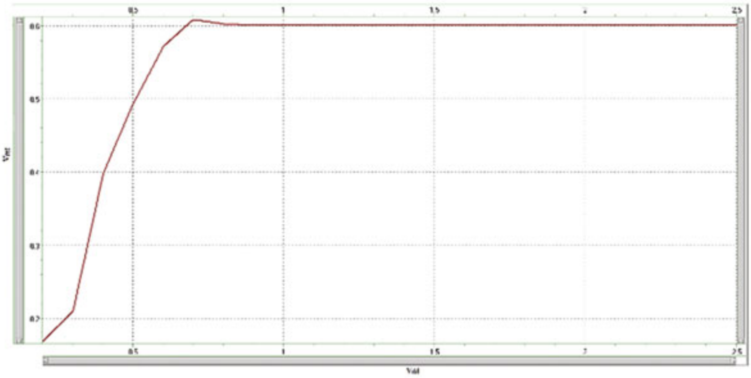


Fig. 8 Variation of output voltage versus variations of supply

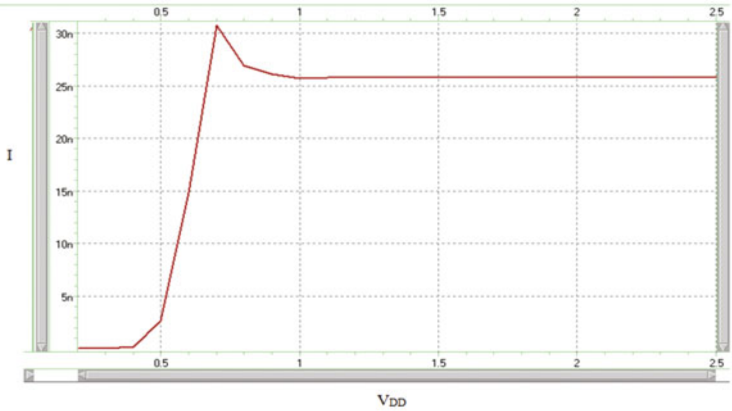


Fig. 9 Variations of current output versus variations of supply voltage

The 0.18 um technology led to 3 mV reduction. The variations are due to increase in voltage range of voltage supply. In the following, analysis of variation of voltage output per variation of supply at 0.5 um scale as well as variation of current output per variation of variation of supply at 0.5 um scale are represented. The results signify proper performance of suggested circuit. The only difference is that in this technology, final and stable solution is obtained when voltage output to the supply exceeds 2.2 V.

The layout of suggested circuit was made through Cadence Software (Fig. 10).

Then, variations of output voltage and output currents at different temperatures ranging from −20 to 100 °C as well as different supply voltages were determined (Fig. 11).

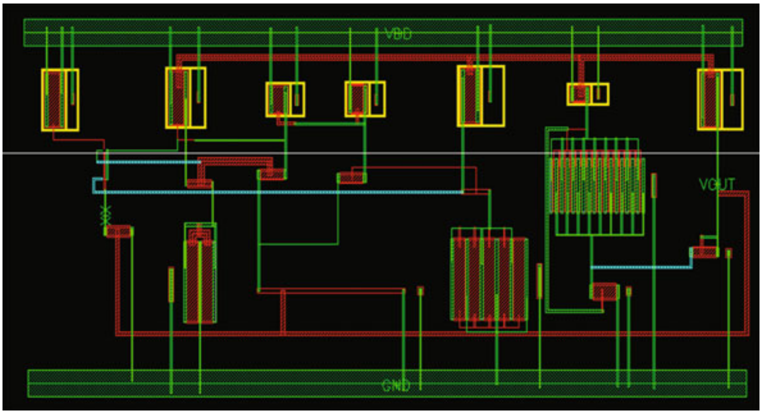


Fig. 10 Schematic representation of circuit

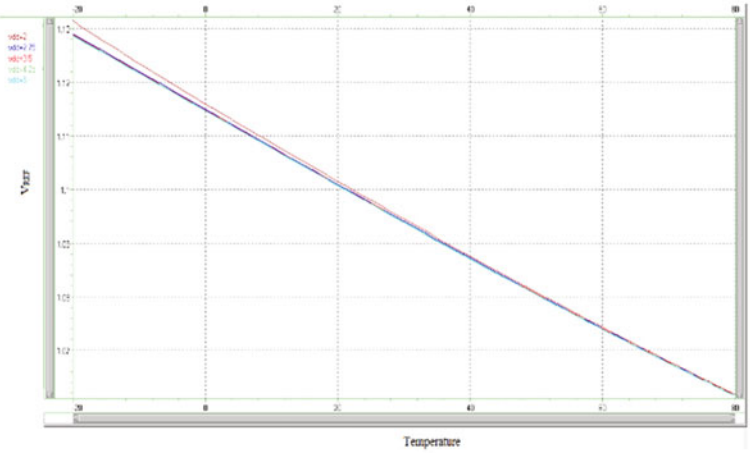


Fig. 11 Variation of output voltage for different temperatures after layout

In this case, variations of output voltage for different temperatures ranging from −20 to 80 °C are represented and they are similar to findings of simulation in H-Spice Software (Fig. 12).

Finally, simulation was done at different corners of the process and obtained results are as shown in the following Fig. 13.

Observably, time to attain steady state is dissimilar for different corners. In SS corner, more time is needed for circuit to attain steady state. The results were obtained after circuit layout (Fig. 14).

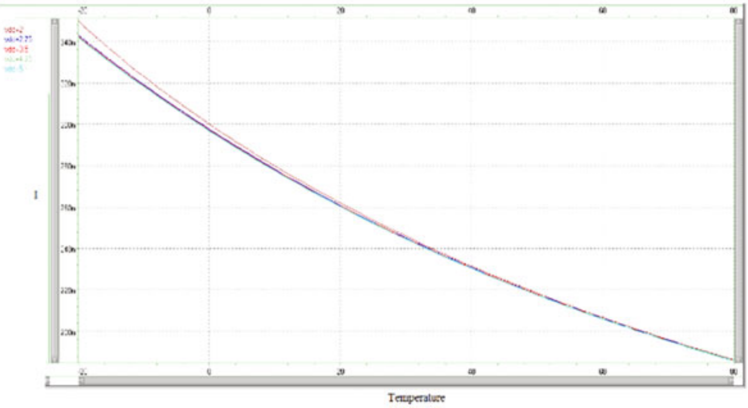


Fig. 12 Variation of output current for different temperatures after layout

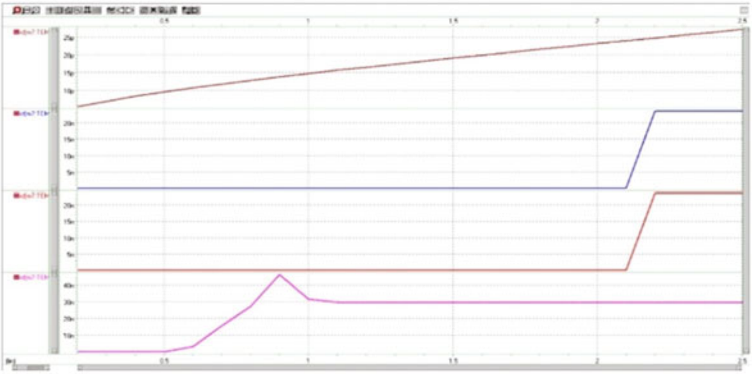
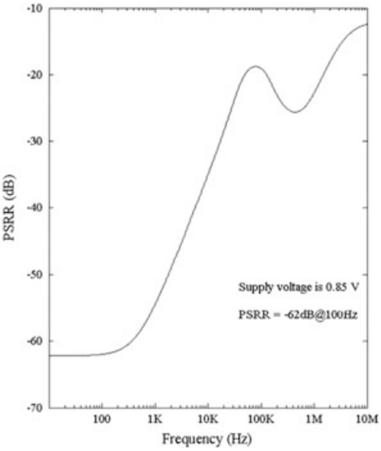


Fig. 13 Variation of output current for different corners after circuit layout

Fig. 14 Calculation of PSRR based on frequency variation (0.85 V)



Maintaining power supply rejection ratio (PSRR) is possible when reference voltage of amplifiers is from a voltage supply in which there are voltage dividers. Usually this issue (i.e. isolation of variation of voltage supply and its noises from reference voltage of amplifiers) is ignored during design of circuits.

This is a significant problem since real voltage supplies of circuits are not quite ideal and any AC signal on supply line could be fed back into the circuit and be amplified. Under logical conditions, this situation may lead to unwanted variation.

This problem is addressed in internal designs of modern op-amps as a parameter called PSRR is included in relevant data sheets which represents power and ranges from −80 to −100 dB. In common op-amps, value of this parameter ranges from −40 to −100 dB and this should be noted in future designs (Fig. 15).

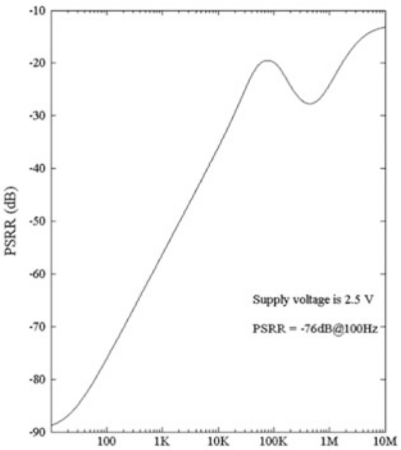


Fig. 15 Calculation of PSRR based on frequency variation (2.5 V)

Observably, PSS for 0.85 V voltage supply is −62 dB at 100 Hz frequency while for 2.5 V voltage supply and similar frequency it is equal with −76 dB. At the end of this chapter, a comparative table is included so as to compare the results with findings of previous studies.

2 Conclusion

Based on the comparison, one could suggest that low technology used to complete this thesis led to coverage of a supply voltage ranging from 0.085 to 2.5 V. In terms of operating voltage range, the results signify improvement in comparison with previous findings. In addition, significant improvement of operating temperature range were made and output current was lower than previous studies. Finally, comparison of circuit layouts with previous works suggest that the circuit occupies less space than integrated circuit (Table 1).

Table 1 Results of comparison between present study and previous studies

Parameter	Reference [4]	Reference [5]	Reference [8]	Reference [9]	This work
Process (μm)	0.35	0.18	0.35	0.13	0.18
Supply voltage	2–3.5	0.45–2	1.4–3	1–2.3	0.085–2.5
Temp (°C)	0–80	0–125	−20 to 80	1–2.3	0.085–2.5
Vref (mV)	800	263.5	747	0–100	−20 to 100
Supply current (μA)	0.4@2 V	0.4@0.45 V	2.1@1.4 V	8.1@1.2 V	0.235@2 V
PSRR (dB)	−51.4 @100 Hz	−45 @100 Hz	−45 @100 Hz	−44 @100 Hz	−62 @100 Hz
Area (mm^2)	0.041	0.043	0.055	0.053	0.023

References

1. Razavi B (2000) Design of analog CMOS integrated circuits. McGraw-Hill, New York. ISBN 0-07-238032-2
2. Johns DA, Martin K (1996) Analog integrated circuit design. Wiley, New York
3. Ma B, Fengqi Y (2014) A novel 1.2-V 4.5-ppm/C curvature-compensated CMOS bandgap reference. IEEE Trans Circuits Syst I Regul Pap 61(4):1026–1035
4. Duan Q, Roh J (2015) A 1.2-V 4.2-ppm/C high-order curvature-compensated CMOS bandgap reference. IEEE Trans Circuits Syst I Regul Pap 62(3):662–670
5. Amaravati A, Dave M, Baghini MS, Sharma DK (2013) 800-nA process-and-voltage-invariant 106-dB PSRR PTAT current reference. IEEE Trans on Circuits Syst II: Express Briefs 60(9):577–581
6. Leung KN, Mok PKT, Leung CY (2013) 5.3-ppm/rC curvature-compensated CMOS bandgap voltage reference, circuits. IEEE J Solid-State 38(3):561–564
7. Abbasi, MU, A high PSRR (2015) ultra-low power 1.2 V curvature corrected Bandgap reference for wearable EEG application. 13th International 2015 IEEE Conference New Circuits and Systems Conference (NEWCAS)
8. Chahardori M, Atarodi M, Sharifkhani M (2011) A sub 1 V high PSRR CMOS bandgap voltage reference. Microelectron J
9. Chouhan SS, Halonen K (2015) Design and implementation of a micro-power CMOS voltage reference circuit based on thermal compensation of Vgs. Microelectron J
10. Crepaldi PC, Pimenta TC, Moreno RL, Zoccal LB, Ferreira LHC (2012) Low-voltage, low-power Vt independent voltage reference for bio-implants. Microelectron J
11. Salehjoo N, Kazeminia S, Hadidi K (2014) Producing flat supply voltage using a temperature compensated BGR within LDO regulator loop. IEEE, 2014, 22nd Iranian Conference on Electrical Engineering (ICEE 2014), 20–22 May 2014, Shahid Beheshti University, pp 150–155
12. Naro GD, Lombardo G, Paolino C, Lullo G (2006) A low-power fully MOSFET voltage reference generator for 90 nm CMOS technology. ICICDT 2006
13. Lin H, Chang D (2006) A low-voltage process corner insensitive subthreshold CMOS voltage reference circuit. Proceedings of ICICDT 2006
14. Serra-Graells F, Huertas JL (2003) Sub-1-V CMOS proportional-to absolute temperature references. IEEE J Solid-State Circuits 38(1):84–88
15. Cheng M, Wu Z (2005) Low-power low-voltage reference using peaking current mirror circuit. Electron Lett 41(10):572–573
16. Luo H, Han Y, Cheung RCC, Han X, Zhu D (2010) Bulk compensated technique and its application to subthreshold ICs. Electron Lett 46(16):1105–1106
17. De Vita G, Iannaccone G (2007) A sub-1-V, 10 ppm/8C, nano powervoltage reference generator. IEEE J Solid-State Circuits 42(7):1536–1540

Micro—Electromechanical Switches Application in Smart Grids for Improving Their Performance

Shariati Alireza and Olamaei Javad

Abstract Wireless sensor networks are as series sensor node in very small dimensions with the capability of sensing the surrounding environments, processing the data which is sensed; and, sharing the data between each other in wireless form. Despite numerous capabilities of these nodes, as their energy is supplied by batteries with limited power, they have limited life. In fact, the restrictions in the energy of nodes and the life span of a network poses as one of the important challenges in (using) these networks. The sensor nodes shall have the characteristics of low consumption capability; thus, in designing the nodes hardware, one must try to use designs and parts with low consumption. Furthermore, providing the sleep mode for the whole node or each section separately is highly important. Therefore, we suggest a combo-switch in which, the Micro—Electromechanical switch is used as a MOSFET switch functions as a gate MOSFET driver with the applicability of energy collection systems. The Power Administration Circuits which use combo-switch have the capacity of very low loss and no leakage, autonomous property and high current transmission capability. The measurements show solar energy collection circuits that use combo-switches collect energy without any power supply sources and voltage source, they charge the battery or drive resistive load. The current leakage during energy collection is very low; therefore, power administration which uses the proposed combo switch could serve as an ideal solution for autonomy of wireless sensor nodes in smart grid systems.

Keywords Wireless sensor networks · Smart grid · Wireless sensor nodes
Micro electromechanical switch · Combo switch · Self-powered

S. Alireza · O. Javad (✉)
Department of Electrical Engineering, South Tehran Branch,
Islamic Azad University, Tehran, Iran
e-mail: olamaee1345@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_42

1 Introduction

These days, wireless sensor nodes have played significant roles in the monitoring and control system for various industrial applications to achieve automation, self-referencing and real-time control with great flexibility. With respect to the smart grid, a widely distributed, reliable and maintenance-free sensor network based on nonintrusive, miniaturized, and cost-effective sensors is quite essential to measure the operating status of the power grid, to monitor conditions of on-line critical equipment and to do fault detection and diagnosis, etc. [1, 2]. Wireless sensor nodes are powered by batteries with limited capacity. Interesting approaches to solve the energy problem are to attempt to achieve self-powered systems through the use of energy harvesters (EH), and various storage devices such as batteries and super-capacitors [3–15]. However, they need a time period for exchanging energy and re-charging. Instead of batteries for re-charging energy source, it is possible to supply energy for equipment and facilities which have low operation period or low energy consumption is available from light, heat or vibrations and ambiances in the environment [16, 17]. Recently, due to the unique specifications of environmental energies, the power electronic researchers have been showing more attention to power administration circuits in energy collection systems [18–22]. In real world environments, the collected power is around less or more micro watt, which is equal to the apparatus energy consumption in idle modes; which could vary in specific situations and various times [23]. Therefore; first, the energy must be collected in a storage capacitor. In the meantime, to prevent energy leakage to the load compartment in the energy collection stage, the storage capacitor shall be separated and disconnected from the load compartment. The storage capacitor can be connected to the load when sufficient energy has been already collected for load drive. This means, the collected energy must be controlled at all time. Besides, on wireless sensor nodes, at the end of data receiving process and data transmission, the storage capacitor shall be disconnected from the rest of sensor node circuit in order to start the energy collection. The main problem that poses here is these functions need a certain amount of energy while the only energy available is the same previously collected energy.

The Micro-Electro Mechanical System (MEMS) switches which are actuated electro-statically and have low actuation voltage, could serve as a suitable option in fulfilling those functions with no need to any constant voltage source.

These switches are made with Micromatching technology and the basis of their function is that they can be used to create a short-circuit or open circuit in the circuits through motion of a mechanical compartment which is actuated in a certain way and can be put in the connection or disconnection modes.

First, we will explain the structure of a combo-switch or a Micro Electromechanical -MOSFET link in which, the Micro Electromechanical switch is used as a gate MOSFET; followed by explaining the function of energy collection circuit which benefits from a Micro Electromechanical—MOSFET switch for capacitive loads with no fixed charging source. Rechargeable batteries are examples of capacitive loads which are used in wireless facilities, including wireless sensor nodes. The Micro Electromechanical—MOSFET combo switches are mostly applicable in energy collection circuits and load driving circuits.

2 The Combo-Switch

Following picture shows a schematic view of the combo switch or Micro Electromechanical—MOSFET switch. The Micro Electromechanical switch works as a gate MOSFET switch driver. Micro Electromechanical switch includes a crystal silicon motion section and a fixed glass compartment the Micro Electromechanical switch was used as an RF signal switch [24].

As it can be observed in the Fig. 1, A and B are signal lines which are separated physically in 2.5 μm distance from each other. A is connected to the motion section. The motion section is suspended by springs elastic forces; and C is connected to the electrical ground. When the voltage between the motion section and C increases and reaches to a certain voltage, the electrostatic forces exceeds the elastic force of springs; therefore, the mobile section moves down and connects A to B. This voltage is also called pull in voltage of Micro Electromechanical switch, which makes the electrical signal to be transported from A to B. On the other hand, when voltage goes lower than a certain level, the electrostatic force becomes less than elastic force of the springs; and, consequently, the mobile section moves up and A is disconnected from B. This voltage is called the pull out of Micro Electromechanical switch. Since the signal lines are built on a glass material and are physically apart, the chance of current leakage in the off mode of the Micro Electromechanical switch is relatively zero. Since the pull in voltage is always

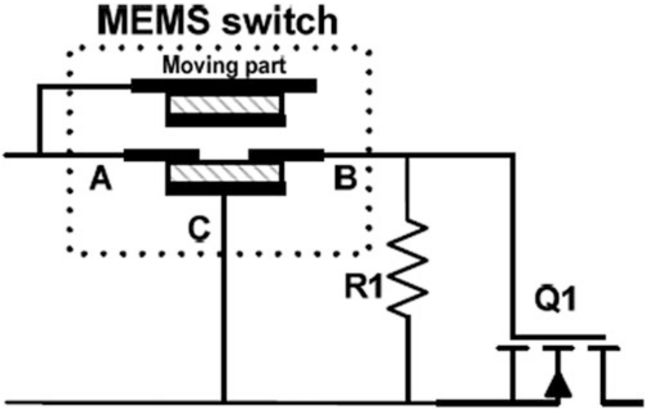


Fig. 1 Structure of micro electromechanical-MOSFET switch

higher than pull out voltage, the Micro Electromechanical switch has hysteresis characteristics.

As it has been seen in the Fig. 1, the mobile section is in the same potential level with A; therefore, if the voltage between A and C becomes less than the *pull in* voltage, the gate-source voltage increases suddenly and the MOSFET switch turns on. On the other hand, if the voltage becomes lower than the pull out voltage, A becomes isolated from B, and the gate-source voltage in R1 discharges to 0 V; and the MOSFET switch turns off in turn. Therefore, the Micro Electromechanical hysterias switch, the combo Micro Electromechanical—MOSFET switch will not need the pulse generation circuit to control the MOSFET switch.

3 Analysis and Simulation of Circuit Function

As the Fig. 2 shows, to simulate a rechargeable battery, a 470 Micro Faraday capacitor has been used as a capacitive load; and we also used a diode to prevent the reverse current leakage. We used four solar cells which were connected in series as energy collector to generate 15 V open circuit voltage in the indoor light conditions which is usual in the shop (300 lx). Since the output power of solar cells in such situation is very low, first, sufficient energy should be saved in storage capacitor Cs (100 μF) to drive the voltage regulator. The storage capacitor Cs in the course of energy collection must be isolated from the remaining circuit; therefore, since Micro Electromechanical switch is ordinarily off, the energy which is saved does not discharge in the remaining parts of the circuit.

The 750 K. Ohm resistance and 100 nf capacitor; too, have been used in the voltage regulator entry for the impedance matching [25]. The R2 resistor (0/00001 Ω) is used to prevent the error caused by connecting the source to the capacitor.

The L2 resistor (0/4 Ω) and L2 inductor (0/000001H) is used to prevent the capacitive loop by C1, Cs and MOSEFT. As the solar cells convert light energy into electric energy, the capacitor voltage gradually increases to approach the open circuit voltage of the solar cell. When the storage capacitor voltage reaches the pull in voltage (12 V) of Micro Electromechanical switch, the mobile part moves downward and the signal lines come close to each other. At this moment, the Cs voltage is used for gate of MOSEFT and the MOSEFT turns on. Then, the saved energy is transported to voltage regulator for load capacitor; therefore, in turn, the energy collected in the storage capacitor is transferred to the load capacitor. Consequently, the storage capacitor voltage decreases within 15 ms from 12 to 9.2 V (pull out voltage), most of which time is due to the latency in turning the voltage regulator on, which could be observed in Fig. 3.

When the storage capacitor voltage reduces to pull out voltage of Micro Electromechanical switch, the mobile part returns to up mode and opens the signal lines; and, the storage capacitor separates from the rest of the circuit; and, the next energy collection process is resumed. This period repeats automatically until the

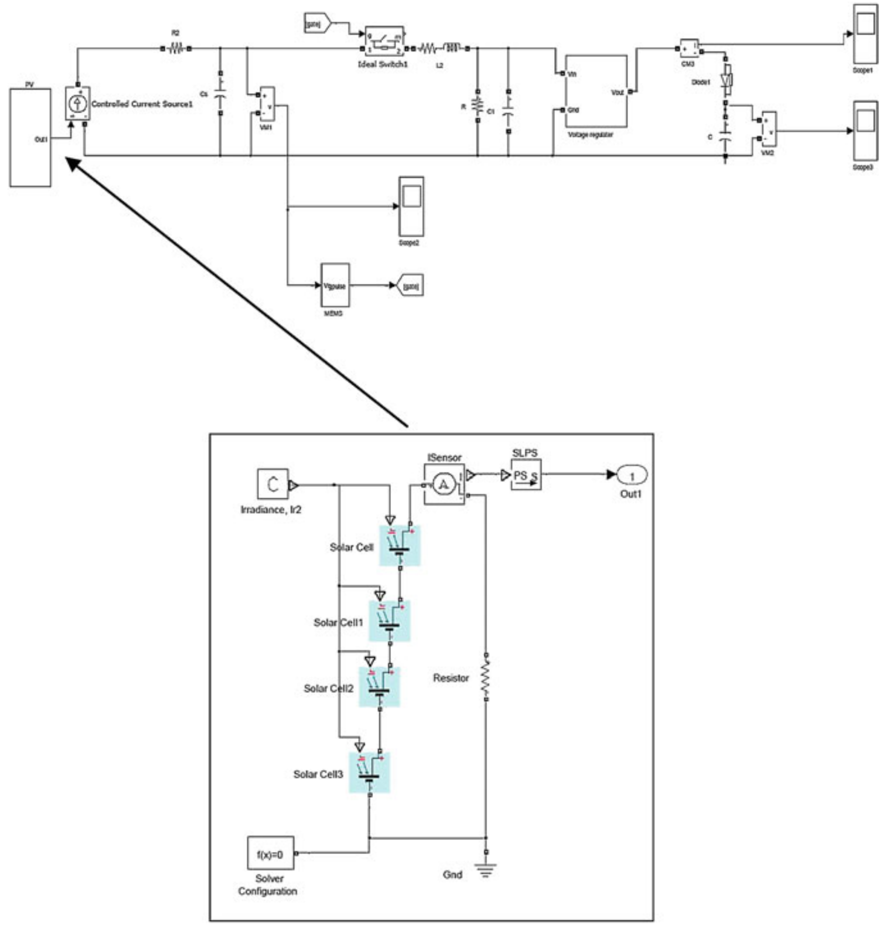


Fig. 2 Circuit architecture

storage capacitor voltage increases stepwise until it reaches the full voltage of the final charge; then the charging process stops as it could be shown in Fig. 4.

In the instances when the load capacitor voltage (or chargeable battery) reduces for different reasons, the cycle repeats except when the storage capacitor starts charging from the lowest level. The advantage of this circuit is that, the energy collection and charging is very low in dim illumination intensity. In energy collection stage, since the signal lines are made on the glass materials and are physically apart, leakage through glass material is very low and trivial. In addition, the drain–Source leakage current is trivial and could be ignored in energy collection; which means, they are fully isolated.

Now, if the solar cells connect to voltage regulator directly and without combo or Micro Electromechanical—MOSFET switch, in very low illumination conditions, the entire power or the energy generated by solar cells should be spent to

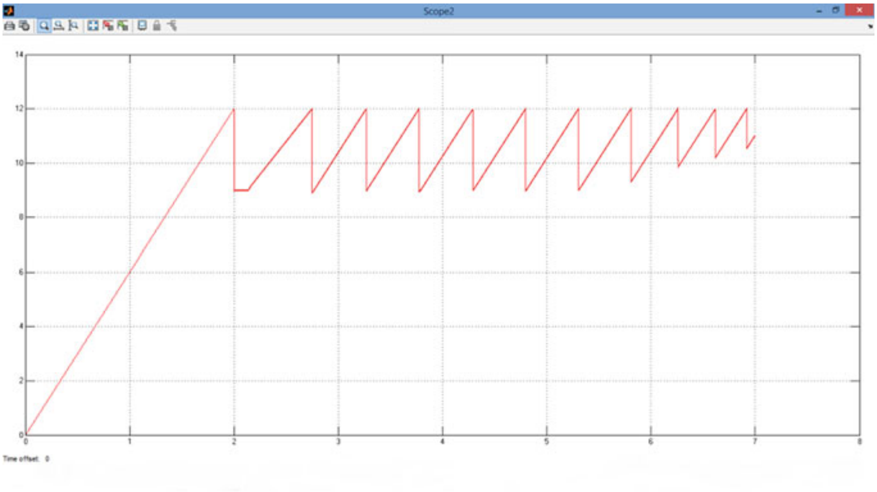


Fig. 3 Storage capacitor voltage

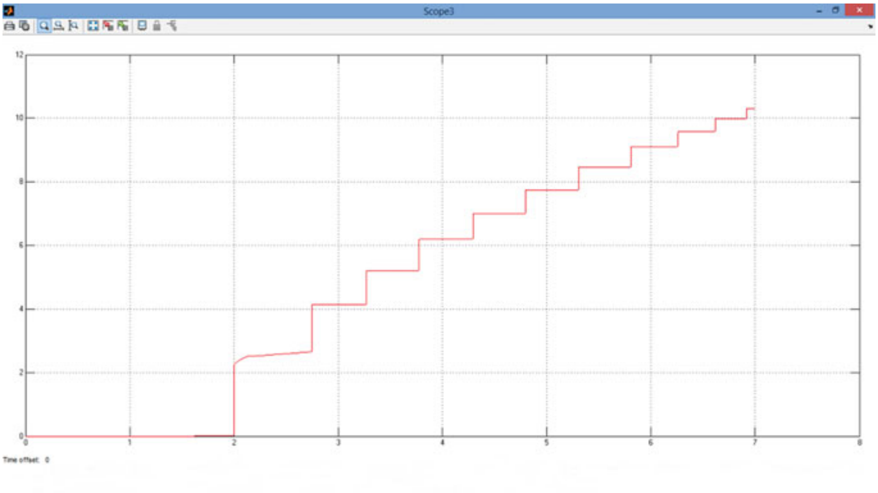


Fig. 4 Load capacitor voltage

supply voltage regulator in off power mode with no output. Therefore, the load capacitor (battery) stays in not charging mode. In addition, when we use other semiconductors [25], as an example BJT and a Zener diode, a little amount of current leakage (around few tens of nano ampere) changes into little volt (in millivolt).

When Micro Electromechanical switch is used solely as the main switch of the line without MOSEFT, the Micro Electromechanical switch, after few cycles stop functioning due to the high peak of the initial charging current and the switch breaks (does not function). Then Micro Electromechanical switch is on, as the storage capacitor 100 μF lowers in 15 ms from 12 to 9.2 V, it discharges 3 mJ, therefore, our energy collector circuit whether it is a combo switch or a Micro Electromechanical—MOSFET makes it possible to collect energy and recharge of the battery even if the power generated in the energy collection systems is very low.

4 Sum-up and Conclusion

The first advantage of combo switch or Micro Electromechanical—MOSFET is that power loss in the energy collection process is near zero; for, the Micro Electromechanical switch signal lines are made on the glass material and are physically separated, the current leakage through the glass material is trivial (less than 12 Pico Ampere).

The second advantage is its simple structure and function with very low power. A plain Micro Electromechanical switch for the gate drive is the alternative of one form of pulse production circuit with a number of items; as, we know, in MOSEFT discussion, we anticipate in ideal situation that Rds resistance (Drain-source resistance) is very high in off mode (infinite) and very low (zero) in off mode. The main problem in using MOSEFT with micro-controllers is that, most MOSEFTs need between 10 and 15 V potential difference between gate and source in order to minimize their RDs; and since micro-controllers usually operate in 3.3 V, they lack the ability to supply this difference in potential; therefore, we need additional circuits known as MOSEFT starter; however, by using MEMS switch, we no longer need them. In the meantime, the Micro Electromechanical switch does not have an external power source, nor a voltage source that would require consuming power.

And finally, since the main power is transported to the load via MOSEFT rather than Micro Electromechanical switch, the combo switch can pass sufficient amount of current.

Therefore, these advantages can change the combo-switch or Micro Electromechanical—MOSEFT switch into a suitable solution for the practical development of energy collection circuits. In addition, it is a suitable alternative in the wireless sensor networks when the sensor nodes are in 99% standby mode most of the time and energy collection is limited.

References

1. Ouyang Y, He J, Hu J, Wang SX (2012) A current sensor based on the giant magnetoresistance effect: design and potential smart grid applications. Sensors 12(11): 15520–15541
2. Han J, Hu J, Ouyang Y, Wang SX, He J (2015) Hysteretic modeling of output characteristics of giant magnetoresistive current sensors. IEEE Trans Industr Electron 62(1):516–524

3. Dallago E, Danioni A, Marchesi M, Nucita V, Venchi G (2011) A self-powered electronic interface for electromagnetic energy harvester. IEEE Trans Power Electron 26(11):3174–3182
4. Jiang X, Polastre J, Culler D (2005) Perpetual environmentally powered sensor networks. In: IPSN 2005. Fourth international symposium on information processing in sensor networks. IEEE, New York, pp 463–468
5. Torah R, Glynne-Jones P, Tudor M, O'Donnell T, Roy S, Beeby S (2008) Self-powered autonomous wireless sensor node using vibration energy harvesting. Measur Sci Technol 19(12):125202
6. Gungor VC, Hancke GP (2009) Industrial wireless sensor networks: challenges, design principles, and technical approaches. IEEE Trans Industr Electron 56(10):4258–4265
7. Park C, Chou PH (2006) Ambimax: autonomous energy harvesting platform for multi-supply wireless sensor nodes. In: 2006 3rd annual IEEE communications society on sensor and Ad Hoc communications and networks, vol 1. IEEE, New York, pp 168–177
8. De Brito MAG, Galotto L, Sampaio LP, de Azevedo e Melo G, Canesin CA (2013) Evaluation of the main MPPT techniques for photovoltaic applications. IEEE Trans Industr Electron 60(3):1156–1167
9. Weddell AS, Grabham NJ, Harris NR, White NM (2009) Modular plug-and-play power resources for energy-aware wireless sensor nodes. In: 2009 6th annual IEEE communications society conference on sensor, mesh and Ad Hoc communications and networks. IEEE, New York, pp 1–9
10. Tan YK, Panda SK (2011) Energy harvesting from hybrid indoor ambient light and thermal energy sources for enhanced performance of wireless sensor nodes. IEEE Trans Industr Electron 58(9):4424–4435
11. Magno M, Jackson N, Mathewson A, Benini L, Popovici E (2013) Combination of hybrid energy harvesters with MEMS piezoelectric and nano-Watt radio wake up to extend lifetime of system for wireless sensor nodes. In: Proceedings of 2013 26th international conference on architecture of computing systems (ARCS). VDE, pp 1–6
12. Moser C, Thiele L, Brunelli D, Benini L (2008) Robust and low complexity rate control for solar powered sensors. In: 2008 design, automation and test in Europe. IEEE, New York, pp 230–235
13. Moser C, Brunelli D, Thiele L, Benini L (2006) Real-time scheduling with regenerative energy. In: 18th Euromicro conference on real-time systems (ECRTS'06). IEEE, New York, p 10
14. Weddell, AS, Magno M, Merrett GV, Brunelli D, Al-Hashimi BM, Benini L (2013) A survey of multi-source energy harvesting systems. In: Proceedings of the conference on design, automation and test in Europe. EDA Consortium, pp 905–908
15. Morais R, Matos SG, Fernandes MA, Valente ALG, Soares SFSP, Ferreira PJSG, Reis MJCS (2008) Sun, wind and water flow as energy supply for small stationary data acquisition platforms. Comput Electron Agric 64(2):120–132
16. Amirtharajah R, Chandrakasan AP (1998) Self-powered signal processing using vibration-based power generation. IEEE J Solid-State Circ 33(5):687–695
17. Mathúna CÓ, O'Donnell T, Martinez-Catala RV, Rohan J, O'Flynn B (2008) Energy scavenging for long-term deployable wireless sensor networks. Talanta 75(3): 613–623
18. Kong N, Ha DS (2012) Low-power design of a self-powered piezoelectric energy harvesting system with maximum power point tracking. IEEE Trans Power Electron 27(5):2298–2308
19. Szarka GD, Stark BH, Burrow SG (2012) Review of power conditioning for kinetic energy harvesting systems. IEEE Trans Power Electron 27(2):803–815
20. Cheng S, Sathe R, Natarajan RD, Arnold DP (2011) A voltage-multiplying self-powered AC/DC converter with 0.35-V minimum input voltage for energy harvesting applications. IEEE Trans Power Electron 26(9): 2542–2549
21. Sun Y, Hieu NH, Jeong C-J, Lee S-G (2012) An integrated high-performance active rectifier for piezoelectric vibration energy harvesting systems. IEEE Trans Power Electron 27(2): 623–627

22. Chung G-B, Ngo KDT (2005) Analysis of an AC/DC resonant pulse power converter for energy harvesting using a micro piezoelectric device. J Power Electron 5(4):247–256
23. Roundy S, Wright PK, Rabaey JM (2003) Energy scavenging for wireless sensor networks. Norwell
24. Kim J-M, Park J-H, Baek C-W, Kim Y-K (2004) The SiOG-based single-crystalline silicon (SCS) RF MEMS switch with uniform characteristics. J Microelectromech Syst 13(6): 1036–1042
25. Shenck NS, Paradiso JA (2001) Energy scavenging with shoe-mounted piezoelectrics. IEEE Micro 21(3):30–42

Phase Balancing in Distribution Network Using Harmony Search Algorithm and Re-phasing Technique

Saeid Eftekhari and Mahmoud Oukati Sadegh

Abstract A large part of the power losses is generated in the distribution network, partly due to the load unbalance in the network. Often, the number of splits is not the same in different phases, and even in the case of equilibrium, due to the difference in the behaviour of the consumers, the unbalance in the network phases are observed and for this reason the distribution network is generally an unbalanced network. The unbalance of the distribution network and the flow of the current in the neutral wire will result in various consequences such as increased power losses, voltage drop, unbalance of three-phase voltages and ultimately consumer dissatisfaction. To reduce this unbalance, various methods and algorithms are presented. In this paper, the re-phasing method using the Harmony Search Algorithm is used to balance the phases in the distribution network. In addition, the proper method of load flow in unbalanced distribution networks is expressed, the selection of the objective function in the re-phasing process is examined and the most appropriate objective function is introduced. In order to demonstrate the efficiency of this method, the simulation for the unbalanced 25-bus network is implemented and compared with the results of other proposed methods. The results show that the proposed method has a very good performance.

Keywords Distribution network · Loss reduction · Re-phasing
Harmony search algorithm

1 Introduction

Unbalance in the distribution network causes zero and negative current components, which will have adverse effects on the distribution network. The zero component will cause additional losses in the network and interfere with the

S. Eftekhari · M. O. Sadegh (✉)
Department of Electrical and Electronic Engineering,
University of Sistan and Baluchestan, Zahedan, Iran
e-mail: oukati@ece.usb.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_43

performance of network protection systems. The negative component of the current also causes transformers to saturate, motor warming, distortion in the operation of the rectifiers, and instability in generators. So far, many studies have been done on the balancing of loads in distribution networks. In [1], a statistical estimation method with the ability to detect the level, position, and unbalanced voltage effects of distribution networks is presented. In [2], the effects of voltage unbalance on induction motors are investigated. In [3], a number of unbalance indicators were introduced and compared. In [4], researchers have been using the Fuzzy Logic Composition and the Newton-Raphson method to address the balancing problem in the distribution network. In [5], an optimal load flow method is proposed that can minimize the unbalanced voltage of the system by using voltage control equipment and other network components. In [6], the reconfiguration is used to transfer the load from heavy load feeders to light load feeders and to load balances in an unbalanced distribution network. The authors in the Ref. [7] have presented a genetic algorithm based approach that can solve the phase balancing problem with several other objective functions in the weighted shape.

In [8], power electronics devices such as SVC to compensate for reactive power, and in Ref. [9], fuzzy logic and hybrid optimization are used to balance the phases in distribution systems. The application of distributed generation, location and optimal size of it in an unbalanced distribution network is investigated in [10]. One of the main methods of phase balancing is the phase displacement method (rephrasing), which was first introduced in [11] as a hybrid integer programming method. Due to its long calculation time and the inadequacy of this technique for large networks, the researchers solved the problem of phase displacement using intelligent methods such as simulated annealing [12]. In [13], an innovative method called Backtracking Search is used for the phase displacement problem in order to balance the distribution network. The authors in [14] designed an expert system to apply re-phasing method to balance the phases in distribution systems. In [15], the authors compared several intelligent algorithms for the phase shift problem and concluded that the dynamic programming algorithm performs better. In [16], the phase displacement method in radial and mesh distribution networks is applied using the BF-PSO algorithm. In Ref. [17], taking into account the unbalanced phase current a safety-based algorithm is also proposed to balance the phases. As can be seen in the above, intelligent methods are very effective in solving the re-phasing problem. Harmony Search Optimization is one of the most successful and effective ways to solve optimization problems that has been widely used by researchers. The efficiency of this method has led to its widespread use in issues related to power systems and distributed networks. For this reason, in this paper, the re-phasing method using the Harmony Search Algorithm is used to balance the phases in the distribution network. Also, to identify the effective objective function in the optimization process, two objective function, which are RMSI and network losses are used, respectively, and the results are compared and appropriate objective function is introduced.

2 Re-phasing Process

In this paper, re-phasing method is used to perform phase balancing. So that for each bus is a key that can transfer the load from one phase to another. The general scheme of this approach is presented in Fig. 1. There are six modes for phase location, each of which is assigned numbers 1 to 6, mode 1 (ABC) mode 2 (ACB) mode 3 (BAC) mode 4 (BCA) mode 5 (CAB) and mode 6 (CBA). The phase change matrices for each mode are given in Table 1.

The re-phasing process for a n-buses network leads to the determination of the vector S = [S1, S2, … Sn] as the key vector that determines the status of the keys to move the phases. Each member of this vector is identified by one to six for example, S = [1 1 5 1 2 4 1 2 3 4 5 5 2 2 5 1 4 6 3 4 1 5 4 1 3] Can be an answer vector for a 25-bus network. The number 1 in the first component of the vector means that the phase arrangement in the first bus should be in the form of ABC or the first mode, or the number 3 in the twenty fifth component requires the phase arrangement to be BAC or the third mode in 25-th bus. So for the bus k, we can consider the following relation for loads:

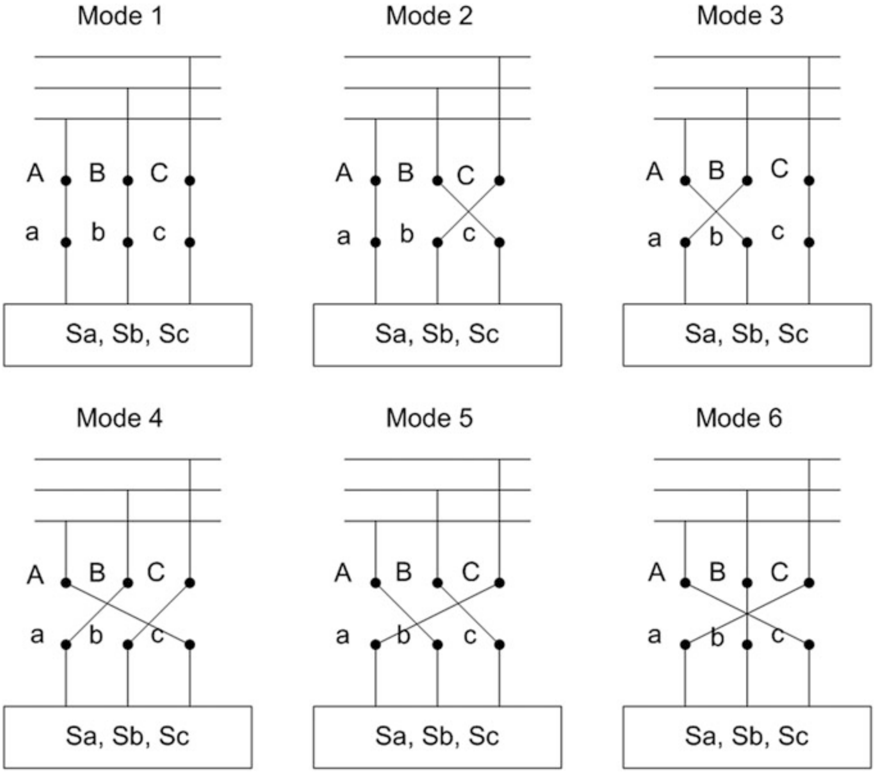


Fig. 1 Six possible modes for re-phasing the phases

Table 1 The phase change matrices for each mode

Mode 1	Mode 2	Mode 3
$s_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; (ABC)	$s_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$; (ACB)	$s_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 0; (BAC)
Mode 4	**Mode 5**	**Mode 6**
$s_4 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$; (BCA)	$s_5 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$; (CAB)	$s_6 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$; (CBA)

$$P_k + jQ_k = S_k(P_{0k} + jQ_{0k}) \tag{1}$$

Then, performing the load flow, the objective function is calculated for the re-phasing process using the following equation:

$$RMSI = \frac{\sqrt{|I_0|^2 + |I_2|^2}}{|I_1|} \tag{2}$$

$$\begin{bmatrix} \mathrm{I}_0 \\ \mathrm{I}_1 \\ \mathrm{I}_2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} \mathrm{I_a} & \mathrm{I_b} & \mathrm{I_c} \\ \mathrm{I_a} & \mathrm{aI_b} & \mathrm{a^2I}_c \\ \mathrm{I_a} & \mathrm{a^2I_b} & \mathrm{aI_c} \end{bmatrix} \tag{3}$$

where I_0, I_1 and I_2 are the zero, positive, and negative components of the current calculated from Eq. 3 (in which a = e^{j120}). *RMSI* is root mean square of current and is called unbalance index. As the index is closer to zero, the network is less unbalance [3].

Another objective function used in this paper is the network losses that are calculated from the following relationships.

$$\mathrm{P_{LOSS}} = \sum_{\mathrm{m}=1}^{3} \mathrm{p^m} \tag{4}$$

$$\mathrm{p^m} = \sum_{\substack{\mathrm{i}=1 \\ \mathrm{j}=1}}^{\mathrm{n}} \mathrm{real}((\mathrm{V_i^m} - \mathrm{V_j^m}) \times (\mathrm{I_{i,j}^m})^*) \tag{5}$$

where $\mathrm{P_{LOSS}}$ is the total loss of the network, $\mathrm{p^m}$ is the loss of m-th phase, where m$\in$ (a, b, c). $\mathrm{V_i}$ and $\mathrm{V_j}$ are the bus voltage of i-th and j-th buses. $\mathrm{I_{ij}}$ is the branch current connected to the bus, i and j, n is the number of network buses.

3 Unbalance Network Load Flow

In this paper, the so-called backward forward sweep load flow method is used which calculates the voltages in each bus and currents in each branch. The superiority of this approach is in ensuring its speed and convergence [18]. The backward sweep is used to obtain the relationship between the branches current and the injection current of the buses, and the forward sweep uses the branches current to calculate the buses voltage. The steps are as follows:

1. Calculate the bus current

$$\begin{bmatrix} I_{ia} \\ I_{ib} \\ I_{ic} \end{bmatrix}^{(k)} = \begin{bmatrix} (S_{ia}/V_{ia})^* \\ (S_{ib}/V_{ib})^* \\ (S_{ic}/V_{ic})^* \end{bmatrix}^{(k)} \tag{6}$$

In which I_i, is the injected current of the i-th bus to the load, $\mathrm{S_i}$ is the apparent power in the i-th bus, and V_i is the voltage of the i-th bus in k-th iteration.

2. Calculate the flow of lines (sweep back)

$$\begin{bmatrix} J_{la} \\ J_{lb} \\ J_{lc} \end{bmatrix}^{(k)} = \begin{bmatrix} I_{ia} \\ I_{ib} \\ I_{ic} \end{bmatrix}^{(k)} + \sum_{m \in M} \begin{bmatrix} I_{ma} \\ I_{mb} \\ I_{mc} \end{bmatrix}^{(k)} \tag{7}$$

In which J_i shows the current flowing in i-th branch and M, the set of lines that feed on the node i.

3. Calculating bus voltages (forward sweep)

$$\begin{bmatrix} V_{ia} \\ V_{ib} \\ V_{ic} \end{bmatrix}^{(k)} = \begin{bmatrix} V_{ja} \\ V_{jb} \\ V_{jc} \end{bmatrix}^{(k)} - \begin{bmatrix} Z_{aa} & \mathrm{Z_{ab}} & \mathrm{Z_{ac}} \\ \mathrm{Z_{ba}} & \mathrm{Z_{bb}} & \mathrm{Z_{bc}} \\ \mathrm{Z_{ca}} & \mathrm{Z_{cb}} & \mathrm{Z_{cc}} \end{bmatrix} \begin{bmatrix} J_{la} \\ J_{lb} \\ J_{lc} \end{bmatrix}^{(k)} \tag{8}$$

where $\mathrm{Z_{aa}}$, $\mathrm{Z_{bb}}$, $\mathrm{Z_{cc}}$ are self-impedance and $\mathrm{Z_{ab}}$, $\mathrm{Z_{bc}}$, $\mathrm{Z_{ca}}$ are mutual impedances of the lines.

The above steps will continue until the following convergence condition is reached:

$$\mathrm{V_i^{(k)}} - \mathrm{V_i^{(k-1)}} < \varepsilon$$

4 Harmony Search (HS) Algorithm

The Harmony Search Algorithm is an optimization algorithm developed in 2001 [19]. One of the simplest and most recent meta-heuristic methods which is inspired from the process of playing simultaneously the music orchestra group. The Harmony Search algorithm has become one of the most used optimization algorithms in recent years due to its applicability for discrete and continuous optimization problems, low mathematical calculations, simple concepts, low parameters, and easy implementation. This algorithm has less mathematical requirements than other meta- heuristic methods, and can be adapted to different engineering issues with changes in parameters and operators. Harmony Search algorithm parameters include Harmony Memory Size (HMS), Harmony Memory Consideration Rate (HMCR) and Pitch Adjustment Rate (PAR) and Bandwidth (BW). This algorithm uses all of its memory solutions. The HS algorithm consists of five steps:

1. Initializing the optimization problem and parameters: At this stage, the optimization problem is first defined:

Minimize $f(x)$
Subject to $g(x) \geq 0$

In these relations, $f(x)$ is the objective function and $g(x)$ is the problem constraint. The next step at this stage is to define the parameters of the algorithm, in which the parameters of the HMS, HMCR, PAR, BW are set at this stage.

2. Initialize Harmony Memory: At this point, the harmony memory is set to:

$$HM = \begin{bmatrix} X_1^1 & \mathrm{X}_2^1 & \dots & \mathrm{X}_N^1 \\ X_1^2 & \mathrm{X}_2^2 & \dots & \mathrm{X}_N^2 \\ \vdots & \vdots & \dots & \vdots \\ X_1^{HMS} & \mathrm{X}_2^{HMS} & \dots & \mathrm{X}_N^{HMS} \end{bmatrix} \tag{9}$$

HMS is the size of the harmony memory, or the number of harmonies in memory, and *N* is the number of variables for each harmony.

3. Create an Improved New Harmony: To generate a new harmony vector, if r_1 is smaller than the HMCR value and the random value r_2 is greater than PAR, then:

$$X_{new}(i) = X_{old}(i) \tag{10}$$

If r_1 is smaller than HMCR and r_2 is smaller than PAR, then:

$$X_{new}(i) = X_{old}(i) + r_3 \times BW \tag{11}$$

If r_1 is larger than *HMCR*, then for X_{new} (i), a random value is considered within its permitted range. r_1, r_2 and r_3 are random numbers.

4. Updating Harmony Memory: In the process of updating the harmony memory, if the harmonies or new harmonies are more competent than the worst harmony in the memory, they replace it, otherwise they will be set aside.
5. Repeat steps 3 and 4 until the final condition is satisfied or repetitions are ended.

5 Calculations Results

To carry out the balancing process, a 25-bus test system, the information of which is presented in [6] and shown in Fig. 2, has been used. To apply the harmony search algorithm to the process of re-phasing of the 25-bus network, first the vector S = [S_1, S_2, … S_{25}] is considered as the initial amount of harmony memory, where S_i is one of the six modes of phase shift keys that are shown in Fig. 1. The parameters of HS algorithm are chosen as: maximum iteration = 1000, HMS = 10, HMCR = 0.9, PAR = 0.1 and BW = 1. The base bus voltage is considered to be 1.05 pu. Base apparent power and base voltage are respectively 100 KVA and 2.4 kV respectively. The process of balancing is done on the network. Figure 3 shows the optimization process of objective function. Table 2 shows the network load before and after the balancing process. According to the results, the loads supplied by the three phases are balanced in a satisfactory manner.

Table 3 shows the values of the unbalance index and the current components before and after the re-phasing. Before re-phasing, the unbalance index is equal to

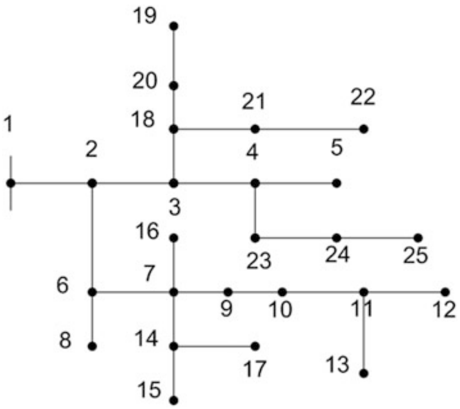


Fig. 2 25-bus network [6]

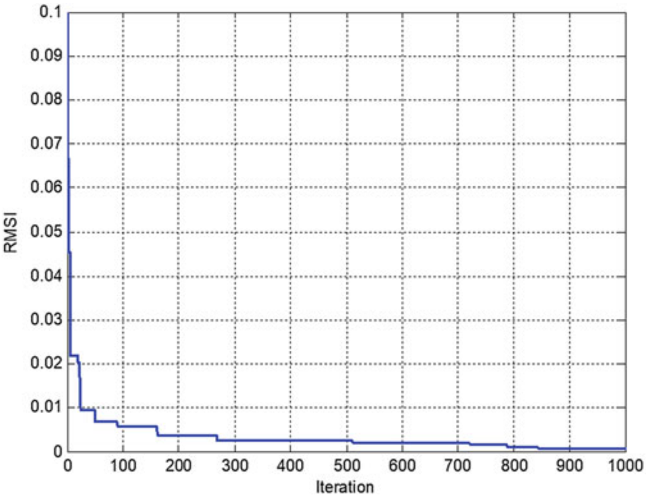


Fig. 3 The convergence process of objective function

0.194, which is equal to 0.0006 after the re-phasing. The zero component of the current is before the re-phasing of 1.28 pu, which after the re-phasing has decreased to 0.0022. Also, the negative component of the current is prior to the re-phasing of 1.27 pu, which is reduced to 0.005 after re-phasing. It is noticeable that after the re-phasing, the negative and zero components, as well as the RMSI value, are significantly reduced. Table 4 shows network currents before and after re-phasing. Before re-phasing due to load unbalance in the grid, different flows are drawn from each phase, but after re-phasing with the balanced distribution of loads on the three phases, the flows drawn from the phases are equal.

Prior to re-phasing a phase current is 11.4 pu, the b phase current is 6.99 pu and the c phase current 9.43 pu. After re-phasing the flow of a, b and c phases are equal to 9.26, 9.26, and 9.25 pu respectively. Also, before the re-phasing due to the inequality of the flow of the phases, the flow of the current in neutral wire is equal to 3.84 pu, which is a significant amount. After re-phasing, the neutral wire current is equal to 0.006 pu. Table 5 shows the active losses of the system before and after re-phasing. The network losses before re-phasing were 68.48 kw. As it can be seen, the active losses have fallen to 65.81 kw after re-phasing.

In this network, the capacity of the main bus transformer is 1100 kV for each phase. S_a represents the amount of apparent power drawn from each phase of the main bus transformer. Before re-phasing, the main transformer core capacity is not used equally between phases. For example, phase a has been loaded more than existing capacity and other phases less than existing capacity. After balancing, it is observed that the main bus transformer loading between the phases is more balanced.

S_{margin}, represent the margin of difference of loading on the main bus transformer and its nominal capacity (1100 kV) on each phase ($S_{margin} = 1100 - S_a$).

Table 2 Network loads before and after balancing for the 25-bus network

Bus No.	Before balancing			After balancing		
	$P^a{}_d + jQ^a{}_d$	$P^b{}_d + jQ^b{}_d$	$P^c{}_d + jQ^c{}_d$	$P^a{}_d + jQ^a{}_d$	$P^b{}_d + jQ^b{}_d$	$P^c{}_d + jQ^c{}_d$
1	00.00+00.00i	00.00+00.00i	00.00+00.00i	00.00+00.00i	00.00+00.00i	00.00+00.00i
2	00.00+00.00i	00.00+00.00i	00.00+00.00i	00.00+00.00i	00.00+00.00i	00.00+00.00i
3	36.00+21.60i	28.80+19.20i	42.00+26.40i	28.80+19.20i	36.00+21.60i	42.00+26.40i
4	57.60+43.20i	04.80+03.36i	48.00+30.00i	48.00+30.00i	57.60+43.20i	04.80+03.36i
5	43.20+28.80i	28.80+19.20i	36.00+24.00i	28.80+19.20i	36.00+24.00i	43.20+28.80i
6	43.20+28.80i	33.60+24.00i	30.00+30.00i	33.60+24.00i	30.00+30.00i	43.20+28.80i
7	00.00+00.00i	00.00+00.00i	00.00+00.00i	00.00+00.00i	00.00+00.00i	00.00+00.00i
8	43.20+28.80i	28.80+19.20i	03.60+02.40i	28.80+19.20i	43.20+28.80i	03.60+02.40i
9	72.00+50.40i	38.40+28.80i	48.00+30.00i	72.00+50.40i	48.00+30.00i	38.40+28.80i
10	36.00+21.60i	28.80+19.20i	42.00+26.40i	36.00+21.60i	28.80+19.20i	42.00+26.40i
11	50.40+31.68i	24.00+14.40i	36.00+24.00i	24.00+14.40i	50.40+31.68i	36.00+24.00i
12	57.60+36.00i	48.00+33.60i	48.00+36.00i	57.60+36.00i	48.00+33.60i	48.00+36.00i
13	64.80+21.60i	33.60+21.12i	36.00+24.00i	33.60+21.12i	64.80+21.60i	36.00+24.00i
14	57.60+36.00i	38.40+28.80i	60.00+42.00i	57.60+36.00i	38.40+28.80i	60.00+42.00i
15	07.20+04.32i	04.80+02.88i	06.00+03.60i	07.20+04.32i	04.80+02.88i	06.00+03.60i
16	57.60+04.32i	03.84+28.80i	48.00+36.00i	57.60+04.32i	03.84+28.80i	48.00+36.00i
17	57.60+43.20i	33.60+24.00i	54.00+38.40i	33.60+24.00i	57.60+43.20i	54.00+38.40i
18	57.60+43.20i	38.40+28.80i	48.00+36.00i	48.00+36.00i	38.40+28.80i	57.60+43.20i
19	50.40+36.00i	38.40+28.80i	54.00+38.40i	38.40+28.80i	50.40+36.00i	54.00+38.40i
20	08.64+06.48i	04.80+03.36i	06.00+04.80i	06.00+04.80i	04.80+03.36i	08.64+06.48i

(continued)

Table 2 (continued)

Bus No.	Before balancing			After balancing		
	$P^a{}_d + jQ^a{}_d$	$P^b{}_d + jQ^b{}_d$	$P^c{}_d + jQ^c{}_d$	$P^a{}_d + jQ^a{}_d$	$P^b{}_d + jQ^b{}_d$	$P^c{}_d + jQ^c{}_d$
21	05.76+04.32i	03.36+02.40i	05.40+03.84i	03.36+02.40i	05.40+03.84i	05.76+04.32i
22	72.00+50.40i	57.60+43.20i	60.00+48.00i	57.60+43.20i	60.00+48.00i	72.00+50.40i
23	08.64+64.80i	04.80+03.84i	60.00+42.00i	08.64+64.80i	04.80+03.84i	60.00+42.00i
24	50.40+36.00i	43.20+30.72i	04.80+03.60i	50.40+36.00i	43.20+30.72i	04.80+03.60i
25	08.64+06.48i	04.80+02.88i	06.00+04.20i	06.00+04.20i	08.64+06.48i	04.80+02.88i
Total load	946.08+648i	573.6+430i	781.8+554i	765.60+543i	763.08+548.40i	772.80+540.24i

Table 3 The current components and the unbalance index before and after the re-phasing

	I_0(pu)	I_1(pu)	I_2(pu)	RMSI
Before the re-phasing	1.28	9.27	1.27	0.194
After the re-phasing	0.0022	9.26	0.005	0.006

Table 4 The flow of phases and the current of the neutral wire before and after the re-phasing

	I_a(pu)	I_b(pu)	I_c(pu)	I_n(pu)
Before the re-phasing	11.4	6.99	9.43	3.84
After the re-phasing	9.26	9.26	9.25	0.006

Table 5 Network losses before and after re-phasing

Losses	(kW)P_{LOSS}
Before the re-phasing	68.48
After the re-phasing	65.81

Before balancing, this amount is negative for the a phase, which indicates that phase a has been loaded more than the limit. For other phases this is a positive amount, which indicates that those phases have not used their maximum capacity. According to Table 6, before the balancing, the capacities used in phase a, b and c, is equal to 1197, 734 and 991 KVA, respectively, which indicates that an additional 97 KVA of phase a has been drained over its capacity.

After balancing, as shown in Table 6, the S_a values of each phase are partially equalized, so that S_{margin} is equal to each phase. The amount of capacity drawn from the network phases after the balancing is equal to 972.85, 972.74, and 972.97 KVA, which indicates the proper utilization of the network capacity. Table 5 also shows that the margin of capacity of the phases a, b, and c is 127.15, 127.26, and 128.03 KVA, respectively.

In order to validate the proposed method, in Table 7 the results of the method of this paper are compared with the results of Ref. [20], which is one of the most recent researches in this field. Significant decrease in the currents of the zero and negative components as well as the unbalance index (RMSI) of the network in the proposed method is evident in comparison with the results obtained from the Ref. [20].

Network voltages before and after re-phasing

To illustrate the state of the bus voltages before and after re-phasing, the amount of voltage unbalancing is defined as follows:

Table 6 Loading of each phase of the main bus transformer before and after balancing

	A phase	B phase	C phase	A phase	B phase	C phase
Parameter	Before the re-phasing			After the re-phasing		
S_a(kVA)	1197.6	734	991.2	972.85	972.74	972.97
S_{margin}(kVA)	−97.592	365.28	108.88	127.15	127.26	128.03

Table 7 Comparing the results of the Harmony search algorithm with the Ref. [20]

	P_{LOSS}(kW)	I_0(pu)	I_2(pu)	RMSI
Initial condition	68.48	1.28	1.27	0.194
Using the Ref. [20] method	65.86	0.013	0.026	0.0032
Proposed method	65.81	0.0022	0.005	0.0006

Table 8 Maximum amount of voltage unbalancing before and after re-phasing

max(δVmax)	
Before the re-phasing	After the re-phasing
0.0266	0.007

$$\delta V_{\max} = \max(|V|_{a,b,c}) - \min(|V|_{a,b,c}).$$

The results show that δV_{max} is very large for all buses and the maximum amount of voltage unbalancing [max (δVmax)] is significant before re-phasing. After re-phasing, this amount is noticeably reduced. Table 8 shows the amount of voltage unbalancing before and after re-phasing. According to the Table 8 the maximum amount of voltage unbalancing is equal to 0.0266, which is reduced to 0.007 after re-phasing.

Figure 4 shows the network voltage profile before re-phasing. In this figure, the three-phase voltage difference is quite evident. The phase a, due to the fact that it has a higher load, also has a higher voltage drop and a lower loaded b phase has less voltage drop than the other two phases. In Fig. 5, which shows the network voltage profile after balancing, the voltage unbalancing reduction is noticeable.

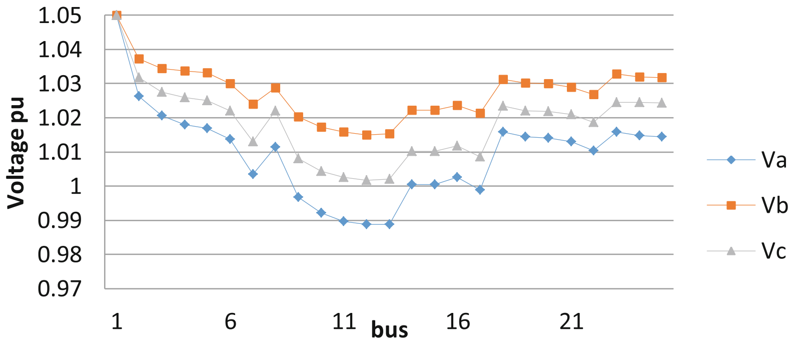


Fig. 4 Voltage profile of the 25-bus network before re-phasing

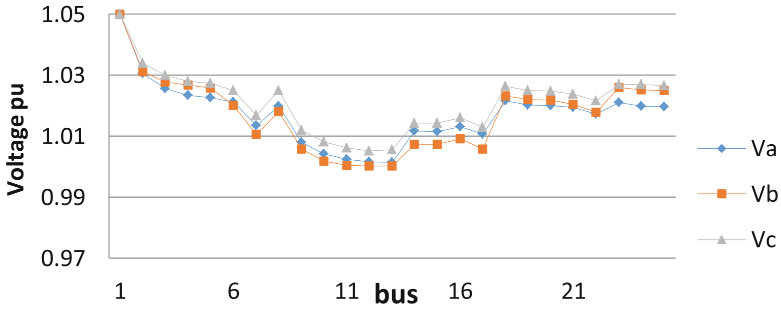


Fig. 5 Voltage profile of the 25-bus network after re-phasing

5.1 *Considering Network Losses as the Objective Function*

Since one of the important goals in re-phasing phases is reducing network losses, the network loss function can also be considered as an objective function in the optimization process. In addition, the equilibrium problem is considered by considering the network loss function as the main objective function and the results are compared with the results of the previous section that are considered in the MRSI index as the target function. Table 9 shows the status of the network keys in the above two steps. The results indicate a significant difference between the keys in the above mentioned conditions. In the next step, the re-phasing problem is solved with the consideration of the network loss function as the main objective function and the results are compared with the results of the previous section, which is considered in the MRSI index as the objective function.

In Table 10, the results of the other network indicators are compared with each other. However, in the latter case, the loss rate is lower than the first one, but this reduction is not significant. Also, the difference in the margin of transformer capacity (S_{margin}) is not significant in the above mentioned cases. However, the difference between the neutral wire current, zero and the negative current and the resulting unbalance index (RMSI) is significant. The results from the first state are better for zero and negative currents as well as for RMSI indexes. In sum, it can be concluded that solving the re-phasing problem with the consideration of the RMSI index as the objective function will lead to better results.

Table 9 The status of the 25-bus switches before and after re-phasing

The status of switches	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Before the re-phasing
1 1 3 5 4 4 1 3 2 1 3 1 3 1 1 1 3 6 3 6 4 4 1 1 5	After re-phasing (objective function: RMSI)
1 1 5 3 1 6 1 3 4 5 3 5 1 1 3 1 5 3 6 3 4 1 1 3 4	After balancing (objective function: network loss function)

Table 10 Comparing the results of selecting the various objective functions

Results after re-phasing	S_{margin}			Losses kw	I_n	I_2	I_0	RMSI
	A phase	B phase	C phase					
Objective function: RMSI	127.15	127.26	128.03	65.81	0.006	0.005	0.0022	0.0006
Objective function: network loss function	126.9410	127.5371	127.9256	65.7952	0.10	0.0401	0.0352	0.0058

6 Conclusion

In this paper, the re-phasing method has been used to carry out the balancing process. This method, by placing the six-state keys on each of the distribution buses, distributes the loads in such a way that the index of unbalance is minimized. Different indexes for unbalanced distribution systems have been defined. In this paper, the RMSI index, the ratio of the sum of the components of the negative and the zero components to its positive component, and network loss function, were used to examine the unbalanced distribution system. To optimize the objective function, the Harmony Search algorithm, due to its applicability for discrete and continuous optimization problems, low mathematical calculations, simple concepts, low parameters, and easy implementation has been used. re-phasing is one of the best methods for phases unbalance reduction, reduction of null current and reduction of negative and zero current components as well as freeing network capacity. By comparison of different objective functions, the choice of RMSI index could lead to better results in the optimization process. Combining and applying the re-phasing method with other existing methods, such as network reconfiguration or the use of distributed generation resources, can lead to better results in reducing network losses.

References

1. Woolley NC, Milanovic JV (2012) Statistical estimation of the source and level of voltage unbalance. IEEE Trans Power Del 28(3):1450–1460
2. Youb L (2014) Effects of unbalanced voltage on the steady state of the induction motors. Int J Electr Energy 2(1)
3. Bina MT, Kashefi A (2011) Three-phase unbalance of distribution systems: complementary analysis and experimental case study. Elect Power Energy Syst 33:817–826
4. Siti WM, Jimoh A, Nicolae D (2011) Distribution network phase load balancing as a combinatorial optimization problem using fuzzy logic and Newton-Raphson. Electr Power Syst Res 81:1079–1087

5. Araujo LR, Penido DRR, Carneiro S, Pereira JLR (2013) A three-phase optimal power-flow algorithm to mitigate voltage unbalance. IEEE Trans Power Del 28(4):2394–2402
6. Vulasala G, Sirigiri S, Thiruveedula R (2009) Feeder reconfiguration for loss reduction in unbalanced distribution system using genetic algorithm. Int J Comput Electr Autom Control Inf Eng 3(4)
7. Chen T-H, Cherng J-T (2000) Optimal phase arrangement of distribution transformers connected to a primary feeder for system unbalance improvement and loss reduction using a genetic algorithm. IEEE Trans Power Syst 15(3):994–1000
8. Quintela FR, Arévalo JMG, Redondo RC, Melchor NR (2011) Four-wire three-phase load balancing with Static VAR Compensators. Elect Power Energy Syst 33:562–568
9. Ukil A, Siti W (2008) Feeder load balancing using fuzzy logic and combinatorial optimization-based implementation. Electr Power Syst Res 78:1922–1932
10. Gupta I, Gupta V (2015) Unbalanced radial distribution system load flow and voltage profile enhancement in the presence of distributed generators. Int J Eng Res Technol 4(5)
11. Zhu J, Chow M-Y, Zhang F (1998) Phase balancing using mixed-integer programming. IEEE Trans Power Syst 13(4):1487–1492
12. Zhu GBJ, Chow M (1999) Phase balancing using simulated annealing. IEEE Trans Power Syst 14(4):1508–1513
13. Lin C-H, Chen C-S, Chuang H-J (2005) Heuristic rule-based phase balancing of distribution systems by considering customer load patterns. IEEE Trans Power Syst 20(2):709–716
14. Lin C-H, Chen C-S, Chuang H-J (2008) An expert system for three phase balancing of distribution feeders. IEEE Trans Power Syst 23(3):1488–1496
15. Wanga K, Skienab S, Robertazzia TG (2013) Phase balancing algorithms. Electr Power Syst Res 96:218–224
16. Hooshmand RA, Soltani S (2012) Fuzzy optimal phase balancing of radial and meshed distribution networks using Bf-PSO algorithm. IEEE Trans Power Syst 27(12):47–57
17. Huang MY, Chen CS, Lin CH, Kang MS (2008) Three-phase balancing of distribution feeders using immune algorithm. IET Gener Transm Distrib 2(3):383–392
18. Murthy KK, Kumar SVJR (2012) Three-phase unbalanced radial distribution load flow method. Int Refereed J Eng Sci 1(1)
19. Geem ZW, Kim JH, Loganathan GV (2001) A new heuristic optimization algorithm: harmony search, vol 76, no 2. Sage, Beverley Hills, pp 60–68
20. Singh D, Misra RK, Mishra S (2016) Distribution system feeder re-phasing considering voltage-dependency of loads. Elect Power Energy Syst 76:107–119

Study on Performance of MPPT Methods in WRSG-Based Wind Turbines Utilized in Islanded Micro Grid

Arash Khoshkalam and Seyed Mohammad Mahdi Moosavi

Abstract Wind energy systems are progressively used in different micro grids. Maximum power point tracking (MPPT) plays a critical role to improve the efficiency of wind energy conversion system (WECS). This Paper provides a comparison between common MPPT methods, including power curve, tip speed ratio (TSR) and optimal torque control (OTC). Then their effects on behavior of an islanded micro grid during and after overload has been investigated. The results indicate that OTC has better behavior than other methods. The studied system is a micro grid with high share of renewable energy resources containing a WECS driven by wound rotor synchronous generator (WRSG), a synchronous diesel unit and some active and reactive loads.

Keywords Micro grid · Wind turbine · Maximum power point tracking (MPPT) Wound rotor synchronous generator (WRSG) · Overloading · Short circuit fault

1 Introduction

In the last decade, due to environmental concerns and reducing source of fossil fuels, a rapid growth of renewable energy systems especially wind turbines was experienced in the world. Integration of wind turbines in power systems encounter new challenges that one of them is congenital variations or unpredictable nature of wind. These variations which affect electrical output power of wind energy conversion system (WECS), is harmful to power system stability [1, 2]. It is obvious that these consequences are more pronounced in isolated micro grids than grid connected WECS [3]. So, voltage and frequency stability is essential for micro

A. Khoshkalam · S. M. M. Moosavi (✉)
Department of Electrical Engineering, Hamedan University
of Technology, Hamedan, Iran
e-mail: moosavi@hut.ac.ir

A. Khoshkalam
e-mail: Arashkhoshkalam91@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_44

grids especially in islanded mode [2, 4–6]. Different methods have been proposed to overcome these problems. While some plans are based on additional hardware such as energy storage systems [1], most of suggested procedures are implemented by proper control of micro grid system. Some examples for the latter are virtual inertia, load shedding, fuzzy control and evolutionary algorithms [7, 8].

Apart from micro grid control, individual control of WECS could affect stability of a micro grid. An important part of WECS control is Maximum Power Point Tracking (MPPT) which is employed to extract maximum available power from the wind at different wind speed [7, 8]. Especially in isolated micro grids, control trend should be very accurate to ensure overall system stability by providing enough power in order to make a balance between generation and demand and prevent power mismatch [9, 10].

One of the major factors affecting the stability of traditional power systems is the voltage dip on grid side. This factor is also very important in the stability of micro grid [11, 12]. In grid-connected micro grid, overload or grid faults cause voltage drop in grid side. This issue leads to imbalance between input and output power of wind turbine and eventually rises the dc link voltage dramatically [13]. In islanded micro grid, control and management is more complex than grid-connected micro grid. In this case, due to lower equivalent physical inertia even small interruptions could give rise to stability issues [14].

Although various MPPT methods have been proposed [4, 5, 17–19], and their stability issues have been studied in grid-connected WECS [13] but their effect on islanded micro grids has been rarely considered.

In this paper after a comparison between common MPPT methods, including power curve, tip speed ratio (TSR), and optimal torque control (OTC), their effects on stability of an islanded micro grid during overload has been investigated. The studied system is a micro grid containing a wind farm driven by wound rotor synchronous generators (WRSG), a synchronous diesel unit and some active and reactive loads.

2 Basic Wind Turbine Concepts

2.1 Wind Energy Extraction

Mechanical power extracted from the wind turbine is given by the following equation [4, 13–18]:

$$P_m = \frac{1}{2}\pi R^2 \rho V_w^3 C_p(\lambda, \beta) \tag{1}$$

where R is rotor's radius (blade length), ρ is the air density in kg/m^3, V_w is the wind speed in m/s, and C_p is the power coefficient of the blade. The power coefficient is a function of rotational speed and number of blades [4, 17]. This parameter is a nonlinear function of λ and β where λ is the tip speed ratio (TSR) and β is the pitch angle.

$$C_p(\lambda, \beta) = \frac{P_m}{P_w} = 0.5\left(116\frac{1}{\lambda_i} - 0.4\beta - 5\right)e^{-(21/\lambda_i)} \tag{2}$$

where λ_i is obtained from:

$$\frac{1}{\lambda_i} = \frac{1}{\lambda + 0.08\beta} - \frac{0.035}{1 + \beta^3} \tag{3}$$

TSR is a very important parameter in WECS and is described as a ratio between tip speed ratio and wind speed given by the following equation:

$$\lambda_T = \frac{\omega_M^* R}{V_w} \tag{4}$$

where ω_M is the rotating speed of the blade. The maximum power coefficient occurs at the optimal tip speed ratio $\lambda_{T,opt}$ with the rated (optimal) pitch angle. The optimal tip speed ratio $\lambda_{T,opt}$ is a constant for a given blade [16]. Turbine speed that gives maximum power generation, is related to $\lambda_{T,opt}$, and wind speed v_w by [8, 17, 18] (Fig. 1):

$$\omega_M = \lambda_{T,opt}\frac{V_w}{R} \tag{5}$$

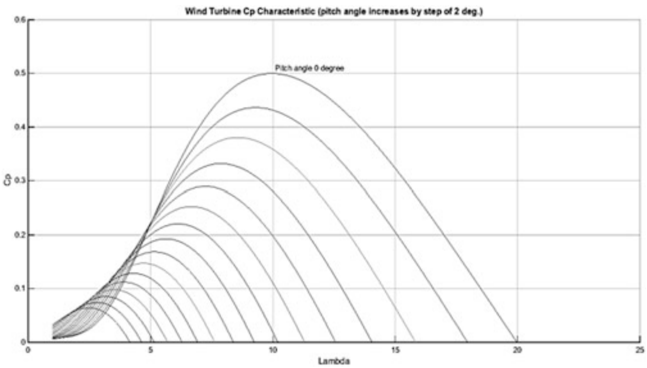


Fig. 1 C_p versus λ_T for used wind turbines

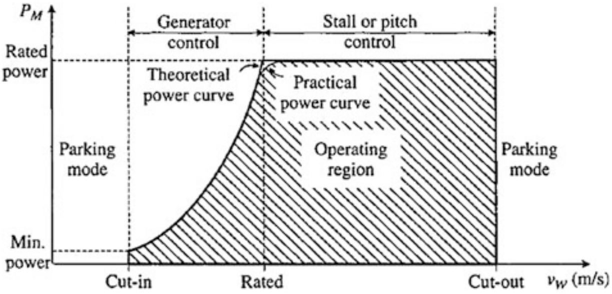


Fig. 2 P_M versus V_w for wind turbines [7]

Equation (5) demonstrates that for maximum power and efficiency, wind turbine speed should be controlled properly according to the wind speed variations.

2.2 *Wind Turbine Power Characteristics*

The power characteristic of a typical wind turbine is surveyed by power curve, that is a relation between mechanical power of turbine and wind speed. A typical power curve is characterized by three wind speeds: cut-in, rated, and cut-out wind speed, as shown in Fig. 2, where P_M is the mechanical power of turbine and V_w is the wind speed. The cut-in wind speed, is the wind speed at which the turbine starts to operate and generate power. The blade should be able to extract enough power for compensation of likely losses and faults. The rated wind speed is the speed at which the turbine produces nominal power. The cut-out wind speed is the highest wind speed at which the turbine is allowed to operate before it is shut down. For wind speeds over the cut-out speed, the turbine should stop its operation by brakes to prevent from eventual damages. To deliver extracted power to the grid at different wind speeds, the wind generator should be controlled appropriately at various speeds [7, 8, 17].

For wind speeds higher than rated speed and lower than cut-out, pitch control method is usually utilized due to good efficiency, proper cost and desired control accuracy [7, 8, 17].

3 Common MPPT Methods

A lot of MPPT methods have been proposed but the most prevalent ones are power curve, tip speed ratio, optimal torque and perturb & observe [4]. A brief description for each technique is introduced in this section.

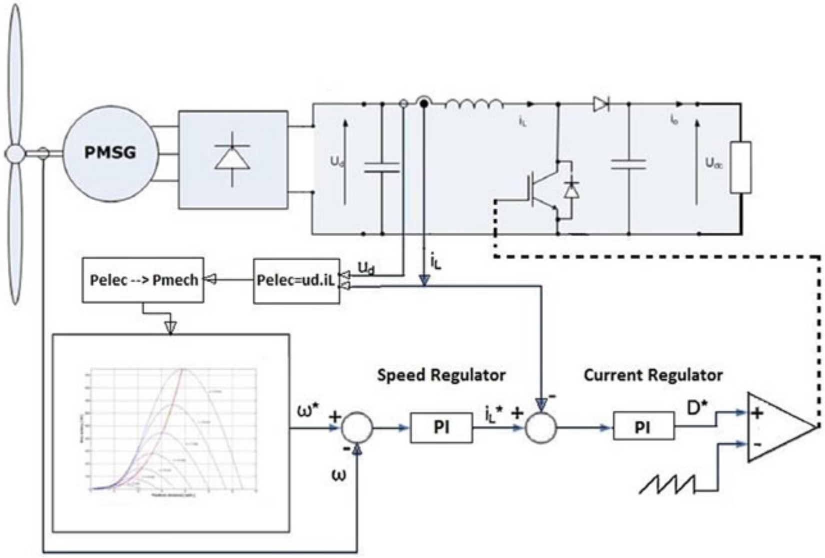


Fig. 3 Block diagram for power curve control method [7, 18]

3.1 MPPT with Power Curve Method

MPPT algorithms based on $P_m(\omega)$ characteristics are well known and widely used. There are two structures for this method. One of them is based on wind speed, which is monitored continuously. Then the reference active power is specified by wind speed [7]. Another structure is based on DC power which is measured in DC link. Then reference speed is determined according to power value. Control diagram of second method is shown in Fig. 3. Maximum power points obtained from experimental results can be saved into a polynomial or a micro controller memory as a lookup table. As illustrated in Fig. 3, the second method needs no anemometers and wind sensors and only DC power measurement is necessary. As precise measurement of wind speed is impractical and increases the system costs [7, 16–20], the second approach is more preferable. This method has fast dynamic response, simple structure and good efficiency [4].

3.2 MPPT with TSR Control

In this method, wind turbine's maximum power is obtained by keeping λ_T at optimum value. By maintaining TSR at optimum value, the extracted energy is maximized. The optimum TSR can be determined experimentally or theoretically

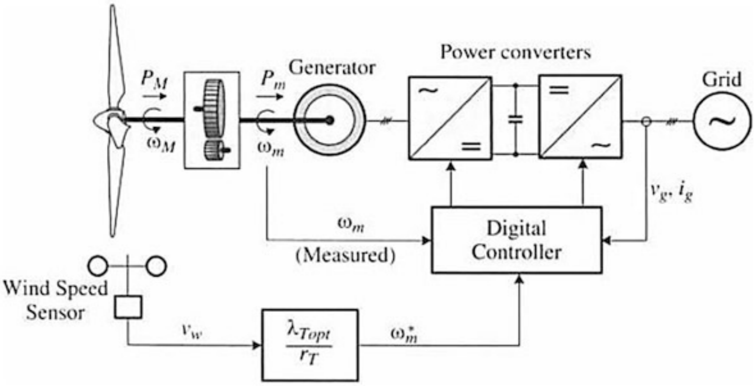


Fig. 4 Block diagram for TSR method [7]

and stored as a reference. Based on $\lambda_{T,opt}$ and wind speed value, desired generator speed (ω_m^*) is specified [18].

$$\omega_m^* = \frac{\lambda_{opt} V}{R} \tag{6}$$

Then, power converter adjusts the generator speed at desired value. This method which depends on exact determination of λ_{opt}, has low flexibility and complexity due to dependence on a constant parameter. Continuous wind speed measurement leads to fast dynamic response [4, 18]. The block diagram of TSR control method is shown in Fig. 4.

3.3 MPPT with Optimal Torque Control (OTC)

In this method, generator mechanical torque T_m is controlled in relation to its speed ω_m to attain maximum power point. It is notable that in this method speed sensors are not required.

With a little mathematical manipulation, mechanical power could be obtained by [18–20]:

$$P_m = \frac{1}{2} \rho \pi R^5 \frac{\omega_m^3}{\lambda^3} C_p \tag{7}$$

If the rotor is running at λ_{opt}, it will also run at $C_{p\ max}$. Thus:

$$P_{m-opt} = \frac{1}{2}\rho\pi R^5 \frac{C_{p,\max}}{\lambda_{opt}^3}\omega_m^3 = K_{p-opt}\omega_m^3 \tag{8}$$

Considering that $P_m = \omega_m T_m$, T_m can be rearranged as follows:

$$T_{m-opt} = \frac{1}{2}\rho\pi R^5 \frac{C_{p,\max}}{\lambda_{opt}^3}\omega_m^2 = K_{opt}\omega_m^2 \tag{9}$$

As shown in Fig. 5, power converter with the aid of proper controller adjusts the generator torque to reference value. This method has a simple structure and doesn't need to wind speed sensors, on the other hand its response and recovery time and also efficiency is absolutely better than other methods [7, 17, 18, 20].

4 Proposed Islanded Micro Grid and Its Response to Overload

4.1 Description of Micro Grid

The studied micro grid includes high share of renewable energy resources. A 10 MW wind farm, involving five WRSG with rated power of 2 MW, supplies the main demands. In addition, a synchronous diesel unit provides 3.125 MVA with power factor equal to 0.8. On the whole, 12.5 MW active power could be delivered to the loads when the wind farm works with maximum capacity (Fig. 6). As the main purpose of this study is only the evaluation of MPPT effect on system performance, it is assumed that the wind speed has a constant value in MPPT region.

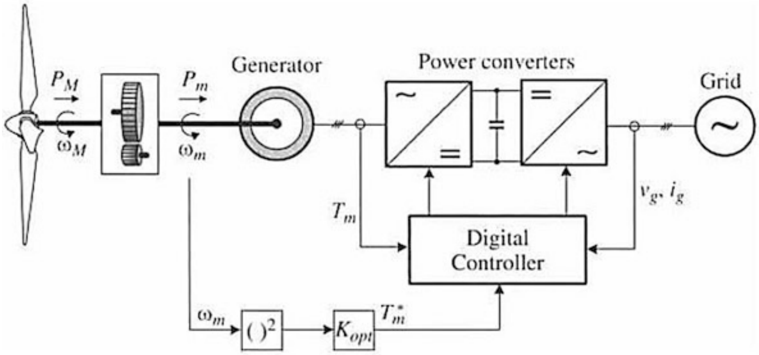


Fig. 5 Block diagram and T_m-ω curve for optimal torque structure [7, 18]—Optimal torque block diagram

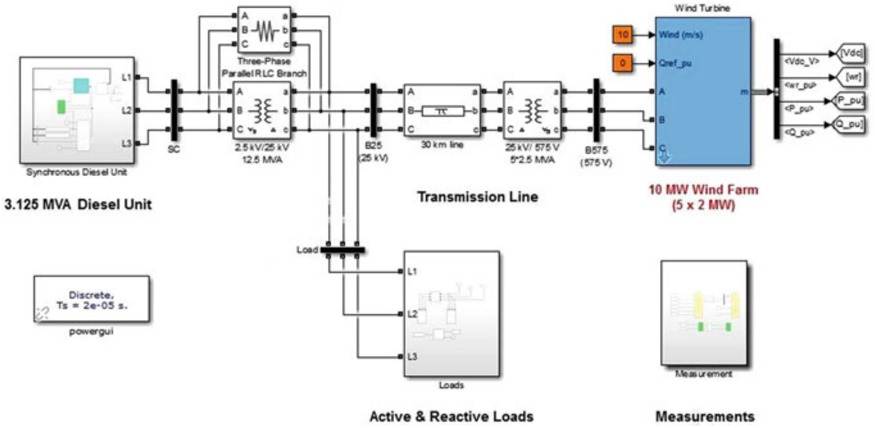


Fig. 6 Block diagram of the micro grid studied in islanding mode

4.2 System Behavior Under Overload

In this section, influence of MPPT methods on micro grid stability is investigated during and after overload. The micro grid loads initially consume 5 MW active power and 1.875 Mvar reactive power. According to wind speed chosen under rated value, the maximum power generated by wind farm in this situation is limited to 5 MW. After a couple of seconds, some loads are added which leads to a power mismatch. The added loads include 2 MW active power so the synchronous diesel unit is forced to operate at maximum capacity. The overloading occurs at t = 7 s when the total load reaches 7 MW active power and 1.875 Mvar reactive power.

Some major parameters such as active power, reactive power, DC link voltage and frequency are considered here. Figure 7a shows the active power. As illustrated, the settling time in OTC is shorter than two other methods and steady state is attained sooner. While in power curve method overload leads to remarkable power oscillations, TSR and OTC methods haven't been obviously impressed by this disturbance. In Fig. 7b reactive power is displayed. To achieve unity power factor, its value has been set to zero. In this figure, again oscillations in power curve method is much more severe and after overload reactive power does not track the desired value. But, in other methods, the set point has been followed.

Figure 7c presents DC link voltage which has the reference value of 1100 V. As illustrated, in TSR and OTC methods the level of voltage has been maintained after some oscillations caused by overloading. However, in power curve method voltage level has increased after overloading. This phenomenon is mainly due to dependence of power set point to DC power which is highly affected by terminal disturbances.

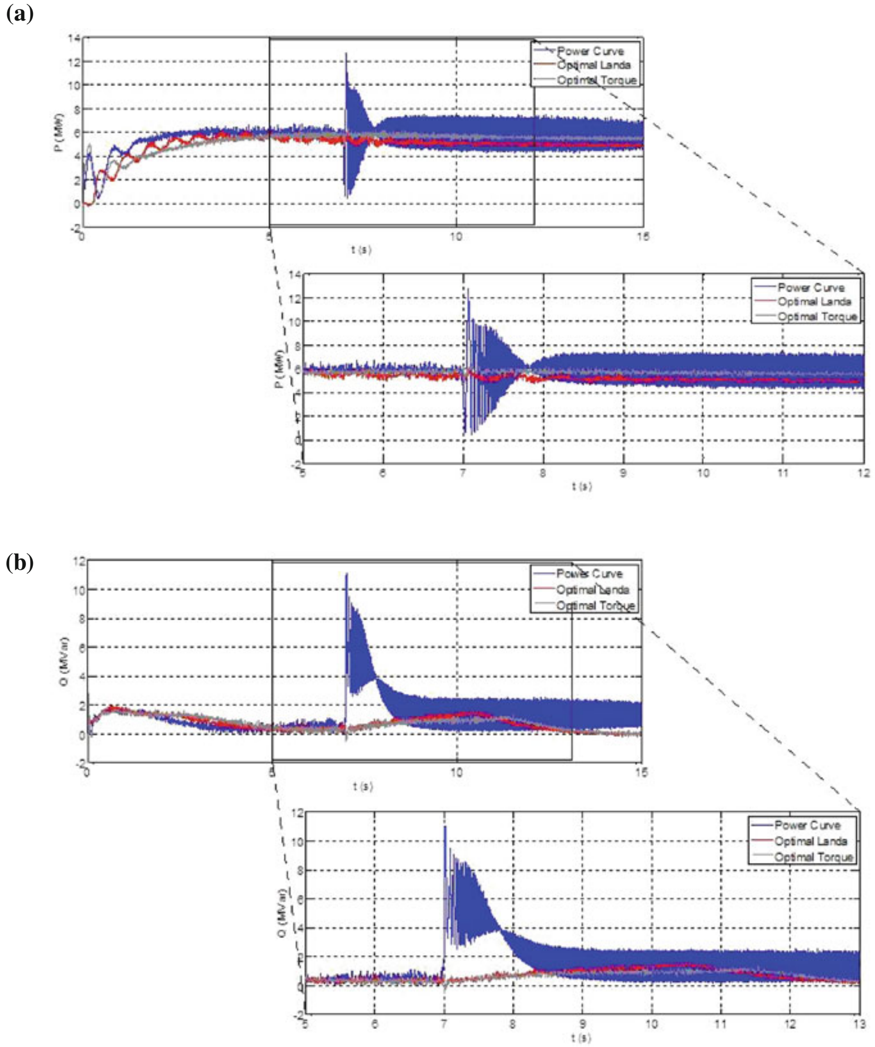


Fig. 7 Simulation results on micro grid in overloading scenario—**a** wind farm active power **b** wind farm reactive power **c** wind turbine's DC link voltage **d** Micro grid overall frequency

Eventually, Fig. 7d demonstrates the system frequency. According to this figure, in OTC method, the transient state elapses faster with lower overshoots. Nevertheless, final values are close to each other in all methods.

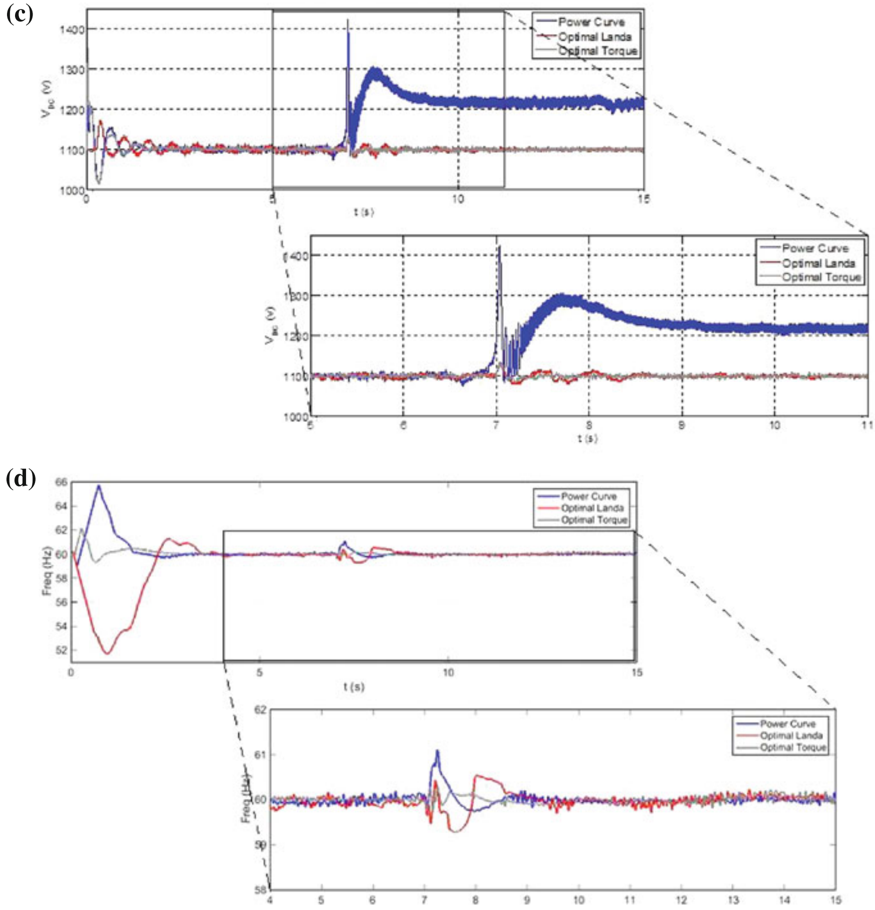


Fig. 7 (continued)

Table 1 Comparison between MPPT methods

	Complexity	Performance	Overshoot	Settling time	Memory required	Wind speed measurement
Power curve	High	Medium	High	Good	Yes	No
TSR	Simple	Medium	Low	Good	No	Yes
OTC	Medium	High	Low	Very good	No	No

4.3 Discussion on Results

It is inferred from the simulation results that OTC is more optimized than other methods. Each of them has their special features which can be useful in different conditions. Settling time in TSR is slightly better than Power curve method while OTC is the best. Overshoot is more pronounced in power curve method compared to other approaches. Due to needed lookup table in power curve method some memory units should be considered which is not mandatory in other methods. Although wind speed sensor is required in TSR method, it is not utilized in presented power curve and OTC. Comparison between studied MPPT methods has been included in Table 1.

5 Conclusion

In this paper, effect of MPPT methods on an isolated micro grid performance during overload has been studied. Considered micro grid included a wind farm with high share of active power injection, a synchronous diesel unit and consumption loads. Common MPPT methods including power curve, TSR and OTC have been introduced and utilized in the micro grid.

Based on simulation results, OTC shows better performance during overload compared to other methods. Regarding its simplicity and low cost besides transient performance, this method is determined as the best among other ones.

Appendix

Parameters of system

Wind Farm	
Turbine data for 1 wind turbine	
Nominal mechanical output power (W)	2 × 10^6 W
Wind speed at nominal speed and at C_p max (6–30 m/s)	11 m/s
Drive train data for 1 wind turbine	
Wind turbine inertia constant H (s)	4.32 s
Turbine initial speed (pu of nominal speed)	1 p.u
Initial output torque (pu of nominal mechanical torque)	1 p.u
Generator data for 1 wind turbine	
Nom. power, L–L volt. and freq.: [P_n (VA), V_n (V_{rms}), f_n (Hz)]	[2.22 × 10^6 730,60]

(continued)

(continued)

Wind Farm	
Reactances [X_d $X_d^{'}$ $X_d^{''}$ X_q $X_q^{''}$ X_l] (p.u)	[1.305, 0.296, 0.252, 0.474, 0.243, 0.18]
Time constants [$T_{do}^{'}$ $T_{do}^{''}$ $T_q^{''}$] (s)	[4.49, 0.0681, 0.0513]
Resistance Rs (p.u)	0.006 p.u
Inertia constant, friction factor, and pairs of poles: [H(s) F (p.u) p]	[0.62, 0.01, 1]
Converters data for 1 wind turbine	
Grid-side converter nominal AC voltage (V)	575 V
Grid-side converter maximum AC current (p.u)	1.1 p.u
Grid-side coupling inductor [L(p.u) R(p.u)]	[0.15, 0.003] p.u
Line filter capacitor (Q = 50) (var)	150×10^3 var
Nominal DC bus voltage (V)	1100 V
DC bus capacitor (F)	90×10^{-3} F
Boost converter inductance [L(H) R(ohm)]	[0.0012 5×10^{-3}]
Synchronous diesel unit	
Wound rotor synchronous generator	
Nominal power, line-to-line voltage, frequency [P_n(VA) V_n(V_{rms}) f_n(Hz)]	[3.125e6 2400 60]
Reactances [X_d $X_d^{'}$ $X_d^{''}$ X_q $X_q^{''}$ X_l] (p.u)	[1.56, 0.296, 0.177, 1.06, 0.177, 0.052]
Time constants [$T_d^{'}$ $T_d^{''}$ $T_{qo}^{''}$] (s)	[3.7, 0.05, 0.05]
Inertia coefficient, friction factor, pole pairs [H(s) F(p.u) p]	[1.07 0 2]
Diesel engine governer	
Regulator gain k	40
Regulator time constants [T_1 T_2 T_3] (s)	[0.01 0.02 0.2]
Actuator time constants [T_4 T_5 T_6] (s)	[0.25 0.009 0.0384]
Torque limits [T_{min} T_{max}] (p.u)	[0 1.1]
Engine time delay T_d (s)	0.024 s

References

1. Díaz-González F et al (2012) A review of energy storage technologies for wind power applications. Renew Sustain Energy Rev 16(4):2154–2171
2. Yingcheng X, Nengling T (2011) Review of contribution to frequency control through variable speed wind turbine. Renew Energy 36(6):1671–1677
3. Bleijs JAM (2007) Wind turbine dynamic response-difference between connection to large utility network and isolated diesel micro-grid. IET Renew Power Gener 1(2):95–106
4. Abdullah MA et al (2012) A review of maximum power point tracking algorithms for wind energy systems. Renew Sustain Energy Rev 16(5):3220–3227
5. Agarwal V et al (2010) A novel scheme for rapid tracking of maximum power point in wind energy generation systems. IEEE Trans Energy Convers 25(1):228–236

6. Tan K, Islam S (2004) Optimum control strategies in energy conversion of PMSG wind turbine system without mechanical sensors. IEEE Trans Energy Convers 19(2):392–399
7. Wu B, Lang Y, Zargari N, Kouro S (2011) Power conversion and control of wind energy systems. Wiley, New York
8. Soetedjo A, Lomi A, Mulayanto WP (2011) Modeling of wind energy system with MPPT control. In: 2011 international conference on electrical engineering and informatics (ICEEI). IEEE, New York
9. Kasem Alaboudy AH, Zeineldin HH, Kirtley J (2013) Simple control strategy for inverter-based distributed generator to enhance microgrid stability in the presence of induction motor loads. Gener Transm Distrib IET 7(10):1155–1162
10. Minxiao H, Xiaoling S, Shaobo L, Zhengkui Z (2013) Transient analysis and control for microgrid stability controller. In: 2013 IEEE Grenoble PowerTech (POWERTECH), pp 1–6. IEEE, New York
11. Jayawarna N, Wu X, Zhang Y, Jenkins N, Barnes M (2006) Stability of a microgrid. In: The 3rd IET international conference on power electronics, machines and drives, 2006. IET, pp 316–320
12. Majumder R (2013) Some aspects of stability in microgrids. IEEE Trans Power Syst 28 (3):3243–3252
13. Conroy JF, Watson R (2007) Low-voltage ride-through of a full converter wind turbine with permanent magnet generator. IET Renew Power Gener 1(3):182–189
14. Tang X, Deng W, Qi Z (2014) Investigation of the dynamic stability of microgrid. IEEE Trans Power Syst 29(2):698–706
15. Kim KH, Van TL, Lee DC, Song SH, Kim EH (2013) Maximum output power tracking control in variable-speed wind turbine systems considering rotor inertial power. IEEE Trans Industr Electron 60(8):3207–3217
16. Zou Y, Elbuluk M, Sozer Y (2011) Stability analysis of maximum power point tracking (MPPT) method in wind power systems. In: Industry applications society annual meeting (IAS), 2011 IEEE. IEEE, New York
17. Kot R, Rolak M, Malinowski M (2013) Comparison of maximum peak power tracking algorithms for a small wind turbine. Math Comput Simul 91:29–40
18. Heydari M, Smedley K, Comparison of maximum power point tracking methods for medium to high power wind energy systems. In: The 20th Iranian electrical power distribution conference (EPDC2015), 28–29 April 2015, Zahedan, Iran
19. Xia Y, Ahmed KH, Williams BW (2013) Wind turbine power coefficient analysis of a new maximum power point tracking technique. IEEE Trans Industr Electron 60(3):1122–1132
20. Nasiri M, Milimonfared J, Fathi SH (2014) Modeling, analysis and comparison of TSR and OTC methods for MPPT and power smoothing in permanent magnet synchronous generator-based wind turbines. Energy Convers Manag 86:892–900

Evaluation of Harmonic Effect on Capacity and Location of Optimal Capacitors in Distribution Network Using HBB-BC Algorithm

Vahid Asgari

Abstract In this paper, the locating and determining the optimal capacity of capacitor banks were studied in the 15-Bus standard distribution network and considering the harmonic effect. The aim was to find the location and capacity of capacitor banks with a loss reduction approach and improve the voltage profile of the network. Constraints such as capacity of capacitor banks, voltage limitation and harmonic distortion limitation were considered in the optimization. Due to the nonlinear nature of the issue of the capacity location in the distribution network, non-linear methods were used. For this purpose, a Hybrid Big Bang-Big Crunch (HBB-BC) algorithm has been used. In simulations, firstly, using the power loss indicator on the slacks, the locations of the banks was determined to estimate the proper capacity of the banks by HBB-BC algorithm. Studies were repeated in two scenarios. In the first scenario, it was assumed that there was no harmonic in the network, and in the second scenario, studies were conducted in the presence of harmonic loads. The results indicated that considering the harmonics in the problem of capacity location will have a great effect on the characteristics of the network and should be considered in studies.

Keywords Harmonic distortion · Harmonic loads · Hybrid Big Bang–Big Crunch (HBB-BC) algorithm

1 Introduction

The reactive power compensation in the design of power systems plays an important role. In this idea (reactive power compensation), the first problem is to determine the optimal capacity of the capacitor, which is optimized based on improving the voltage profile and reducing the losses and improving the power factor in the system and ultimately causing increases the loading limit of lines.

V. Asgari (✉)
Mazandaran Electricity Distribution Company, Sari, Mazandaran, Iran
e-mail: Asgari.vahid@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_45

But the issue that matters is the installation location and capacity of the capacitive banks. In recent years, numerous studies were done to determine the location and optimal size of capacitors to reduce losses. The problem of determining the location and optimal size of the capacitor is a nonlinear and hybrid optimization problem. Therefore, it is necessary to use nonlinear methods or meta-heuristic algorithms to solve the problem to locate and determine the capacity of capacitive banks.

In [1], the optimal locating of fixed capacitors in the main and subset feeders of radial networks, which leads to a reduction of the energy losses, the release of existing feeder capacities and the improvement of the voltage profile were studied. In this reference, an optimal capacity problem was presented in the radial distribution network feeders and an analytical solution for problem solving. The proposed method in this reference was independent from the voltage level of the studied system and its only limitation was the radial power of the studied system. Therefore, this method can be used in low-pressure and medium-pressure networks with radial structure. The objective function in the optimization problem is to reduce energy losses by taking into account the cost of capacitors. Firstly, the cost of providing the reactive power required for network loads was considered. In the proposed method, both parameters of the optimal amount and location of fixed capacitors in feeders with non-uniform load flow, taking into account the average load values in certain time periods and with or without the feeders, were calculated by closed mathematical equations. The presented method in this paper, as a new alternative with greater accuracy and providing a model close to existing conditions and eliminating unrealistic assumptions than other analytical methods, can be used in practice for regional electricity. The results of this method was based on a 15-Bus radial feeder was presented in the end of the paper. The purpose of [2] was to obtain the optimal capacity size and location in the radial distribution network using a genetic algorithm was designed to reduce costs and improve the quality of power. In this reference, the obtained results were better than the results of other methods. In this method of study was considered for various load states such as active and reactive loads, constant impedance, etc. In [3] was studied to find the optimal capacity value using the Harmony Search Algorithm, which aims to minimize the cost function, and maximize the feeder efficiency and improve the voltage profile and the power quality. In [4] was used the differential evolution algorithm to obtain the optimum shunt capacitor size in the distribution network for different loading states (constant loads, variable loads and effective loads), so that the objective function for different load conditions was done by a DE algorithm was performed on the standard network. In [3] two methods were used that include: sensitivity analysis and gravitational search algorithm (GSA). Sensitivity Analysis is a systematic technique used to reduce the search space to find the correct solution for determining the location of capacitors. Capacitor value was determined using the GSA for the relevant locations. In order to validate the proposed method, a number of networks with different states were used in different sizes and complexities, and were also compared with the internal point analytical algorithm and one of the meta-heuristic optimization methods called thermal simulation. The results shown that the proposed method was able to provide optimal responses. In the proposed method [5], was conducted an integrated approach of loss sensitivity coefficient and voltage stability index to determine the optimum connection of capacitor banks.

The Bacterial foraging optimization algorithm was used to determine the optimum size of the capacitors [6]. The simulation results was shown the performance and effectiveness of the proposed method. In [7] Taboo search was used to solve the problem for load growth, feed capacity and voltage constraints were considered to minimize investment costs and system energy losses. Also, a sensitivity analysis method was applied to identify selected locations for capacitors and thereby reduce the search space of the problem. Today, the widespread development of nonlinear loads such as fluorescent lamps, as well as power electronics, generate and introduce considerable amounts of harmonic currents to the power system. These harmonic currents are applied to the equipment due to the impedance of the lines in the form of harmonic voltages. Although capacitors are devices that do not produce harmonics themselves, they have effects on the harmonics produced by nonlinear loads that should be studied. One of these effects is resonance in the network [8]. In this research, determining the optimal capacity of capacitor in the distribution network in the presence of harmonics and a comparison between the optimal capacity of the capacitor in the distribution network in the presence of the harmonic as well as the absence of it was studied.

2 The Studied Network

The studied system was an IEEE 15-Bus standard system, that was shown in Fig. 1, the linear diagram of this 1752 KVA network. The total load of this system was 1752 KVA with power coefficient value of lag 0.7. The system losses was estimated by value of 61.95 KW before the compensation.

The Bus load data and also the lines are given in Tables 1 and 2, respectively.

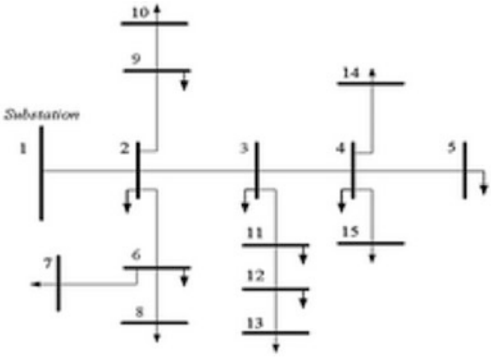


Fig. 1 The single-line diagram of the studied network

Table 1 System loads

Bus number	KVA	Bus number	KVA
1	0	9	100
2	63	10	63
3	100	11	200
4	200	12	100
5	63	13	63
6	200	14	100
7	200	15	200
8	100		

Table 2 Resistance and reactance of lines

Branch number	Sending node	Receiving node	R (Ω)	X (Ω)
1	1	2	1.353	1.323
2	2	3	1.170	1.144
3	3	4	0.841	0.822
4	4	5	1.521	1.027
5	2	9	2.013	1.358
6	9	10	1.686	1.137
7	2	6	2.557	1.725
8	6	7	1.088	0.734
9	6	8	1.252	0.844
10	3	11	1.795	1.211
11	11	12	2.448	1.651
12	12	13	2.013	1.537
13	4	14	2.231	1.504
14	4	15	1.197	0.807

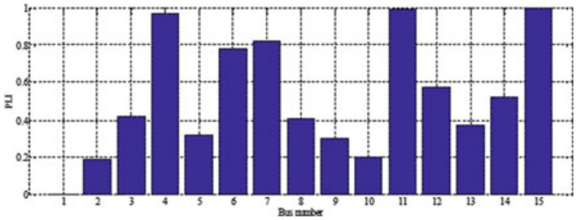
To create harmonic conditions in the network, it was sufficient to use nonlinear load in one or several slacks. The active and reactive power equations were in the form of Eqs. (1) and (2).

$$P(s) = P_0\left(\frac{V}{V_0}\right)^{n_p}\frac{1+T_{p1}s}{1+T_{p2}s} \tag{1}$$

$$Q(s) = Q_0\left(\frac{V}{V_0}\right)^{n_q}\frac{1+T_{q1}s}{1+T_{q2}s} \tag{2}$$

In the above Equations, V is the positive sequence voltage and V_0 is the initial value of positive sequence voltage. Also, the Q_0 and P_0 are values of active and reactive power at the initial point of V_0. n_p and n_q are control parameters of the load, which must be selected between 1 and 3. T_q T_{p1}, T_{p2}, T_{q1} and T_{q2} are control constants of load active and reactive power.

Fig. 2 LPI index



3 Power Loss Index (LPI)

PLI was used to determine the location of capacitor banks in its distribution system. So that any Bus with higher value of LPI, had an increasingly potential for reducing losses in the case of installing reactive power compensators in that Bus. This index was calculated as the Eq. (3).

$$PLI(i) = \frac{I(i) - I_{\min}}{I_{\max} - I_{\min}} \tag{3}$$

In Eq. (3), I(i) is the value of the i-Bus current. I_{max} and I_{min} are the value of the maximum and minimum currents respectively. Value of the PLI for studied network was obtained that shown in Fig. 2 as bar graph.

Based on the obtained results from the simulation, Buses 4, 11, 15 are the most suitable places for the installation of capacitor banks accordance with the loss reduction approach.

4 Constraints and Objective Function

After finding the appropriate location for capacitor banks using LPI, the capacity of these compensators were determined. The problem of determining capacity of capacitor banks in a distribution network is a completely non-linear and high-order problem. As a result, it is beneficial to use meta-heuristic methods to solve this problem.

$$\begin{aligned} &Maximize \\ &f = \max(\Delta P^{R}_{Loss}) \end{aligned} \tag{4}$$

where,

$$\Delta P^{R}_{Loss} = \sum_{k=1}^{n} P_{T,Loss}(k, k+1) - P'_{T,Loss}(k, k+1) \tag{5}$$

In Eq. (5), the loss value of each of the different lines prior to the restructuring was obtained from Eq. (6), the loss value of lines after the reconstruction, which follows from Eq. (7).

$$P_{(k,k+1)} = R_k \frac{(P_k^2 + Q_k^2)}{|V_k|^2} \tag{6}$$

$$P'_{(k,k+1)} = R_k \frac{(P_k'^2 + Q_k'^2)}{|V_k'|^2} \tag{7}$$

In Eqs. (6) and (7), R is impedance of line, P and Q are the active and reactive powers across from line, respectively. The constraints to be considered in determining capacity of capacitor banks are given below.

(A) The voltage value in all Buses must remain constant in the range of 0.95–1.05.

$$0.9 \le V \le 1.05 \tag{8}$$

(B) The amount of reactive power introduced by the capacitor banks of 70% of the reactive power required by the load is lower than the load of the system remains as a lag.

$$\sum_{b=1}^{CB} Q_c(b) \le 0.7 \sum_{q}^{N} Qd(q) \tag{9}$$

(C) Power coefficient for each Bus (PF) must not exceed the considered constraints.

$$PF_{\min} \le PF \le PF_{\max} \tag{10}$$

(D) The total harmonic distortion value for all Buses is less than the maximum value.

$$THD \le THD_{\max} \tag{11}$$

5 Load Distribution of Backward and Forward Sweep Technique

In this paper, a forward–backward load distribution method was used, which is briefly described. The two main steps of this approach are the backward and forward sweep.

Forward sweep: In this step, the whole length of the network from the reference slack, which is the distribution station, is swept to the end of the feeder, and usually one of the network parameters such as the slacks voltages is updated at this step. In other words, at this step, the slacks voltages can be updated with the current of branches from the previous iteration and the use of the radial structure of the network. Generally, in this step, a parameter occurs that is the nature of its changes from the beginning of the feeder to the end of it. The most significant example is the slack voltage.

Backward sweep: At this step from the end of the feeder towards the beginning, the feeder is swept, and usually one of the parameters associated with the forward sweep parameter is updated in this step. Otherwise, if slack voltage is updated in the in the forward sweep state, in this step, due to the obtained slack voltage from the prior sweep, the current of the branches from the end of feeder is updated.

Forward-backward sweep methods due to the high speed and the need for a lower amount of computer memory, as well as their good convergence feature, widely were used in distributed load calculations. Here, the general algorithm consists of two basic steps of the forward and backward sweep that was repeated to be convergence. These methods are divided into three types total current, total power and total impedance methods.

The steps for the forward-backward sweep method can be described as follows for a non-linear network.

1. Determine Slack
2. The size and angle of the voltage for each node assumed as (usually 1pu < 0)
3. The current of a node in k-th iteration is calculated as follows.

$$I_i^{(k)} = \left[\frac{S_i^{sch}}{V_i^{(k-1)}}\right]^* \tag{12}$$

4. From the beginning of the feeder (Slack) towards the end, the voltage in each node in k-th iteration is calculated.

$$V_j^{(k-1)} = V_i^k - Z_{ij} I_{ij}^k \tag{13}$$

5. Power error in each node was calculated if the error value is lower than the allowed error, the process is completed. Otherwise, continues to step 3. The power error value is calculated from the following equation.

$$\Delta S_i^{(k)} = S_i^{sch} - V_i^{(k)} (I_i^{(k)})^* \leq \varepsilon \tag{14}$$

6 HBB-BC Algorithm

HBB-BC algorithm, the combination of the BB-BC algorithm and particle swarm algorithm (PSO) was used to determinate capacity of capacitor banks in distribution system. Each one of BB-BC and PSO algorithms have advantages and disadvantages. But in the hybrid algorithm, disadvantages of the two algorithms were reduced, but the consistency of both algorithms remains suitable.

The particle swarm algorithm was first proposed by Eberhart and Kennedy in 1995. This method has been inspired by the group movement of fishs and the migration of birds. This algorithm is used to explore governed patterns of simultaneous flight of birds and a sudden change in their route. Particle swarm has little quantitative parameters to set, so it's easy to implement. The velocity and position values of each particle are updated according to the mathematical and logical formulas. This algorithm has a special character of memory to store the best place for each particle. In the particle swarm algorithm, since the particle moves through the search for optimal particles in the search space, the convergence rate of this algorithm is greater than other evolutionary algorithms. But the disadvantage of this algorithm is that, in a very sophisticated search environment, particles move at a faster speed to avoid local minima and may lose significant amounts of search space, and also the basic parameters setting such as weight inertia, maximum allowed speed and acceleration coefficients that are capable in finding the optimal response, according to the needs of the problem has little ability. BB-BC algorithm was first introduced by Eksin in 2006. This algorithm is inspired by the phenomenon of the beginning and end of the universe, called the Big Bang in the universe, which is related to the emergence of the universe, and the great contraction or great destruction that relates to the collapse of the universe and the end of its life. It consists of two stages, the BB-BC algorithm from primary population point of view is similar to other evolutionary algorithm [9].

The first-generation population is called the Big Bang, in which the population is randomly distributed over the entire search space. It then turns to the Big Crunch phase, which is actually a convergent operator. This operator with a large number of inputs has only one output, which is called the center of mass, and is calculated using the following equation:

$$X_i^{(k)} = \frac{\sum_{j=1}^{N} \frac{X_i^{(k,j)}}{f_j}}{\sum_{j=1}^{N} \frac{1}{f_j}} \quad \mathrm{i} = 1, 2, 3, \ldots, \mathrm{c} \tag{15}$$

where $X_i^{(k)}$ is i-th component of center of mass in k-th iteration, $X_i^{(k,j)}$ is i-th component from j-th particle generated in k-th repetion. f_j is value of the objective function in j-point, n is number of particles and c is the number of control variables. In the proposed HBB-BC algorithm, using particle swarm algorithm (PSO) capacities, the HBB-BC algorithm searching capabilities improve, and it prevents the BB-BC from falling at the optimal local points. In the HBB-BC and PSO algorithms, to find the center is used from the local optimum points and general optimum points for create new points.

$$X_i^{(k+1,i)} = \alpha_2 X_i^k + (1+\alpha_2)(\alpha_3 X_i^{pbest(k)} + (1-\alpha_3) X_i^{pbest(k,i)}) + (r_i \alpha_1 (X_{i\max} - X_{i\min}))/(K+1) \tag{16}$$

In Eq. (16), $X_i^{pbest(k,i)}$ is the best place of j-th particle until k-th iteration, and $X_i^{gbest(k,i)}$ is the general best place until k-th iteration. α_2 and α_3 are the adjustable parameters that control the effect of the local and general optimal points.

7 Simulation Results

In order to validate the proposed method, locating and capacity capacities of capacitor banks in the studied system were performed. Studies were repeated in two scenarios, and in the first scenario it was assumed that there was no harmony in the network and optimization was done. In the second scenario with adding harmonic loads, the problem of locating and determining the capacity of the capacitor banks was solved by the HBB-BC algorithm. As noted above, locating of the capacitor banks was conducted by HBB-BC algorithm that its parameters are given in Table 3.

As discussed in the previous section, the aim of this research was to determine the optimum capacitor capacity in the loss power reduction network in the distribution system lines. For this purpose, it was assumed that there was only possibility of installing three capacitors in this system. After putting the capacitors in selected buses (4.11 and 15-bus), optimum capacity of each capacitor bank was obtained by the HBB-BC algorithm. The result values from the HBB-bac simulation are given in Table 4.

Table 3 Parameters of the HBB-BC algorithm

Population	Iteration	α_1	α_2	α_3
100	100	0.9	0.6	0.7

Table 4 Optimum capacity of capacitor banks in the first scenario

Slack number	4	11	15
Capacity of capacitor banks	365	350	310

Installing capacitor banks with the defined capacities in the selected slacks by HBB-BC algorithm, total loss of lines were calculated that are shown in Fig. 3 as a bar graphs.

In Fig. 3, the losses of the lines before compensation are shown as red bar and after compensations as blue bars. The highest value of loss was related to Line 1, losses before capacitor locating was 37 KW, while after compensation was reduced to 20.5 KW. Losses values of all lines after the installation were found by 30.3 KW. In Table 5 value of the system losses after the capacitor locating in a 15-bus network is given by several other algorithms.

By examining the obtained results in Table 5, the highest percentage of loss reduction associated with HBB-BC algorithm is 5.50%. However, by genetic algorithm, the number of capacity banks was higher, but due to an incorrect locating, its losses were higher than other methods.

Creating the harmonic conditions in the power system, the issue of locating and determining capacitor banks was solved. For this purpose, the harmonics in the network were created by applying two nonlinear 13 and 15-buses with previous capacities. The specification of these two nonlinear loads are given in Table 6.

By changing the loads of two 13 and 15-buses with non-linear loads, the THD value was obtained for all slacks, that it's values shown as bar graph in Fig. 4. The maximum value of the voltage THD is for the 13 and 15-slacks, which are 4.3 and 3.9%, respectively.

In this scenario, the PLI index was used to determine the location of capacitor banks. Due to no change loads capacity in the second scenario, the proposed location for 15 and 11, 4-buses installation of the capacitor banks was proposed. Using the HBB-BC algorithm, taking into account the harmonic constraint,

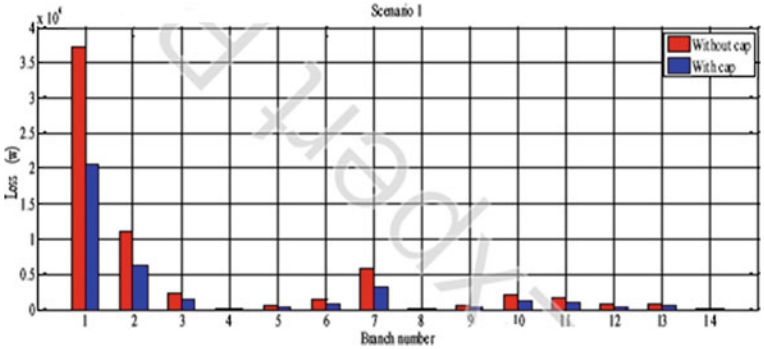


Fig. 3 Losses of electrical lines

Table 5 Performance evaluation of different algorithms

	Without compensation	GA		PSO		DE		HBB-BC	
Loss (KW)	61.95	33.2		32.7		32.3		30.3	
Percentage of loss reduction	–	46.41		47.22		47.86		50.7	
Install location and optimum capacity of capacitor banks		3	150	6	871	3	454	6	370
		4	300	11	321	6	500	11	340
		6	300	–	–	11	178	15	320
		11	150	–	–	–	–	–	–

Table 6 Specifications of nonlinear loads

Slack number	13	15
Capacity (KW)	63	200
Initial value of positive sequence voltage [Mag(pu) Phase (deg.)]	[0.54–11.8]	[0.54–11.8]
n_p	2.3	2.3
n_q	6	6
$T_{p1} = T_{p2} = T_{q1} = T_{q2}$	0	0

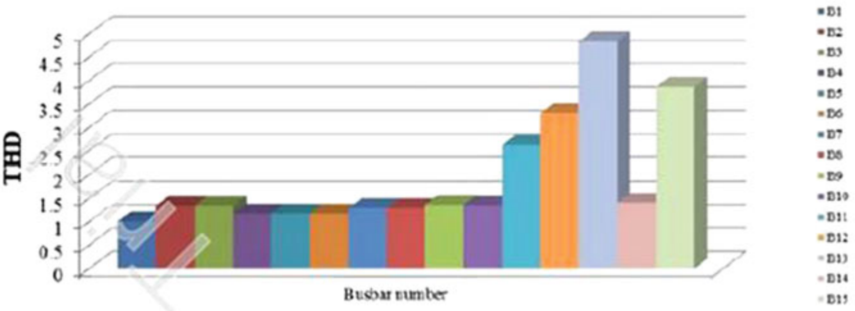


Fig. 4 The THD value was obtained for all slacks

Table 7 Optimum capacity of capacitors in the second scenario

Number of selected bus	4	11	15
Capacity of capacitor (KVA)	360	230	135

the optimal values of capacitor banks were determined, which are given in Table 7. Capacity of capacitor should be set to voltage TDH at all buses remains less than 5%.

In the second scenario, the losses of all lines were calculated. Figure 5 indicates the losses of lines in the new conditions before and after compensation. The amount of losses before compensation was 63.1 KW that is more than the non-harmonic conditions. But after the capacitor locating, total losses of all lines was change to 35.7 KW. As in the first scenario, the highest losses in both cases with a capacitor and no capacitor were calculated more than other lines.

To confirm the performance of the proposed HBB-BC algorithm, two particle swarm and genetic algorithms were used to determine capacity of capacitor banks as shown in Table 8.

The results obtained in Table 8 shown that the proposed method is more suitable than the other two algorithms. The HBB-BC algorithm reduced the amount of losses by 35%, while using of the genetic algorithm, the amount of losses reduction was 36.8%. In the presence of harmonic loads due to THD constraint, possibility of installing the more capacitors was impossible and capacitor banks could not delivery more reactive power on the network. As a consequence, due to the pass of the reactive power from the lines, losses increased. In the following, the voltage THD values of all slacks were recalculated. These values are shown in Fig. 6 in the form of bar graphs.

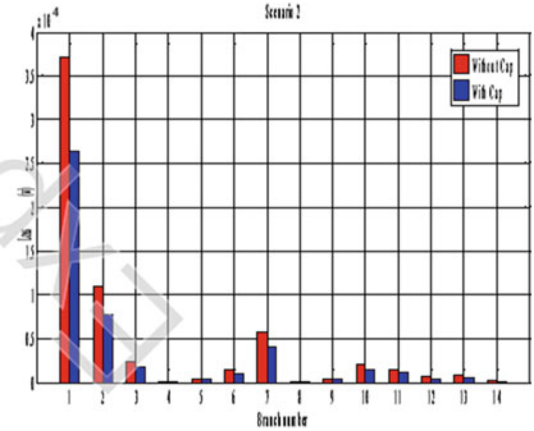


Fig. 5 Lines losses in the second scenario

Table 8 Evaluation of the performance of various algorithms

	Without compensation	GA		PSO			HBB-BC	
Loss (KW)	63.1	40.1		38.2			35.7	
Percentage of loss reduction	–	36.8		39.6			43.4	
Install location and optimum capacity of capacitor banks		6	360	6	420		370	6
		11		300	11	310	340	11
		15		340	15	280	320	15

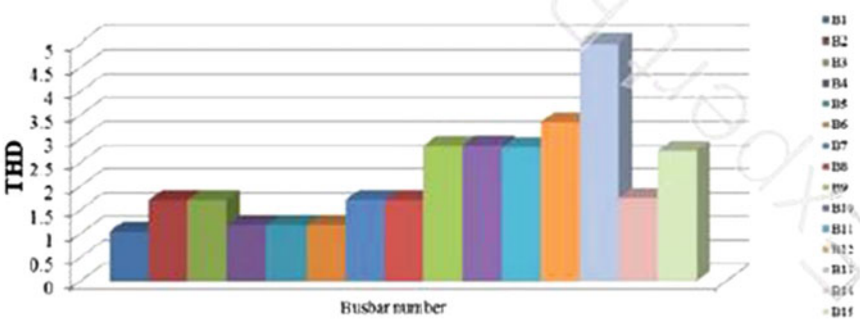


Fig. 6 The value of THD voltage of the slacks after the compensation

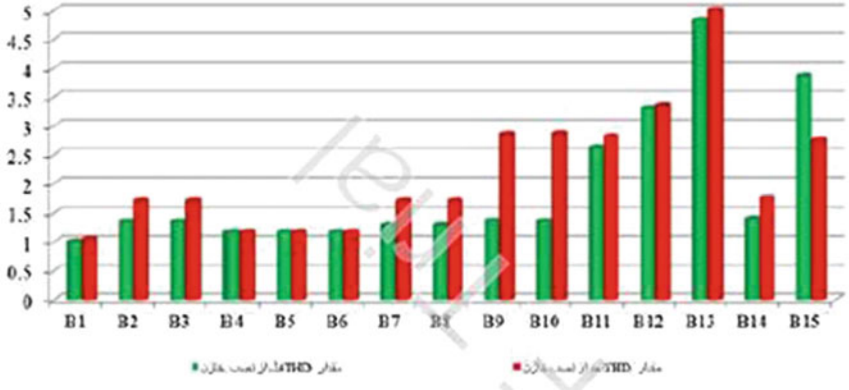


Fig. 7 The TDH voltage of all slacks except the 16-slack installation of capacitor banks

According to the results, it can be deduced that the TDH value in the 13-bus reached by 5%, and was limitated determining the capacity of the capacitor banks, but the TDH voltage of all slacks except the 16-slack increased if installation of capacitor banks. In THD values were increased for other buses Fig. 7.

8 Conclusion

In this paper, the problem of locating and determining capacity of capacitive banks with the aim of reducing the losses of lines in a sample distribution system was studied. The problem was solved by PLI index and was performed determining the optimum location and capacity of capacitor banks by a powerful HBB-BC algorithm. Studied were repeated in two scenarios, with or without of harmonic loads.

In the first scenario, it was assumed that there was no harmonic source in the network. After the determination of optimal location of installation of capacitor banks by PLI index, the capacity of each bank was determined by the HBB-BC algorithm. The results indicated that after installing capacitor banks, the losses of system reduced by value of 50.7%. To confirm the results of the simulation, a comparison was made between the obtained results from the suggested DE, PSO, GA algorithms. The results shown that the HBB-BC algorithm is more suitable than other methods.

But in the second part of the study, with two nonlinear loads in the system, studies were repeated and in this scenario, the THD voltage with the maximum total distortion of all buses less than 5% added to the problem constraint. This limitation has caused the capacitor banks in the new conditions to be lower than the values for the obtained THD in the first scenario. The TDH constraint for a slack with non-linear load, but without the installation of capacitor bank, limited the capacity of capacitor bank. The losses reduced by value of 43% in this scenario. In this section, results were compared with two genetic algorithm and particle swarm algorithm.

The simulation results were defined in two scenarios: in the presence of harmonic loads in the distribution system, by installing a capacitor, increased the voltage harmonics of all slacks, except for the slack with nonlinear capacitor and load. With regard to the results, it could be argued that it was better to install capacitor banks on slacks where there were nonlinear loads.

References

1. Nojavan S, Jalali M, Zare K (2014) Optimal allocation of capacitors in radial/mesh distribution systems using mixed integer nonlinear programming approach. Int J Electric Power Syst Res 107:119–124
2. Sultana S, Roy PK (2014) Optimal capacitor placement in radial distribution systems using teaching learning based optimization. Int J Electr Power Energy Syst 54:387–398
3. El-Fergany AA, Abdelaziz AY (2013) Cuckoo search-based algorithm for optimal shunt capacitors allocations in distribution networks. Electric Power Components Syst 41 (16):1567–1581
4. Das P, Banerjee S (2013) Placement of capacitor in a radial distribution system using loss sensitivity factor and cuckoo search algorithm. Int J Sci Res Manag 2(4):751–757
5. Fard AK, Samet H (2013) Multi-objective performance management of the capacitor allocation problem in distributed system based on adaptive modified honey bee mating optimization evolutionary algorithm. Electric Power Components Syst 41(13):1223–1247
6. Legha MM, Tavakoli M, Ostovar F, Hashemabadi MA (2013) Capacitor placement in radial distribution system for improve network efficiency using artificial bee colony. Int J Eng Res Appl 3(6):228–233
7. El-Fergany AA, Abdelaziz AY (2014) Capacitor allocations in radial distribution networks using cuckoo search algorithm. IET Generation Transm Distrib 8(2):223–232
8. El-Fergany AA, Abdelaziz AY (2014) Artificial bee colony algorithm to allocate fixed and switched static shunt capacitors in radial distribution networks. Electric Power Components Syst 42(5):427–438

9. Das P, Banerjee S (2014) Optimal sizing and placement of capacitor in a radial distribution system using loss sensitivity factor and firefly algorithm. Int J Eng Comput Sci 3(4):5346–5352
10. Hamouda A, Sayah S (2013) Optimal capacitors sizing in distribution feeders using heuristic search based node stability indices. Int J Electr Power Energy Syst 46:56–64

Performance Evaluation of Indicators Effective in Improving Air Cooler Output by Linear Programming

Amir Khayeri Dastgerdi

Abstract The air conditioning unit requires water and electricity. It has household appliances and facilities, and has many manufacturing benefits, including the simplicity of technology and low cost of production. This unit in addition to cooling the environment, it also provides moisture. The purpose of this paper is to examine different solutions for reengineering the production process in order to identify effective indicators for optimizing production costs, outlet air temperature and water and power consumption efficiency of this device.

Keywords Production cost · Water consumption · Power consumption Outlet air temperature · Methodology

1 Introduction

The water cooler is a cooler that cools the air with evaporation of water. Water coolers work with evaporative cooling. Evaporation cooling is a process in which evaporation is used as a natural heat absorber. In this process, the tangible heat of the air is absorbed and used as the necessary heat for evaporation of water. The amount of heat absorbed depends on the amount of water that evaporates [1]. Evaporative cooling is a very old process, which originated thousands of years ago in the ancient civilizations of Iran and Egypt. New evaporative coolants are produced based on original samples made in the 1900s by the United States. Evaporative cooling can be done directly or indirectly, spontaneously or in combination. In direct evaporative cooling, the amount of water in the cooled air increases. In indirect evaporation, evaporation occurs within a heat exchanger, and the amount of moisture in the cooled

The type of research in this paper is applicable.

A. K. Dastgerdi (✉)
Islamic Azad University, Najafabad Branch (IAUN), Najafabad, Iran
e-mail: Amirdastgerdi1975@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_46

air does not change. Since high evaporation rates increase relative humidity and may cause uncomfortable environment, direct evaporative cooling should be performed in areas where relative humidity is low. Evaporation is performed spontaneously when the evaporation process is carried out naturally. If it is possible to cool a room with its own evaporation in which there is a barrier of stagnant water or in a stream like a pond or water fountain. In cases where evaporation is carried out by mechanical devices, the evaporation is mixed. Clearly, this type of evaporation is consumed, but the amount of energy consumed is much lower compared to air conditioning. The basis of the evaporative cooling method is the thermodynamic evaporation of water, or in other words, the change of water from liquid to steam. Due to the use of this kind of cooler in Iran more than anywhere else in the world, sometimes these coolers are known as Iranian coolers [2].

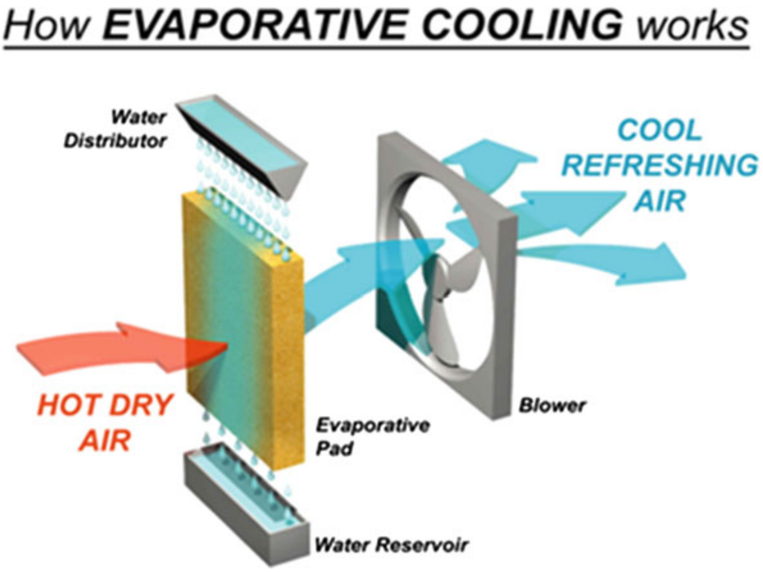
2 Research Method

This article is a descriptive-computational approach in terms of target dimension. Considering that in this analysis of the technical and economical effects of the cooling industry, this question is answered, what is the quality of the products of the cooling industry in the economy of the country? In other words, in this analysis, the flow of cooling-related costs is analyzed to identify changes in sales, production, imports, employment, and some other effects, such as quality and cost-effectiveness. In this study, the prioritization of indicators is used through the scientific method of linear programming. My goals are to change the old index that was evaluated in the production of a cooler. Reengineering the design Through linear planning, all of the old and new indicators, including body material, number of filters, intake air path, production cost, air temperature, are evaluated. To weigh up any of the indices, we must prioritize the indicators of the efficiency of the cooler so that we can optimize the production. We first weigh the linearization method to the indices. If the index is positive, each number will be divided by the largest number of its column, and if the index is negative, the smallest number will be divided by each column number. Then, using the chanol entropy method, we obtain the uncertainty of the index as follows.

3 Findings

Research has shown that most cooler manufacturers have taken into account repetitive indicators, including the amount of vapor in the cooler, and considered the cooling condition of the system only through the creation of moisture and the use of it in the cooling system. In a cooling system, evaporative water coolers, which are formed by splitting the water molecules and increasing its volume, in addition to absorbing the heat of the intake air and cooling it, require the vapors to return to the water's primary state, so that a reversible evaporation cycle can be performed. But in the current coolers, this cycle is cut off after evaporation of the water and getting the heat of the intake air, and that warmed moisture is wasted by leaving the pores of the body and causing excessive consumption of water [3].

In my design, as in the philosophy of refrigerators using frigo gas, I believe that the heated frost gas in the radiator inside the house is re-directed to cool off the radiator outside the room. Therefore, in this scheme, after removing the heat of the intake air by the water vaporized inside the water cooler, it is necessary to prevent the vent from exiting the cooler body and return to the cooling process again. The design of the body has become more important as it has not been addressed so far. The scientific look at the Hollywood production technology actually creates the realization of the knowledge-based economy, which also drives producers' approach to research and development in production. Optimum production has a meaning beyond production, because in the global market, the product will be exploited at a minimum cost. This product will have a finished product, and this will not be possible with a scientific look at the production and implementation of the related scientific process.



4 Statement of the Problem

Our goal is to reduce the temperature of the outlet air, reduce water and electricity consumption, reduce sales prices on the market, and implement a competition strategy. One of the drawbacks to the water cooler system is the high water consumption, which averages 250 L per day. Cooling at 22 °C has also led consumers to buy expensive gas coolers. One of the advantages of producing a cooler is the low cost compared to other coolers in the world. Due to the very low energy consumption of the water pump, it does not appear to have an effect on the energy consumption of the pump, but it should be noted that moisture is the main requirement of the system's cooling system, and the pump water pump exits a lot of air outflow, and the cooler needs a moisture content of 60–80% That the amount or amount of this amount in the entire cooling process affects the performance of the cooling process [4].

It also turns to other remarks about the use of pottery instead of scratching or separating the chrome body or using a battery and a direct current to generate direct electricity and switching the electromotor to DC or using a compressor cooling unit to help cool it. That a plan would be productive and justifiable if the cost of the project, if not less than the previous model, was not technologically, at least not the previous technology, since in this case it would be possible to produce a cooler at a price of 10 million USD for the room temperature Under 30 °C. But should not the consumer be charged? Or compete with other manufacturers as well as guiding them to collaborate in production? There are also statements about designs that are laboratory-related, and the ability to use it is not in practice consistent with existing standards and standards. For example, the use of high-tech acoustic power does not even have the ability to aerate up to 3500 sfm, so only a coil fan model has the ability to produce air with the desired power to carry out the operation of the aerosol and throw. The fan is designed for open spaces such as halls and towers, which is freely available behind the fan, but is closed in the cooler body. Also, changing the impeller blade angle increases the power consumption due to the pressure on the electromotor, which is not justified by the current power consumption. In my invention, I have tried to reduce only water consumption by 140–210 L per day, which is worth the first step in saving the country by one billion and seven hundred million liters per day, so I believe in it. The big stone is showing no signs, and the invention should be taken step by step so, with respect to and respect for previous patents, I should state that if the designs were before it could be produced, the manufacturing companies still did not produce the same model thirty years ago. Based on this, I am basically guided by the fact that reducing the air output of about 4–5 °C and reducing power consumption as well as reducing the cost of making your cooler is a big step, due to the following reasons:

1- The invention should not be in theory and in the production of the laboratory where the investor has the incentive to produce it.
2- The invention should be parallel to the former technology, which the investor does not need to change the whole of his previous production line, since it is costly and less investment is willing to undertake such a new production risk.
3- The invention should not change all the previous technology. It is only necessary to correct the previous faults. In the water cooler, the purpose is to reduce the consumption of water and electricity, as well as to reduce the cost of production. Therefore, in my invention, it is prescribed to use even the wild worm from the same shell Because the use of pottery, a metal washable filter or fiber, the fact that this cooler is likely to be used in the distant parts of the country in terms of after-sales service is very wrong, because in the absence of tufts, sometimes it is used as a desert shrub. The production of clay pile with its high price and inappropriate services of manufacturers have caused consumption. The refrigerator models are not intended to be purchased, so the use of this filter model does not have any effect on efficiency. Therefore, in the new design of the water cooler, we changed the location of the filter to be placed in succession so that the intake air passes through three filters simultaneously and three times Cooling is done, and this design reduces the speed of the airflow by about 1 cm, which is not noticeable.

4- In my invention, I create a cooler body in a mold and polycarbonate material, which eliminates all the losses of water wasting out of the body, vapor and direct throwing of sunlight and warming the body because the polycarbonate body Thermal insulation and anti-corrosion and mass. Despite the thickness and heaviness of the body, remarks have been made of polymer material that is admirable, but it can be used with washable metal chips that are expensive and do not remove the dust from the inlet and it is warming itself, also according to the map. The air does not exactly enter the cooler from one side, but it has a 180° inclination of Hawara, which sometimes does not "pass through all the air entering the entire surface of its filters, and it enters the air side of the filter, which actually has an effect on cooling I do not have to turn to the other inventions. According to my opinion, there is no need to build a two- Or the creation of a canopy for the cooler roof as well as water purification because of the mass of the cooler pan. The cooler process uses water vapor in the ovens called humidity and blowing air between it and only the ability to cool the air at 5°–6° In the invention of the air, there is a process of cooling the air in a closed environment without contact with the outside, due to the fact that it is a piece of the body that reduces the consumption of water, so the moisture in the closed space cools down. And eventually "passes through a nanoscale filter where the dust is taken, and the rest is inward Channel and, finally, the room. The use of air inside the room to move it to the air intake required by the cooler increases the humidity of the pathway of the colmermay process and causes the air to warm up because the cooler is usable in arid areas that can be created by creating Hawara's humidity is cool so reuse of the humidity created in the room space and directing it to the cooler input is completely wrong. The cooler needs 20–30% humidity, which makes it 70 to 80% cooler.

5- In order to reduce the power consumption of my invention, an electromotor without strap and a direct power transmission system using the gearbox has been used that reduces the power consumption of the 7000 air conditioners, which is 4 amps, to 3.2 amps and also reduces the vibration of the body. Is.

6- The current production line for the production of coolers requires 50 workers per production line, but my invention, due to the production of a single piece by molding and injection of polycarbonate materials and the lack of use of straps, requires the worker in each production line to There are 6 people. Because the cost of production is reduced, the investor's desire to use this invention and its production is high.

With regard to all the indicators involved in the construction of a water cooler, the goals to be produced from the investor's point of view, reduce costs and, in the eyes of the consumer, reduce the consumption of water and electricity in this cooler. In the near future, cheap electricity will be achieved with solar technology, but water is not readily obtained, so the crisis is a water crisis. Production must be understandable and logical for the investor. Many laboratory designs can be made, but is the investor entitled to produce it? Is it necessary for a poorly-formed society to build a six-wheel-drive and bulletproof body that can fly? Is it justified? So the plan should include all the needs of the social strata, including the amount of consumer revenue to buy and the investor's desire to produce it, as well as the priority of reducing current crisis resources, such as water or electricity.

5 Results

The goal is to change the old index that was evaluated in the production of a cooler. Reengineering the design through linear planning All the old and new indicators, including body material, number of filters, intake air path, production cost, air temperature, are measured by expert and expert opinion and then weighed To each of the indicators, the effective indicators in the efficiency of the cooler are prioritized. In this case, we can optimize the production. We first weigh the linearization method to the indices. If the index is positive, the column numbers are divided by the largest number of their columns, and if the index is negative, the smallest number will be divided by each column number. Then, using the channel entropy method, we obtain the uncertainty of the index as follows.

$$\mathrm{Ej} = -1/\mathrm{Ln.m}\left[\sum(\mathrm{nij} * \mathrm{Ln(nij)})\right]$$

M	The number of mounts
ij	Also a normalized number of columns and rows
n	Number of indices
$dj = 1 - Ej$	Obtaining the confidence level
$Wj = dj/dj$	Obtain the index weight j

In this regard, we compare three cooler models in terms of the following indicators.

X1 Cost of Cooler Purchase
X2 Cooler Power Consumption
X3 Cooler Water Consumption
X4 Quality Filter
X5 price filter
X6 Output Temperature
X7 Heat Exchanger Body Exhaust
A1: water cooler with metal body and cellulose filter
A2: water cooler with Polycarbonate body and sequential filter (new design)
A3: Freon gas cooler

	X1	X2	X3	X4	X5	X6	X7
Cooler with metal body and cellulose filter—A1	260	5	9	7	5	7	7
New design cooler with polycarbonate body and Opaque filter—A2	315	3	5	3	1	3	7
Cooler Gas Frouw—A3	1315	9	1	9	7	1	5

Cooler with metal body and cellulose filter—A1	260	3	1	7	1
	260	5	9	9	5

260/260 = 1 3/5 = 0.6 1/9 = 0.1111 7/9 = 0.7778 1/5 = 0.2 1/7 = 0.1429 7/7 = 1

New design cooler with polycarbonate body and Opaque filter—A2	260	3	1	3	1	1	7
	315	3	5	9	1	3	7

260/315 = 0.8254 3/3 = 1 1/5 = 0.2 3/9 = 0.3333 1/1 = 1 1/3 = 0.3333 7/7 = 1

Cooler Gas Frouw—A3	260	3	1	9	1	1	5
	1315	9	1	9	7	1	7

260/1315 = 0.1977 3/9 = 0.3333 1/1 = 1 9/9 = 1 1/7 = 0.1429 1/1 = 1 5/7 = 0.7143

	X1	X2	X3	X4	X5	X6	X7
Cooler with metal body and cellulose filter—A1	1	0.6	0.1111	0.7778	0.2	0.1429	1
New design cooler with polycarbonate body and Opaque filter—A2	0.8254	1	0.2	0.3333	1	0.3333	1
Cooler Gas Frouw—A3	0.1977	0.3333	1	1	0.1429	1	0.7143

$$\mathrm{E1} = -1/\mathrm{LN3}[1 * \mathrm{LN1} + 0.8254 * \mathrm{LN.08254} + 0.1977 * \mathrm{LN0.1977}] = 0.0919$$

$$\mathrm{D1} = 1 - \mathrm{E1} = 0.9081 \quad \mathrm{W1} = \mathrm{D1}/6.1554 = 0.1475$$

$$\mathrm{E2} = 1/\mathrm{LN3}[0.6 * \mathrm{LN0.6} + 1 * \mathrm{LN1} + 0.3333 * \mathrm{LN0.3333}] = 0.1291$$

$$\mathrm{D2} = 1 - \mathrm{E2} = 0.9081 \quad \mathrm{W1} = \mathrm{D2}/6.1554 = 0.1415$$

$$E3 = -1LN3[0.1111 * LN0.1111 + 0.2 * LN0.2 + 1 * LN1] = 0.1067$$

$$D3 = 1 - E3 = 0.8933 \quad W3 = D3/6.15544 = 0.1451$$

$$E4 = -1/LN3[0.7778 * LN0.7778 + 0.3333 * LN0.3333 + 1 * LN1] = 0.1078$$

$$D4 = 1 - E4 = 0.8922 \quad W4 = D4/6.1554 = 0.1449$$

$$E5 = -1/LN3[0.2 * LN0.2 + 1 * LN1 + 0.1429 * LN0.1429] = 0.1151$$

$$D5 = 1 - E5 = 0.8849 \quad W5 = D5/6.1554 = 0.1438$$

$$E6 = -1/LN3[0.1429 * LN.01429 + 0.3333 * LN0.3333 + 1 * LN1] = 0.1236$$

$$D6 = 1 - E6 = 0.8764 \quad W6 = D6/6.1554 = 0.1424$$

$$E7 = -1/LN3[1 * LN1 + 1 * LN1 + 0.7143 * LN0.7143] = 0.1704$$

$$D7 = 1 - E7 = 0.8296 \quad W7 = D7/6.1554 = 0.1348$$

$$D1 + D2 + D3 + D4 + D5 + D6 + D7 = 6.1554$$

$$\begin{aligned} WA1 =& 1 * 0.1457 + 0.6 * 0.1415 + 0.1111 * 0.1451 + 0.7778 * 0.1449 + 0.2 * 0.1438 \\ &+ 0.1429 * 0.1424 + 1 * 0.1348 = 0.5451 \end{aligned}$$

$$\begin{aligned} WA2 =& 0.8254 * 0.1475 + 1 * 0.1415 + 0.7778 * 0.1449 + 0.2 * 0.1438 \\ &+ 0.1429 * 0.1424 + 1 * 0.1438 = 0.6666 \end{aligned}$$

$$\begin{aligned} WA3 =& 0.1977 * 0.1475 + 0.3333 * 0.1415 + 1 * 0.1451 + 1 * 0.1449 \\ &+ 0.1429 * 0.1438 + 1 * 0.1424 + 0.7143 * 0.1448 = 0.6256 \end{aligned}$$

$$A2 \ggg A3 \ggg A1$$

6 Conclusion

Nowadays, it is time to assemble and copy the technology of other countries without research and development and localization. Therefore, in a scientific and productive approach to production, in addition to this plan, other designs in the country's products that are being developed based on the technology of previous years can be optimized and localized according to the needs of the country. With the implementation of this plan, about 1,700,000,000 L will be saved in the country's water consumption, which will also lead to the growth of the agricultural economy, which is the most important strategy of the project, which will create employment in agriculture and reduce unemployment and poverty in society. Gets The scientific look into the production of cooler technology in the country actually creates the foundation of the knowledge economy, which drives other manufacturers to research and expand the approach to production. Today, optimal production has a meaning beyond production, because in the world of global production it will be a product that has the lowest cost of goods, and this will not be possible with a scientific look at the production and implementation of the scientific process associated with it. The specific design of the body, considering the indicators such as the chemical behavior of water cells in the path change, which reduces the temperature of water and decreases its molecular mass, and consequently the "fall of the volume of evaporation into the cooler pan and prevent the loss of water vapor. The air inside the system will increase the speed of temperature reduction, reducing the consumption of 140 L of water per 12 h of work compared to the current coolers, which also accounts for about 12 million water coolers in Iran, amounting to one billion and 700 million Liters per day. Creating a standard for all manufacturers of cold storage facilities that include attention to indicators Foots such as cellular behavior of water in rotational motion, the time of evaporation of water and its reversal through condensation, the effective and optimal use of polycarbonate materials in the body, yields productivity, which can capture all the optimal characteristics of this industry. Body quality statement, optimum water consumption and effective air outlet can lead all industry manufacturers to comply with the principles of efficiency so that no investment can enter this area. With this method of operation, you can use the lowest input Gaining the highest output and leading a successful production of low-energy resources.

References

1. Heydari M, Arabipourian F, Gheitrani F, Birjandi AAM (2006) Technology and home appliance repair workshop. Printing and Publishing Company of Iran Textbooks. ISBN 964-05-1265-6
2. Gallipin L, Baktai B (2008) Maintenance of a water cooler, Baktai B ed (trans. Salmani KA). Industrial and Industrial Co., Tehran
3. Tabrizzahi M, Majidi S (2011) Surveying the performance of domestic water coolers in Iran and presenting a new design to improve cooling efficiency. In Third International Conference on Heating, Cooling, and Air Conditioning, Tehran
4. Omid Kashani B (1395) The necessity of the existence of the label of economic efficiency of water consumed in household appliances, including water coolers and practical measures to increase this efficiency. In Iran Water and Wastewater Science and Engineering Congress, Tehran, Iran Water and Wastewater Association Iranian Scientific Association)—Tehran University—Water and Sewage Engineering Company of Iran

Determining the Parameters of Insulation Model by Using Dielectric Response Function

Seyed Amidedin Mousavi and Arsalan Hekmati

Abstract The identification of electrical insulation of power grid's equipment is so important due to increasing the reliability of these networks. To do this, internal insulation of equipment should be accessible. Procedures for conducting such experiments are often costly, time-consuming and impossible. Today using modern methods such as polarization and depolarization current technique (PDC), insulation model can be obtained with an easier way. Insulation's dissipation factor could be calculated using model parameters and plotted at different frequencies. Insulation condition could be observed partially by comparing the dissipation factor curve with its initial curves. Notably, the interpolation of these curves needs enough sciences and skills. In this paper using the curve fitting, Genetic and PSO algorithm the insulation models parameters have been determined by PDC technique. Depolarization current have been calculated by these models parameters and compared with experimental data to modeling validation. Then insulation dissipation factor have been plotted using different models data.

Keywords Insulation model · GA algorithm · PSO algorithm · Dissipation factor

1 Introduction

Transformers are the most important components of the power grids. Their insulations take different stresses during their lifetime and must always have suitable condition during this time. The important factors that could cause serious damage are aging and moisture. For this reason condition of insulation should be monitored continuously. There are several ways to do this. Among these methods polarization and depolarization current technique (PDC) is partly newer. This method has no requirement to information about geometric structure and insulation materials configuration. This is one of the advantages of this approach. In this paper, the

S. A. Mousavi (✉) · A. Hekmati
Department of Electrical Engineerng, Shahid Beheshti University, Tehran, Iran
e-mail: amid_moosavi@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_47

insulation has been modeled as a simple RC model by using of depolarization current. Indeed, the black box model has been used for insulation modeling. There are many techniques for determining the model parameters. In this paper, exponential curve fitting, PSO and genetic algorithm have been used to evaluating them. The power transformer of NEKA power plant has been candidate as a test object.

2 Theory

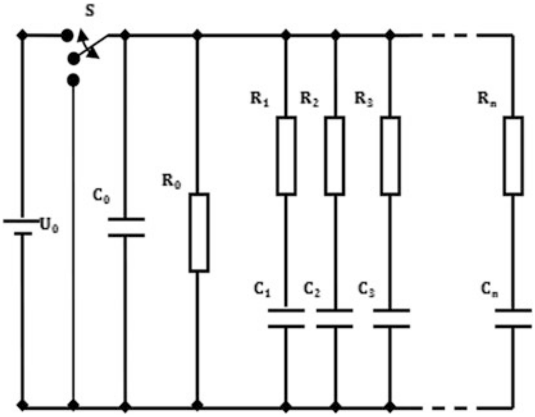
When an electrical field is applied to both ends of a dielectric the polarization current will flow through it. This current flows due to the tendency of dipoles to be in the same direction with electrical field. When the field is removed from the dielectric, after sufficient time, dipoles relax and tend to return to its original state and cause depolarization current that flows in the dielectric. The depolarization current has various relaxation mechanisms that related to different part of insulation. Each part of the insulation has a unique mechanism because of the difference in moisture content and aging of each section [1, 2]. This process can be modeled as a simple parallel RC model cases which is shown in Fig. 1.

The number of branches of model can be varied from five to ten [1]. In this article, the numbers of branches have been considered six. It should be mentioned, the time constant of each branch is larger enough from the previous branch. When a branch is charged, the pervious branch has reached to its steady state. As reviewed above depolarization current can be expressed as [3–5]:

$$\mathrm{i_{depol}} = \sum_{i=1}^{n} \mathrm{A_i} \exp\left[-\frac{t}{\tau_i}\right] \tag{1}$$

$$\mathrm{A_i} = \frac{\mathrm{U_0}\left[1 - \exp\left[-\frac{t_p}{\tau_i}\right]\right]}{\mathrm{R_i}} \tag{2}$$

Fig. 1 Insulation RC model



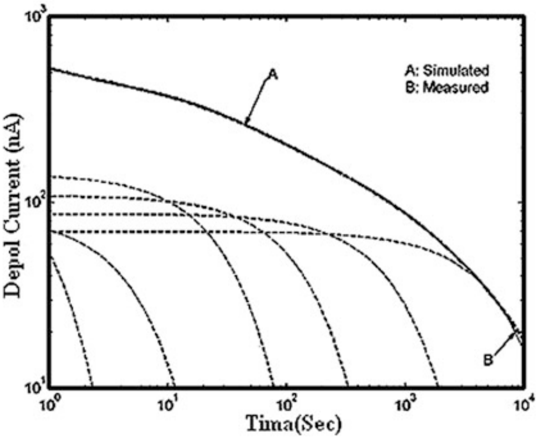


Fig. 2 Depolarization current and its components

which t_p is charging time when DC voltage or DC electrical field is applied on dielectric and τ_i is time constant of each branches [6]. Ri and Ci are series elements of each branches of parallel RC model. Notably, C_0 with using some experimental test in power grid frequency cab be obtained and R_0 is accessible through the difference between depolarization and polarization currents in maximum available time in plots [7–9]. In transformers insulation model, the branch that has smaller time constant models oil performances and the greater ones models paper operations. Figure 2 shows the depolarization current and its components that relative to relaxation mechanisms of different branches.

3 Test Object

Power transformer of NEKA power plant in north of Iran is selected as a test object. A step voltage U_0 = 2 kV according Fig. 3 is applied between high voltage and low voltage winding. Measuring time length has been 50,000 [s]. During the test oil temperature of transformer has been kept constant around 20–25 °C then Dc voltage source has removed from both ends of transformer and these point has been short circuit. The measured depolarization current has been measured and shown in Fig. 4.

4 Parameters Identification Methods

In this paper using exponential curve fitting, genetic algorithm and PSO algorithm insulations model parameters of insulation have been determined and their results have been compared with each other's.

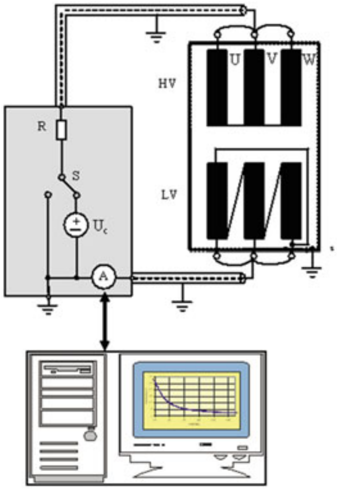


Fig. 3 Test method

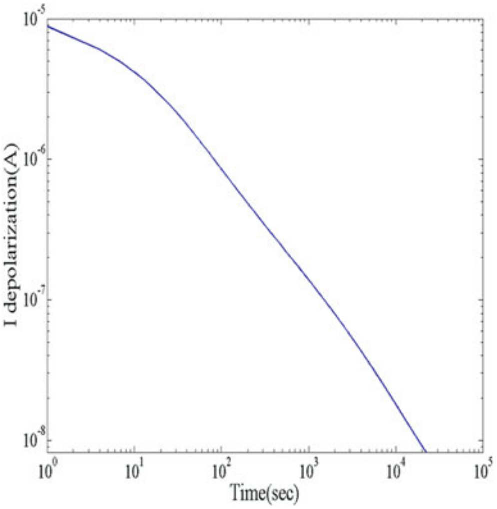


Fig. 4 Measured depolarization current

4.1 Exponential Curve Fitting Method

In this method, the branch which has maximum time constant is selected for parameters estimation. To do this, the end part of depolarization current curve is selected to determining the R_i and C_i values. Then by using these parameters the A_i $\exp[t/\tau_i]$ expression, related to this branch, are calculated. Then this expression is subtracted from total depolarization current. This operation has been continued to obtain all parameters of series RC branches (Table 1). By using this method all of the insulation parameters are determined and illustrated in Table 1.

Table 1 Estimated parameters of model using curve fitting technique

Number of branch	A_i(A)	τ_i (s)	R_i (GΩ)	C_i (nF)
1	2.384e−8	20,842	55	378
2	9.527e−8	2891.8	20	138
3	3.346e−7	569.15	5.9	95
4	1.214e−6	91.5781	1.46	55.5
5	3.475e−6	22.9936	0.57	39.5
6	4.808e−6	3.9936	41.6	959

4.2 Genetic Algorithm and PSO Algorithm

In this section, using genetic algorithm and PSO algorithm [10–12] insulation parameters model has been determined. The chromosomes have been selected as following:

τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	R_1	R_2	R_3	R_4	R_5	R_6

Notably, the fitness function has been considered as below:

$$\text{Fitness} = \frac{1}{\sum \left(i_{\text{depl}-\text{actual}} - i_{\text{depol}-\text{simulated}}\right)^2} \tag{3}$$

A_i's and C_i'c can be obtained easily from relation (4) and $\tau_i = R_iC_i$ respectively. The results have been shown in Tables 2 and 3 and compared in Fig. 5.

Least squared errors (MSE) method is used to demonstrate the error of estimation parameters. The comparison of these results has been shown in Table 4.

Table 2 Estimated parameters of model using genetic algorithm

	A_i(A)	τ_i (s)	R_i (GΩ)	C_i (nF)
1	5.0642e−8	10,161	34.96	290.65
2	1.4811e−7	1408.4	13.5	104.3
3	4.7025e−7	303.5520	4.25	71.373
4	2.0102e−6	54.2026	0.9949	54.480
5	3.6120e−6	13.5507	0.5537	24.473
6	4.1703e−6	2.8214	0.4795	5.883

Table 3 Estimated parameters of model using PSO algorithm

	A_i(A)	τ_i (s)	R_i (GΩ)	C_i (nF)
1	5.0003e−8	10,806	35.02	291.24
2	1.2451e−7	2352	13.46	102.9
3	3.9470e−7	321.55	4.65	73.555
4	1.9329e−6	55.211	1.0112	25.071
5	3.4786e−6	14.001	0.5515	25.009
6	3.9534e−6	2.8341	0.4309	5.395

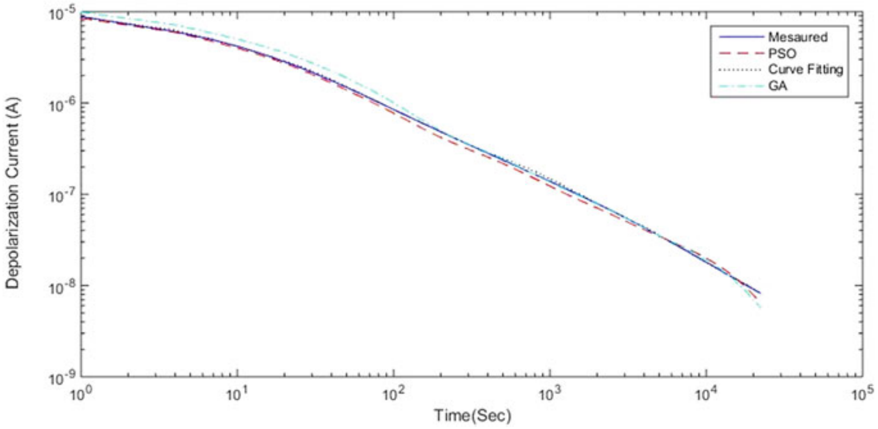


Fig. 5 Depolarization current

Table 4 Comparison of MSE of the three methods

	Curve fitting	GA	PSO
MSE	1.13e−15	6.024e−17	3.481e−17

As seen in Table 4 in PSO method error rate is lower than the others.

5 Using the Model Parameters to Calculation the Dissipation Factor of Insulation

With the insulation model parameters it is possible to calculate the dissipation factor as bellow:

$$\tan\delta(\omega) = \frac{\frac{1}{R_0} + \sum_{i=1}^{n}\left[\frac{R_i(\omega C_i)^2}{1+(R_i\omega C_i)^2}\right]}{\omega C_0 + \sum_{i=1}^{n}\left[\frac{\omega C_i}{1+(R_i\omega C_i)^2}\right]} \tag{4}$$

The values of C_0 and R_0 in this paper have been obtained 10.2 nF and 2.5 GΩ respectively. In Fig. 6 the dissipation factor versus frequency has been demonstrated using several methods.

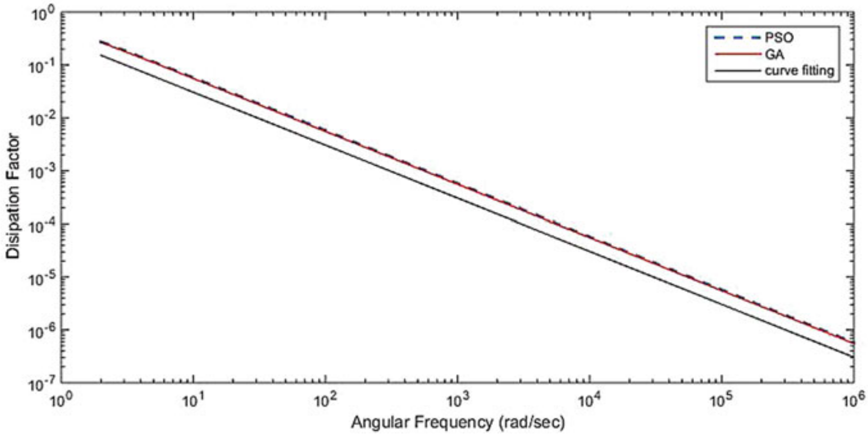


Fig. 6 Dissipation factor

6 Conclusion

In this paper the parameters of insulation RC model have been estimated using PDC test results using exponential curve fitting, GA and PSO algorithm and the dissipation factor of this insulation has been calculated in frequency with these parameters.

References

1. Saha TK, Purkait P, Muller F (2005) Deriving an equivalent circuit of transformers insulation for understanding the dielectric response measurements. IEEE Trans Pow Delivery 20(1)
2. Paithankar AA, Pinto CT (2001) Transformer insulation diagnosis: Recovery voltage measurement and DC absorption test. In: Proceedings of electrical insulation electrical manufacturing coil winding conference, pp 597–600
3. Kuchlar A, Bedel T (2001) Dielectric diagnosis of water content in transformer insulation systems. Eur Trans Electr Power 11(1):65–68
4. Gafvert U, Frimpong G, Fuher J (1998) Modeling of dielectric measurements on power transformers. Proceedings of CIGRE, Paris, France, paper no. 15/103
5. Frimpong G, Gafvert U, Fuher J (1997) Measurement and modeling of dielectric response of composite oil/paper insulation. In: Proceedings 5th international conference on properties and applications of dielectric materials, vol 1, pp 86–89
6. Leibfried T, Kacher AJ Insulation dagnostics on power transformers using the polarization and depolarization current analysis. In: Proceedings of IEEE international symposium on electrical insulation, pp 170–173
7. Hassing M, Braunlic R, Gysi R, Alff J, der Houhanessian V, Zaengel WS (2001) On-Site applications of advanced diagnosis methods for quality assessment of insulation of power transformers. In: Proceedings of IEEE annual report conference electrical insulation and dielectric phenomena, pp 401–447

8. er Houhanessian V, Zaengl WS (1997) Application of relaxation current measurements to on-site diagnosis of power transformers. In: Proceedings IEEE annual report conference on electrical insulation and dielectric phenomena, pp 45–51
9. Mohamad G, Nemeth E (1991) Computer simulation of dielectric processes. In: Proceedings of 7th international symposium on HV engineering, Dresden, Germany, pp 309–312
10. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: IEEE international conference on neural networks conference proceedings. Perth, Australia, vol 4, pp 1942–1948
11. Sun M, Shi Z (2004) Structural model updating based on particle swarm optimization. J Vibr Eng Chin 17:350–353
12. Shi Y, Eberhart RC (1998) Modified particle swarm optimizer. In: Proceedings IEEE conference on evolutionary computation proceedings ICEC, NJ, USA, pp 69–73

Modeling Electrical Arc Furnace (EAF) and Simulating STATCOM Devices for Adjusting Network Power Quality

Behrang Sakhaee, Davood Fanaie Sheilkholeslami, Mohammad Esmailee and Davood Nazeri

Abstract Electrical arc furnace is considered as great and intricate load in power systems. It is unbalanced, non-linear with time variable properties and consumes active and reactive power with a lot of oscillations. Meanwhile, one basic attempt to neutralize its undesired effects on the network is to model it accompanied by compensating network power quality. This paper studies the effect of STATCOM on EAF performance. In this article, a real model modeling random and chaotic properties of EAFs is used. All parameters of this model are modeled based on a real model of a 28.2 MW EAF. It is a dynamic model of EAF developed in MATLAB/SIMULINK environment. Moreover, a real STATCOM system model is simulated in MATLAB separately as the necessary reactive power compensator for this model. It is perceived from simulation results that the transient performance of EAF voltage is better when the line equipped with STATCOM than when it is without compensating devices.

Keywords Electrical arc furnace · STATCOM · MATLAB/SIMULINK

1 Introduction

Electrical Arc Furnaces are among great non-linear loads used in steel industry. Instant oscillations of active and reactive powers with great amplitudes in an EAF originate disturbances in the electrical network. This issue may cause problems for both steel makers and the servicing electricity network. Voltage flicker, harmonic and interharmonic are some problems caused by EAFs, to name a few [1]. Voltage

B. Sakhaee (✉)
Department of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
e-mail: Behrang.s@gmail.com

D. F. Sheilkholeslami
Department of Engineering, Azad University of Mashhad, Mashhad, Iran

M. Esmailee · D. Nazeri
Department of Management, Ferdowsi University of Mashhad, Mashhad, Iran

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_48

harmonics existence and current harmonics flowing within the network affect performance of the network's devices and equipment. Capacitors, electronic equipment, transformers, motors, and so on are elements affected by these harmonics. World standard institutes define harmonic permissible levels in order to preserve harmonic distortions level in an acceptable level and save equipment from their destructive effects. IEEE-519 is one of these standards [2]. It is necessary to perform harmonic measurement to determine harmonic sources and levels in different points. Subsequently, designing and implementing proper corrective solution requires expenditure and a lot of precision. Hence, it is essential to use computerized methods and simulations for harmonic analysis of a network so that time and money can be saved [3].

Therefore, when the EAF is connected to a power network and voltage flicker ranges are over the top that make the network unstable and cause voltage flicking, mitigation methods should be considered to correct such disturbances. In this vein, a comprehensive research has been carried out in Ref. [4] that presents scientific application of static synchronous compensators for improving Power Quality (PQ) and improving voltage flicker in EAF. Thus, obtaining a precise model in time field from which the effect of such loads connected to the power system can be studied is a vital issue for EAFs.

In various literatures, modeling most U-I properties have been recommended in both dynamic performance and steady state of the system. In Refs. [5, 6], non-linear time variable resistance models have been suggested for EAF in such a way that the arc has been simulated by limited band white noise and periodic sinusoid wave rules for flicker compensation suggestion. Conductance model of EAF has been presented for harmonic studies based on Cassie equation that shows one-line U-I properties at the monitoring stage. A time variable resistance model was proposed in Ref. [7] to study the initial stage of melting cycle so that the behavior of the electrical arc was shown by linear function of arc length with random variables. Several chaotic systems have been provided in Refs. [8, 9] to illuminate functional and dynamic behavior of EAF. Random methods have been proposed in Refs. [10] to model linear approximation of U-I properties in order to achieve voltage and current relation as well as obtaining reactive power consumption for enhancing compensating in EAFs. One serious problem of random methods is estimating and modeling the EAF random event especially in initial stages of melting. A precise neural-network-based model has been proposed in [11] for modeling the extreme non-linear feature of U-I of an EAF. The neural-network-based model can be employed more effectively in the form of disturbances wave, voltage oscillations and performance of reactive power compensating equipment concerning an EAF connected to a power system. A Discrete Wavelet Transform (DWT) and the Radial Basis Function Neural Network (RBFNN) method have been proposed in Ref. [12] to model properties of dynamic U-I of the EAF.

In this paper, a dynamic model of an electrical arc furnace (EAF) developed in MATLAB/SIMULINK environment is presented. The model is based on simulating electrical arc resistance per time unit whose behavior is considered random. This model is applicable in estimating voltage flicker in power transmission network in point of common coupling (PCC) due to the EAF performance [13]. Furthermore, a model of a real STATCOM system with 30 MVA capacity is modeled separately in MATLAB as the necessary reactive power compensator for this model.

2 EAF Model

An EAF unit model connected to transmission network with 110 kV voltage level, simulated in MATLAB/SIMULINK, is depicted in Fig. 1 in next page.

The EAF model simulated in MATLAB/SIMULINK is depicted in Fig. 2.

The EAF model simulated in MATLAB/SIMULINK is depicted in Fig. 2.

The EAF with 28.8 MW power range is connected to a 33 kV switching via 33/0.7 kV reducer transformer at the EAF. The 33 kV switching is connected to a 33 kV transmission network via an 110/33/6.3 kV power transformer. The equivalent of the 110 kV transmission network is presented for two various switching states by a voltage supply parallelized with a serial impedance wherein the network functional specification including three-phase short circuit current and X/R rate is shown in Eqs. (1) and (2).

$$I_{SC3} = 14.6kA, \quad \frac{X}{R} = 5.16 \tag{1}$$

$$I_{SC3} = 11.1kA, \quad \frac{X}{R} = 4.88 \tag{2}$$

Properties of some of the network main parameters are provided in Table 1.

The EAF unit is connected to a 33 kV switching via a cable with 800 m length (with $300/25\,\text{mm}^2$ specification). Power transformers are simulated with a three-phase power transformer model in MATLAB. Hence, coil resistance, leakage of coil's inductance and core magnetic specification are considered. Cabals are considered by their resistance and reactance. The modeling of components including power transformer, cabals and network equipment is accomplished by references that were discussed in introduction. In order to estimate the level of voltage flicker in power network, a critical element has to be modeled in EAF details. The proposed model has been explained in Sects. 2.1, 2.2 and 2.3.

2.1 EAF Mathematical Model

Various models of EAF have been illustrated in Refs. [14, 15] from which some present dynamic behavior of electrical arc resistance while others consider electrical arc random behavior. The crucial issue is that how both random and nonlinear behaviors of electrical arc are considered for various functional modes. This issue is important especially for estimation of the voltage flicker levels.

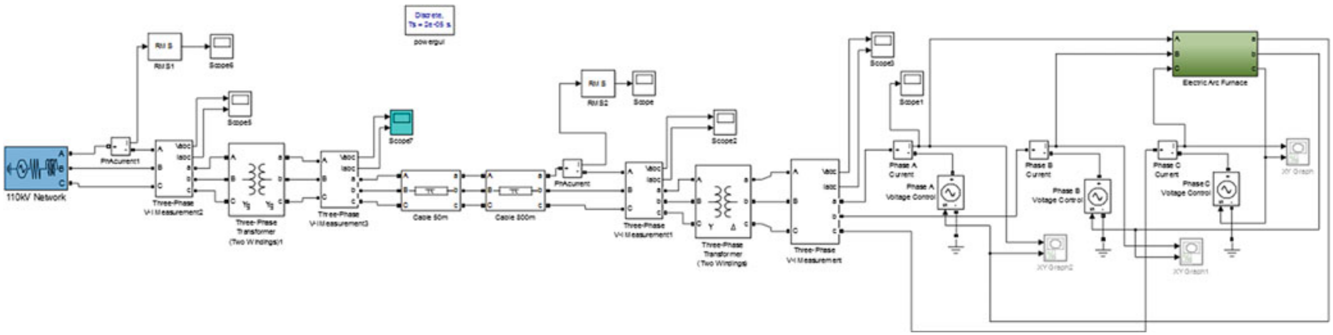


Fig. 1 Model of a substation connecting EAF to 110 kV transmission network in Matlab/Simulink

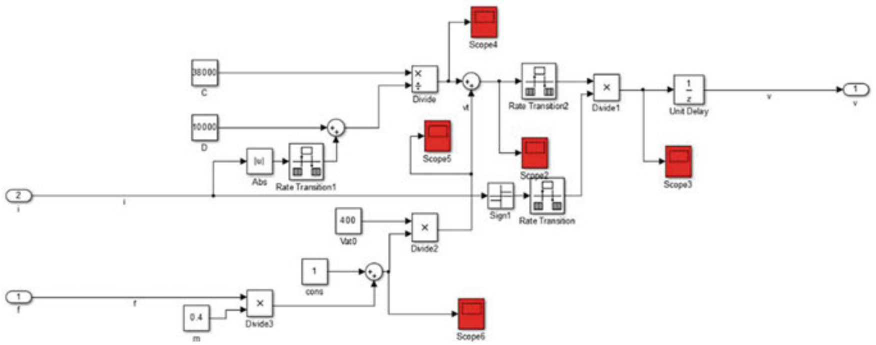


Fig. 2 EAF model simulated in MATLAB/SIMULINK

Table 1 Parameters of three-winding three-phase power transformer and two-winding three-phase EAF transformer

	Three-wending three-phase transformer	Two-wending three-phase transformer
Rated power	63 MVA/63 MVA/21 MVA	60 MVA
Rated voltage	Ur1/Ur2/ Ur3 = 110 kV ± 10 × 1.5%/ 33 kV/6.3 kV	Ur1/Ur2 = 33 kV/(0.7 kV–0.344 kV)
Rated current	In1/In2/In3 = 330.6 A/1102.2 A/1924.5 A	In1/In2 = (1050 A–630 A)/(60,400 A–49,500 A)
Voltage group	YNyn0d5	Yd
Short-circuit impedance	uk, 1-2 = 13.5%, uk, 1-3 = 7.5%	uk = 13.3–20.6%

In this section, the EAF is modeled in the form of a nonlinear resistance controlled by current using available functions in MATLAB in power system simulation section. The electrical arc current and its derivatives are input parameters to the function whereas the time variable nonlinear resistance and the current passes through are controlled. The melting process in EAF can be divided into three parts. At the first part, electrical arc begins to get extinct. Therefore, it is assumed in EAF model, that current and voltage of electrical arc reach the crossing zero point at same time, immediately. Thus, electrical arc voltage reduces until it reaches the u_{ig} area, in where it acts as an open circuit with an electrical circuit. During this session, there is a little leakage flow that flows in foamy slag of circuits with electrical arc. It is assumed here that the foamy slag has a constant resistance R_g whereas it is assumed that u_{ig} is proportional to the electrical arc length. At the second part, the electrical arc is created and a transient state is appeared in the voltage wave shape during melting process starts. This is because the arc resistance drops suddenly from u_{ig} to u_d. This process is described by an exponential function with a time constant τ_1. The electrical arc abates from the melting process at the

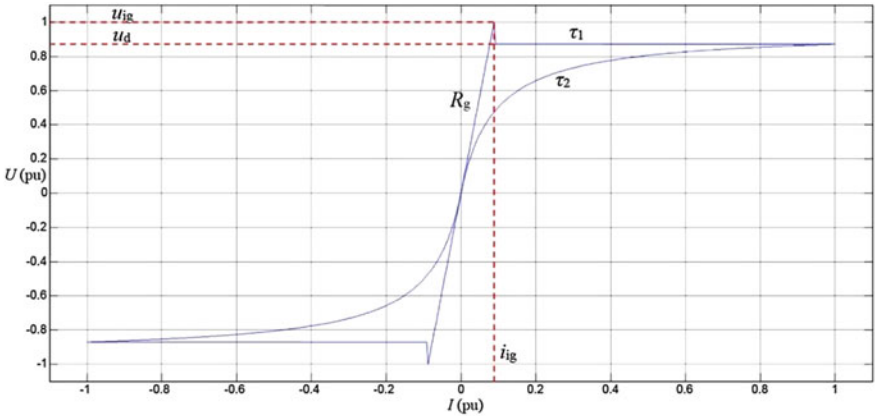


Fig. 3 Dynamic U–I characteristic of electric arc

third part. At this stage, voltage drop reduces with lower range of intense changes after arc abatement. This process is shown as an exponential function with the time constant τ_2. The non-linear behavior of the EAF R_g is shown as below:

$$R_a = R_g, \quad if\, 0 \le I < i_{ig}; \quad \frac{dI}{dt} > 0 \tag{3}$$

$$R_a = \frac{u_d + (u_{ig} - u_d)e^{\frac{-(I-i_g)}{\tau_1}}}{I}, \quad if\, I > i_{ig}; \frac{dI}{dt} > 0 \tag{4}$$

$$R_a = \frac{u_t + (u_{ig} - u_t)e^{\frac{-I}{\tau_2}}}{I}, \quad if\, \frac{dI}{dt} < 0 \tag{5}$$

$$I = |i(t)|; \quad u_{ig} = 1.15 u_d; \quad i_{ig}\frac{u_{ig}}{R_g}; \quad u_t = \left[\frac{I_{\max} + i_{ig}}{I_{\max}}\right] u_d; \quad \Gamma_2 = 2.\Gamma_1 \tag{6}$$

The obtained U-I properties of EAF are shown in Fig. 3 using above relations.

2.2 *Stochastic Behavior of EAF*

The EAF dynamic model must consider the random event of electrical arc to analyze voltage flicker. It is expected that the maximum voltage flicker happens when the electrical arc length is variable. This issue is related to the load current's area in the U-I feature. It is considered for this that the EAF voltage (U_{eaf}) changes randomly based on Eq. (7) in form of a sinus wave.

$$U_{eaf}(t) = \alpha\, U_{eaf}(1 + m\ \sin(\omega_f t)) \tag{7}$$

2.3 The EAF Model Parameters Estimation Process

Parameters R_g, u_{ig}, u_d, $I_{\max}$, τ_1 and τ_2 defined in dynamic properties of U-I, change by operational conditions of EAF's load. To determine these parameters, various active powers and operational conditions are considered for the EAF. Therefore, a process shown in Fig. 4 is carried out.

3 Introducing the Model Used for Simulating STATCOM

The generic base of the system is a parallel compensator depicted curtly in Fig. 5. The compensator is connected to load parallel through L_C and rest of the system have been simplified as limitless voltage source and the system impedance as $R_S + jwL_S$. The compensator is connected to the PCC via L_C whose voltage is point V_t. Capacitor C_f has been used as filter for removing harmonic of the i_C compensator's current. The compensator consumes or generates reactive power and no energy source is connected to the link DC. Voltage of the link DC (V_{dc}) is adjusted by the compensator [16].

Fig. 4 Procedure for estimation of EAF model parameters [13]

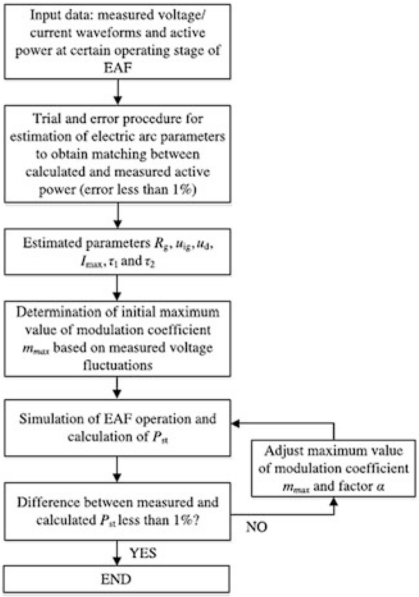
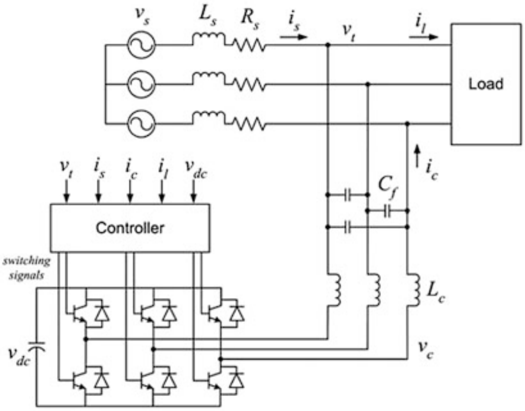

Input data: measured voltage/ current waveforms and active power at certain operating stage of EAF
Trial and error procedure for estimation of electric arc parameters to obtain matching between calculated and measured active power (error less than 1%)
Estimated parameters R_g, u_{ig}, u_d, I_{max}, τ_1 and τ_2
Determination of initial maximum value of modulation coefficient m_{max} based on measured voltage fluctuations
Simulation of EAF operation and calculation of P_{st}
Adjust maximum value of modulation coefficient m_{max} and factor α
Difference between measured and calculated P_{st} less than 1%?
NO
YES
END

Fig. 5 The overall diagram of the voltage and current compensating system using STATCOM



Supplying necessary amount of reactive power, the compensator can eliminate or reduce reactive power, harmonics and lack of balance in the load's current. Moreover, it can adjust the V_t voltage to a specified amount. The compensator is capable of accomplishing these tasks separately or together.

Should the three-phase voltage amplitudes and the three-phase current amplitudes are equal and angle between subsequent phases is $2 \times \pi/3$, voltage and current are balanced. There are two kinds of lack of balance titled load unbalance and voltage source unbalance from the perspective of a compensator connected to the load parallel. In the former kind, If the load i_1 is unbalanced, voltage V_t is unbalanced since voltage drop is different due to the unbalanced current. If the unbalanced load is compensated, the balanced current is obtained from the source and hence, the voltage is balanced and load current's lack of balance is compensated. In the latter type, the load generates unbalanced current due to unbalanced voltage V_t. This lack of balance is compensated by adjusting voltage by STATCOM. Three-phase voltage of PCC is balanced consequently and load's current will be balanced as well.

Plan of controlling compensating unbalanced current and method of controlling compensating unbalanced voltage and mixed control (voltage and current) have been presented. Current control diagram is demonstrated in Fig. 6. The reactive elements in load's current are calculated and are supplied by the compensator. Therefore, the balanced current and power factor unit are gained from the source. Output voltage of compensator vc has been controlled. Thus, the compensator's current follows the reference ic obtained from Eq. (8).

$$v_c^* = v_t + K_{P1}(i_c^* - i_c) + K_{I1}\int_0^t (i_c^* - i_c)dt \tag{8}$$

Output of the PI controller differs from voltage of invertor V_C and PCC voltage due to voltage drop on (L_C). At the steady state, the second phrase is zero in the Eq. (8) and the third phrase is voltage drop on the coupled (L_C).

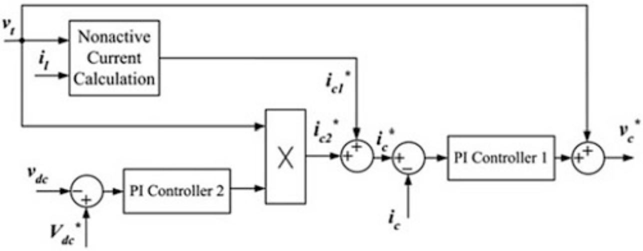


Fig. 6 Block diagram of current compensation

$$i_c^* = i_{c1}^* + i_{c2}^* \tag{9}$$

The compensating reference current of i_c^* has two elements i_{c1}^* and i_{c2}^*. The first element is the reactive element in load's current. It receives some active power from the source to cope with losses in order to adjust voltage of link dc. This active current is related to i_{c2}^* which is coherent with the PCC's voltage, V_t. Consequently, the control loop is designed like Eq. (10).

$$i_{c2}^* = \left[K_{P2}(V_{dc}^* - V_{dc}) + K_{12}\int\limits_0^t (V_{dc}^* - V_{dc})dt\right] v_t \tag{10}$$

Sum of i_{c1}^* and i_{c2}^* is a current i.e. i_c^* which has to be supplied by the compensator. It is mentioned in Eq. (10).

Most voltage unbalance in power systems are because voltage amplitudes are not equal. Voltage unbalance in this paper refers to inequality in voltage amplitude which is compensated by parallel compensation.

Base of controlling voltage compensator is that its output must be exclusively active power (losses are not considered). Its methodology is that the output voltage of compensator (V_c) should be coherent with V_t and the three-phase amplitude of voltage dc is controlled independently. Whenever V_c amplitude is greater than V_t in each phase, the compensator generates reactive power. If V_t amplitude is greater than V_c, the compensator consumes reactive power.

$$v_c^* = \left[1 + K_{P1}(V_t^* - V_t) + K_{I1}\int\limits_0^t (V_t^* - V_t)dt\right] v_t(\omega t - \theta^*) \tag{11}$$

In Fig. 7, the first input to the phase angle changing block is the reference voltage of the compensator calculated via adjusting needed voltage. The phase of this reference voltage of the compensator is changed in the concerned block diagram and phase angle θ^* changing is determined by control loop of the dc link voltage which is the second input to the phase angle changing block.

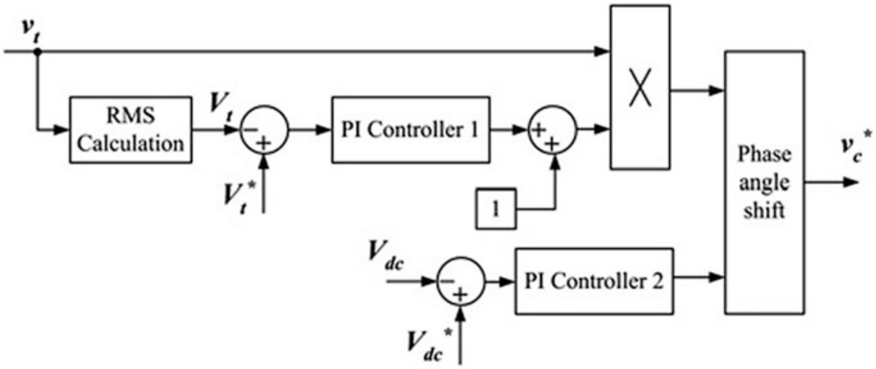


Fig. 7 Voltage control diagram

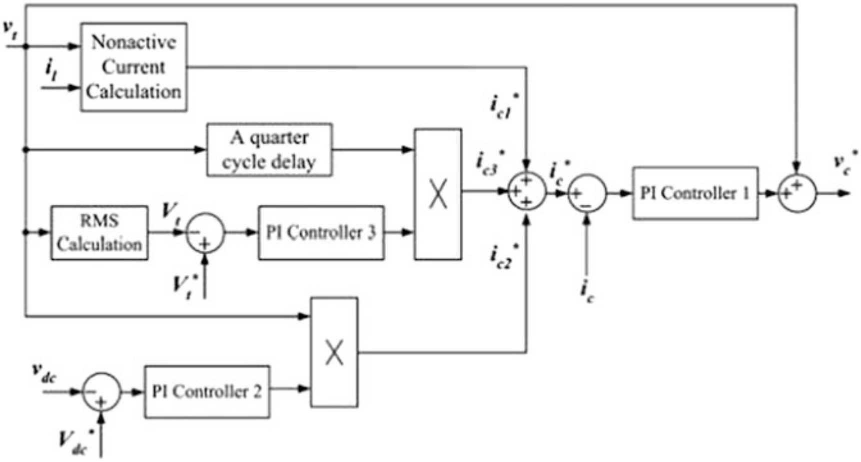


Fig. 8 Control block diagram

$$\theta^* = \left[K_{P2}(V_{dc}^* - V_{dc}) + K_{I2} \int_0^t (V_{dc}^* - V_{dc})dt \right] \tag{12}$$

The question emerged subsequently is that whether the compensator is capable of adjusting the voltage and compensate reactive power of the load concurrently? The answer is yes. This compensator is a mixed compensator and its control diagram is depicted in Fig. 8.

Fundamentally, the current control loop has been combined with the voltage adjusting loop in this structure. The reference current is a combination of three elements: the load reactive element i^*c_1, the voltage adjusting element i^*c_2, and the dc link controlling element i^*c_2 from which i^*c_3 is in Eq. (13):

$$i_{c3}^{*} = \left[K_{P3}(V_t^{*} - V_t) + K_{I3}\int\limits_{0}^{t} (V_t^{*} - V_t)dt \right] v_t(\omega t - \frac{\pi}{2}) \tag{13}$$

The compensation consists of combination of current control loop and voltage control loop together. These two controlling loops can work separately where one is active and the other is inactive, or they can work simultaneously. This issue brings flexibility for STATCOM. It can be observed from Fig. 8 that gain of PI controller 1 has to be determined initially to determine amplitude of the output voltage of the compensator to compensate the load reactive current. Then, gain of PI controller 2 corrects amplitude of output voltage of the invertor for the reactive power and also adjusting the voltage. Ultimately, gain of PI controller 2 is chosen to correct the phase angle of output voltage of the invertor.

In three controlling methods, the controller has been used in three phases in order to controlling current and voltage of each phase.

Kp for the controller is determined by how fast the STATCOM respond to step changes I voltage.

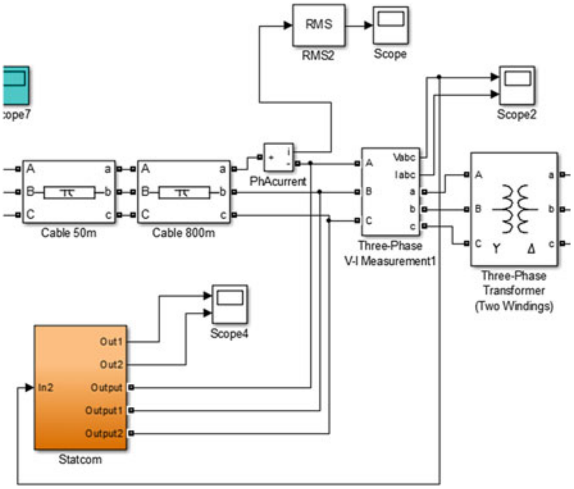
Greater KP leads to rapider response. But if it be very great, it causes over-shoot with the compensator.

KI is determined with the maximum discrepancy between two steady operational states. If KI be very great, problems emerged even with steady conditions.

According to the subjects presented, STATCOM model has been simulated in MATLAB/SIMULINK like Fig. 9.

Its internal circuit has been simulated as Fig. 10.

Fig. 9 Modeling of STATCOM in MATLAB



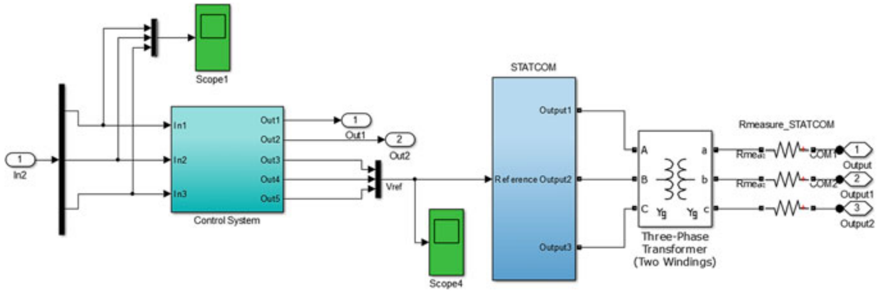


Fig. 10 Internal circuit of STATCOM

4 Aftermaths of Simulation

In no load conditions (Fig. 11) (EAF load separator) network's voltage and current chart is like Fig. 12.

As it is obviously observable from the above fig, when the network is exploited in no load conditions, voltage and current charts are presentable with high quality and no perturbation in the signal.

When non-linear load of EAF is connected to the aforementioned network, and charts of the network's voltage and current are extracted in this condition, charts will be like Fig. 13.

Regarding the aforementioned fig, one may clearly observe that in EAF performance condition and imposing its current to the network, EAF non-linear behavior creates conditions for lack of quality of the network's power which requires to be corrected rightfully. Should no solution is considered to cope with it, the harmonic current of this non-linear load is injectable even to the global network which may have undesired effects on the network.

One practical solution considered is to use power electronic compensating devices such as SVC, DVR, TCR, and STATCOM. Each of these devices has

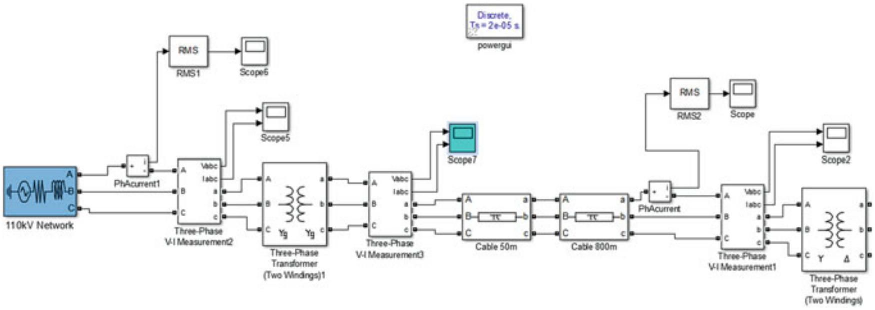


Fig. 11 No load conditions of power network

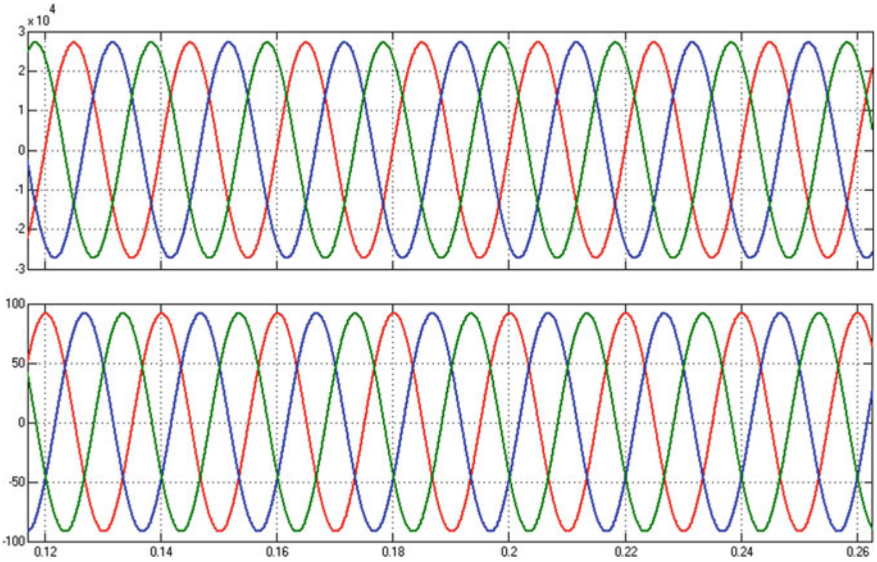


Fig. 12 Voltage and current in no load conditions of power network

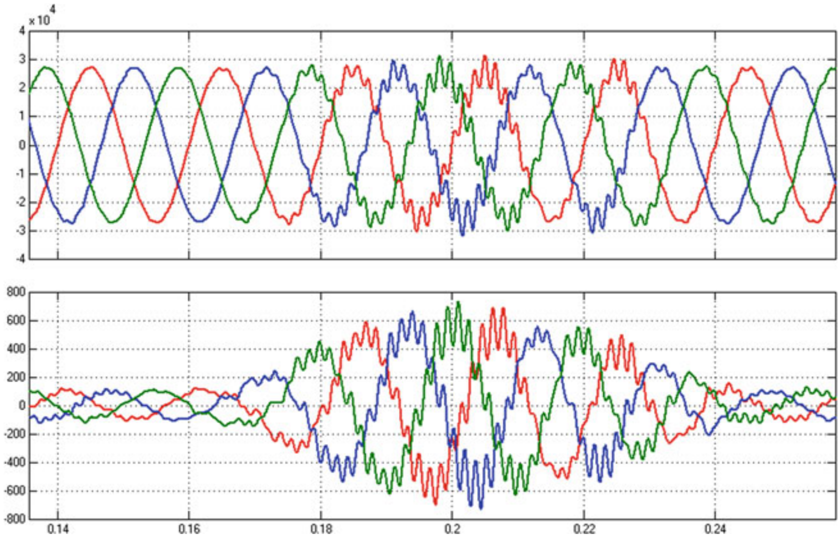


Fig. 13 Voltage and current in non-linear load of EAF conditions of power network without STATCOM

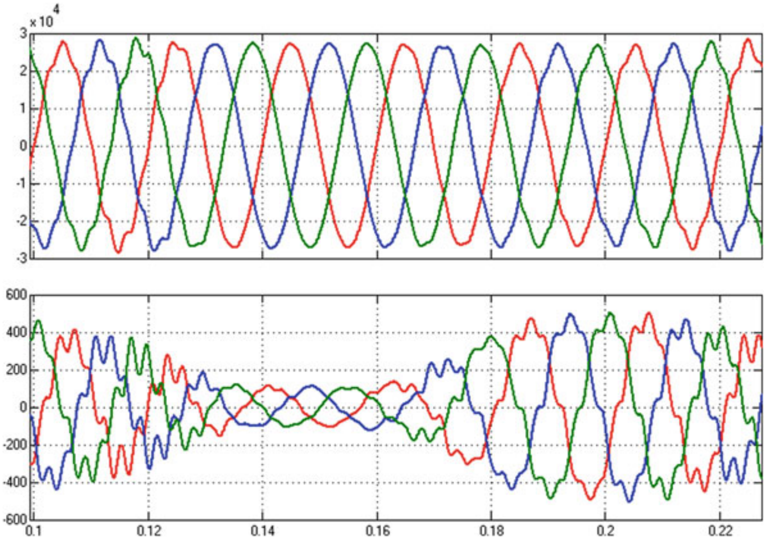


Fig. 14 Voltage and current in non-linear load of EAF conditions of power network with STATCOM

different advantages and disadvantages from which STATCOM is more popular due to being more affordable and its parallel connection to the network.

As it is observed from Fig. 10, the STATCOM's circuit generates the necessary amount of voltage for compensating by receiving feedback from the network's voltage and by injecting it to the network, prevents voltage changes from original form of the nominal voltage. Form of the network's current and voltage wave is like Fig. 14 when STATCOM is connected.

Form of the injected current to the network and voltage wave of STATCOM is as Fig. 15. This wave form is supplied per time unit to compensate the amount of network current and voltage by STATCOM.

It is deductible from the above fig that adjusting STATCOM devices optimally and connecting it to the network parallel, the possibility of compensating perturbation of the network voltage is provided to a high extent due to presence of EAF.

5 Conclusions

In this paper, a dynamic model of an Electrical Arc Furnace (EAF) simulated in MATLAB/SIMULINK environment has been developed. This model is based on simulating electrical arc resistance per time whose behavior is considered random. Regarding the simulation results, it is clearly observed that in EAF performance condition and imposing its current to the network, EAF non-linear behavior creates conditions for lack of quality of the network's power which requires to be corrected

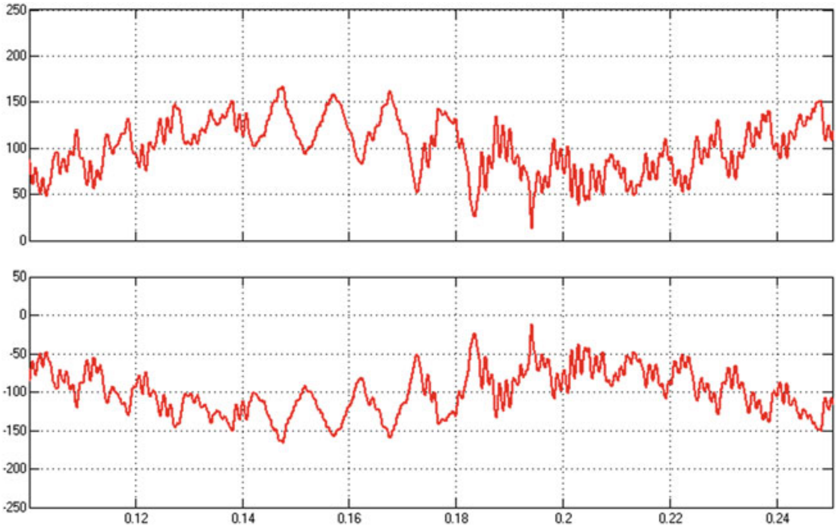


Fig. 15 Measured voltage and current waveforms

rightfully. For this purpose, a real system model of STATCOM was simulated separately in MATLAB as the compensator of the necessary reactive power for this model. The simulation results imply that by adjusting STATCOM devices optimally and connecting it to the network parallel, the possibility of compensating perturbation of the network voltage is provided to a high extent due to presence of EAF.

References

1. Dionise TJ (2014) Assessing the performance of a static var compensator for an electric arc furnace. IEEE Trans Ind Appl 03:1619–1629
2. Ji F, Zhou L, Yao G, Chen C (2005) Static var compensator on the method of synchronous symmetrical component. Proc CSEE 06:27–32
3. IEEE Std 519-2014 (Revision of IEEE Std 519-1992)—IEEE recommended practice and requirements for harmonic control in electric power systems
4. CIGRE Working Group B4.19 (2003) Static synchronous compensator (STATCOM) for arc furnace and flicker compensation. Technical brochure no. 237
5. Montanari GC, Loggini M, Cavallini A, Pitti L, Zaninelli D (1994) Arc-furnacemodel for the study of flicker compensation in electrical networks. IEEE Trans Power Delivery 9(4):2026–2036
6. Sharma H, McGranaghan M, Smith J (2008) An efficient module for flicker assessment of electric arc furnaces. In: Proceedings of the IEEE power and energy society general meeting, July 2008
7. Alonso MAP, Donsion MP (2004) An improved time domain arc furnace model for harmonic analysis. IEEE Trans Power Delivery 19(1):367–373

8. Horton R, Haskew TA, Burch RF (2009) A time-domain AC electric arc furnace model for flicker planning studies. IEEE Trans Power Delivery 24(3):1450–1457
9. O'Neill-Carrillo E, Heydt GT, Kostelich EJ (1999) Nonlinear deterministic modeling of highly varying loads. IEEE Trans Power Delivery 14(2):537–542
10. Carpinelli G, Iacovone F, Russo A (2004) Chaos-based modeling of DC arc furnaces for power quality issues. IEEE Trans Power Delivery 19(4):1869–1876
11. Cano-Plata EA, Ustariz-Farfan AJ, Soto-Marin OJ (2015) Electric arc furnace model in distribution systems. IEEE Trans Ind Appl 51(5):4313–4320
12. Chang GW, Shih MF, Chen YY, Liang YJ (2014) A hybrid wavelet transform and neural-network-based approach for modelling dynamic voltage-current characteristics of electric arc furnace. IEEE Trans Power Delivery 29(2):815–824
13. Teklić AT, Grˇci´ BF (2017) Modelling of three-phase electric arc furnace for estimation of voltage flicker in power transmission network. Electric Power Syst Res 146:218–227
14. Ozgun O, Abur A (2002) Flicker study using a novel arc furnace model. IEEE Trans Power Delivery 17(4):1158–1163
15. Bhonsle DC, Kelkar RB (2011) Simulation of electric arc furnace characteristics for voltage flicker study using MATLAB. In: International conference on recent advancements in electrical, electronics and control engineering, Sivakasi, India
16. Zhang X-P, Rehtanz C, Pal B (2012) Flexible ac transmission systems: modelling and control. Springer, Heidelberg

Distributed Generation Optimization Strategy Based on Random Determination of Electric Vehicle Power

Mohammad Ali Tamayol, Hamid Reza Abbasi and Sina Salmanipour

Abstract In this paper, an optimal strategy is presented for the participation of distributed generations in the energy market, considering the effect of uncertainty in the generation of wind turbines, uncertainty in market prices and uncertainty in the demand for electric vehicles. Virtual power plant is a collection of distributed generations that are co-located in order to participate in the market. The uncertainties make planning difficult for a virtual power plant. Four strategies have been proposed for participation in the energy market for the virtual power plant and the optimization problem has been solved with the help of the learning and training algorithm. In this strategy, the problem is solved by the probabilistic estimation method which has both an acceptable profit and needs a little time to perform calculations confirmed by the Monte Carlo method.

Keywords Electric vehicle · Distributed generation · Optimization

1 Introduction

Global concerns caused by increasing environmental problems on the one hand, and a significant increase in the need for electric power generation, increased instability in the forecasting and planning of generation and the like on the other hand, have resulted in distributed generation. As a result, electric vehicles are increasing sharply. In the future, electric vehicles will dominate the vehicle market in the world. Investigating the network in the presence of these vehicles is very important.

M. A. Tamayol (✉) · H. R. Abbasi · S. Salmanipour
Science and Act of Tohid Jam Co. (SATJCo.), Isfahan, Iran
e-mail: tamayol.ali@satjco.com; info@satjco.com
URL: http://www.satjco.com

H. R. Abbasi
e-mail: abbasi.hamid@satjco.com

S. Salmanipour
e-mail: salmanipour.sina@satjco.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_49

In this paper, the term "Plug-in Hybrid Electric Vehicle" (PHEV) is referred to as vehicles with bidirectional power transmission. If a significant number of networked vehicles are coordinated and managed under the control of a vendor, such as the electric car parking lot, they can act as a small virtual plant (VPP) with a very high startup speed and no set up cost. Also, these types of vehicles receive energy from the grid when charging their batteries [1] (G2 V).

The advantages and disadvantages of a wide range of PHEVs in the network have been studied in several articles. In [2–4], the issue of the participation of units has been optimized by considering the connection of electric vehicles to the network using particle algorithm. In these papers, no attention has been paid to the duration of the connection of vehicles to the network. In these papers, the supply body is not obliged to observe the minimum amount of battery charge of the car when it is disconnected from the network. In [5], the effect of PHEV (Plug-in Hybrid Electric Vehicle) on network load curves, generation capacity, costs, and emissions of greenhouse gases has been investigated. Further, the effect of the presence of PHEV on minimizing the distribution network losses from the viewpoint of the distribution network operator, as well as the connection of electric vehicles to the network in order to provide the reservation and frequency regulation, have been examined in [6, 7].

In most of the papers that have studied the planning of electric vehicle charging, optimization has been made through the network operator [8, 9]; while it is possible to optimize the charging of vehicles through electric vehicle parking lots in order to provide more materials for consumers, lower the cost of charging the vehicles, improve distribution network profile. In order to achieve this goal, electricity tariffs, which are a function of the current momentum of the electricity market, should be designed in such a way that electric vehicle parking change their consumption patterns in order to flatten the load curve in peak hours.

In this paper, the charging the PHEV battery is carried out in a parking lot with the goal of minimizing parking costs using a random-based planning method. The program is put into operation with a controller in the parking and V2G1 is also considered for this parking lot. Also, the effects of important factors such as the time spent leaving PHEV from parking, the distance traveled daily, the time to return PHEV to the parking lot, as well as the traffic conditions and driving habits that should be considered in planning of charging the PHEV, are modeled using probabilistic random functions. Finally, the effect of the presence of electric car parking with and without charge control is evaluated on costs and peak load curves.

2 Introducting Electric Vehicles

Considering the oil crisis and rising prices of fossil fuel in recent years, especially in industrialized countries, the issue of replacing existing vehicles has been taken seriously. Failure over long distances and the need for recharging is the most important disadvantage of electric vehicles. In contrast, the low cost of electrical energy and environmental issues are the most important advantages they provide. The importance of shifting the transport energy from fossil fuels to electricity is evident, since the transportation sector accounts for about two thirds of the oil consumption and 97% of the energy consumed in this sector depends on fossil fuels.

With the rise of electric vehicles and the need for daily charging, they can be introduced as new and big loads for the power grid. But these loads have the characteristics that separate them from other cases which is the ability to store power and inject into the network as a distributed power generation [2].

The patterns of using private vehicles vary from one country to another, but on average, these vehicles are typically used less than 4–5 h a day. In fact, vehicles are kept in the parking lot for about 20 h. Studies have shown that the greatest reduction in fuel consumption, the lowest limit, high efficiency, and most importantly, the general public's acceptance is made possible by PHEV vehicles. In the following, we summarize the various types of technologies in electric vehicles [5].

2.1 Electric Vehicles (EV)

These vehicles have an electric engine along with batteries for supplying electric power, and the energy of the batteries is used as a driving force for the electric engine of the car as well as for the supply of energy for other equipment. The batteries can be recharged by connecting to the power grid and the energy generated by the car's brakes, and even from non-grid electrical sources such as solar cells. The main advantages of these vehicles are:

Absolutely free of greenhouse gas emissions

- Producing very low noise
- Their efficiency is much higher than the internal combustion engines.
- The price of their electric engines is low.

The main disadvantages of these vehicles are fully dependent upon the battery (whose technology has not yet reached the capacity and density of energy comparable to fossil fuels).

2.2 Hybrid Electric Vehicles (HEV)

These vehicles have both a fuel-driven engine and an electric engine with enough battery (1–3 kW) to store fuel from the fuel-driven engine and the vehicle's brakes. The batteries come at the required time to help with the vehicle to produce auxiliary power or provide the propulsion force for the vehicle at low speeds when the fuel-driven engine is turned off.

Over the past decade, about one and a half million vehicles have been sold. In developed countries such as the United States, about 3% of vehicles are hybrid. Disadvantages of these vehicles are as follows:

- Failure to charge the batteries from the power grid
- Dependence on the fossil fuel consuming engine (the inability to move the car only using an electric motor).

2.3 Plug-in Hybrid Electric Vehicles (PHEV)

These vehicles designed to overcome the disadvantages of hybrid electric vehicles are rechargeable through the network and therefore require more batteries than they do. There's a complete fossil fuel system in the vehicles. Plug-in hybrid electric vehicles have more batteries than HEVs (about 5 times). The essential difference between these two types of electric car batteries is that the PHEV batteries should be capable of rapid discharge and fast charging, while HEV batteries operate only when fully charged and discharge rarely occurs in them. The price of PHEV batteries is between 1.3 and 1.5 times the cost of EV batteries, however, the total cost of batteries in PHEV vehicles is less than EV due to lower battery life. The following criteria can be mentioned for these vehicles:

With a mass production of batteries, it costs $750 per kWh; the total cost of the batteries will be about $6000 for a medium-range vehicle (40 km with a battery of 8 kWh).

With a life span of 200,000 km, the saved cost of fuel is about $ 4000 which is less than the battery life.

Reducing battery costs to $500 per kWh creates a common competition for PHEVs and gasoline vehicles.

2.4 Fuel Cell Vehicles

2.4.1 Simple Fuel Cell Vehicle

In fuel cell vehicles, the cell itself and fuel cell are sources of power generation and no auxiliary battery is used. A fuel cell similar to the battery of the electric vehicles generate the required electric power and propulsion power. The propulsion system includes a reverser to convert the fuel cell flow from DC to AC with variable frequency and voltage, an AC rotor and a power transmission system from the engine to the wheels of the vehicle.

2.4.2 Hybrid Fuel Cell Vehicles

A hybrid fuel cell has a battery with a high-capacity capacitor parallel to the fuel cell system. The hybrid fuel cell simultaneously utilizes the highest efficiency of fuel cell energy and high power within the battery. When the energy consumption rate is high (e.g. acceleration mode), the power required by the vehicle will be supplied by the battery and fuel cell set. When the energy consumption rate is low (e.g. moving in the street), the fuel cell supplies the required power. Batteries will be charged during low power consumption. Therefore, the fuel cell is designed for normal movement and battery life with maximum power to supply the power and energy needed. Selecting a battery set depends on factors such as fuel cell costs and performance, battery technology and movement cycle. Using the battery allows quick start up to the fuel cell vehicle and protects it against the reverse reaction of fuel cells during fuel performance and fuel consumption. In addition, the battery supplies the maximum power needed. And the energy generated can be recovered. The response time of the vehicle system to load changes in the presence of battery is faster. The hybrid fuel cell has a good performance, long operating time and a quick refueling time. It can cover an acceptable distance.

Other advantages of the battery include:

- No need for pre-heating fuel cell for starting the vehicle.
- The ability of the vehicle to operate in a fully electric mode during the time when the system is not capable of operating at its nominal voltage level.
- Much faster response time for load changes

Some of the disadvantages of the presence of battery include the cost, weight, recharge times that are imposed on the car. The cost of a single battery set is usually proportional to the amount of energy that it can store and the cost of the fuel cell is proportional to the amount of power required. Therefore, a high-capacity battery and medium energy storage may be a bit expensive.

3 Application Requirements of the Joint Vehicles

Queuing theory [10] can be used to describe the charging process of several vehicles connected simultaneously to a bus. Different queues can be selected for different charging conditions. Here are two things to consider: one assumes that several vehicles will be charged at the same station, and the other will charge several vehicles in a residential complex. The difference between these two modes is the number of vehicles that are being charged at the same time. At a station, it is assumed that the number of applicants required for charging at that point is unlimited. But it is assumed that a limited number of residents simultaneously demand a charge in a residential complex due to private ownership. The limited or unlimited number of charge applicants at a single point leads to the selection of two different queues in queuing theory. In accordance with this theory, the probability distribution of the number of vehicles that are being charged at the same time at a station and complex respectively follows from Eqs. (1) and (2).

$$p_n = \begin{cases} \left(\sum_{i=0}^{c-1} \frac{(c\rho)^i}{i!} + \frac{(c\rho)^c}{c!} \cdot \frac{1}{1-\rho} \right)^{-1} & n = 0 \\ \frac{(c\rho)^n}{n!} .p_0 & n = 1, 2, \ldots, c. \end{cases} \tag{1}$$

$$p_n = \begin{cases} \left(\sum_{i=0}^{c} \binom{N_{\max}}{i} .(c\rho)^i + \sum_{i=c+1}^{k} \frac{N_{\max}!.(c\rho)^i}{(N_{\max}-1)!c!c^{i-c}} \right)^{-1} & n = 0 \\ (c\rho)^n . \binom{N_{\max}}{n} .p_0 & n = 1, 2, \ldots, c. \end{cases} \tag{2}$$

In the above equation, ρ determines the length of the queue. By choosing this number less than one, one can be sure that the number of vehicles waiting to receive the service remains within the range. c represents vehicles that are being charged at the same time. k shows the number of vehicles in the queue. Nmax represents the maximum number of vehicles which is, in fact, the capacity of the charging location. N is the number of vehicles that are currently being charged.

It can also be assumed that the charging time of each vehicle varies exponentially in the interval [Tmin, Tmax].

$$T = \begin{cases} T_{\min} & T \leq T_{\min} \\ -T.\ln(U) & T_{\min} < T < T_{\max} \\ T_{\max} & T \geq T_{\max} \end{cases} \tag{3}$$

In addition, reference [10] states that there are currently only three levels for charging vehicles that are assumed for a level 3 station with a voltage of 400 V and a current of 63 amps as well as a residential complex of Level 1 with a voltage of 230 V and current of 16 amps. Knowing the charging level of the voltage V and the maximum charge current, Imax is determined. As a result, the current by the vehicle for charging I is determined according to the following equation.

$$I = \min\left\{ \frac{D_E}{V.T}, I_{max} \right\} \tag{4}$$

Now, knowing the flow rate for n number of vehicles that are being charged simultaneously, the total power demand is as follows.

$$P = \sum_{i=1}^{n} V.I_i \tag{5}$$

After the above definitions, the power demanded for charging vehicles can be obtained by a randomized simulation of the probable distribution in accordance with the following steps:

According to the Eqs. 4 and 5, we randomly generate the number of vehicles being simultaneously charged and repeat the following steps:

According to market penetration rate, the vehicle category is randomly selected.

According to the Eq. 2, the vehicle parameters are randomly determined.

The T charging time is randomly obtained according to the Eq. 3.

According to the Eq. 4, the charge current is determined.

By summing up the power of each vehicle, the total demand is calculated in terms of 6.

The above steps are repeated to ensure that sufficient samples are produced. By performing the above simulations and producing a sufficient number of samples, we conclude that a probable distribution station will be in accordance with what is shown in Fig. 1 [9], and this distribution is normal for a residential complex in Fig. 2. The results are presented in Ref. [5] in detail.

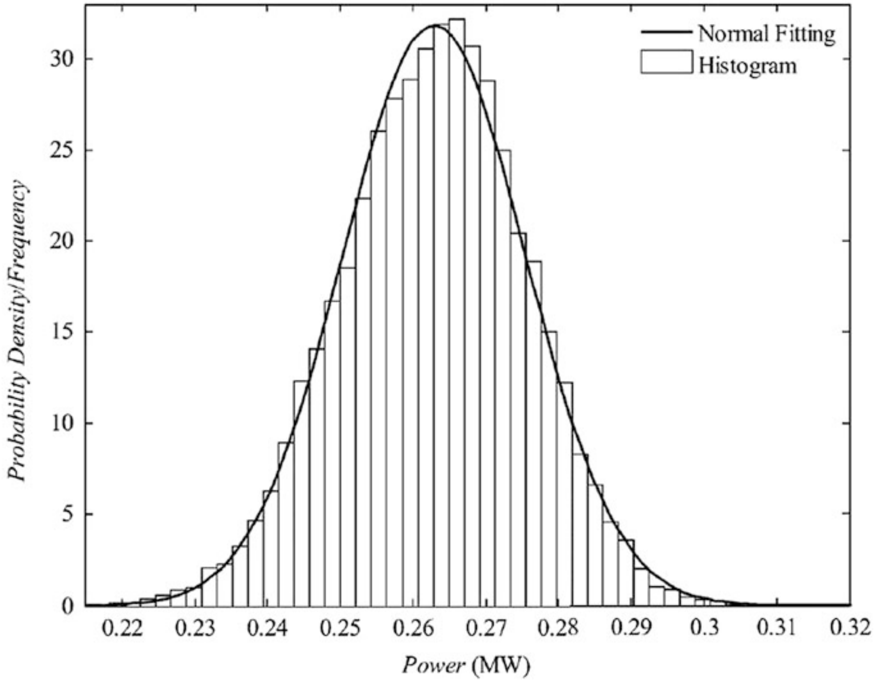


Fig. 1 Probabilistic distribution of power demand for an electric vehicle in a residential complex

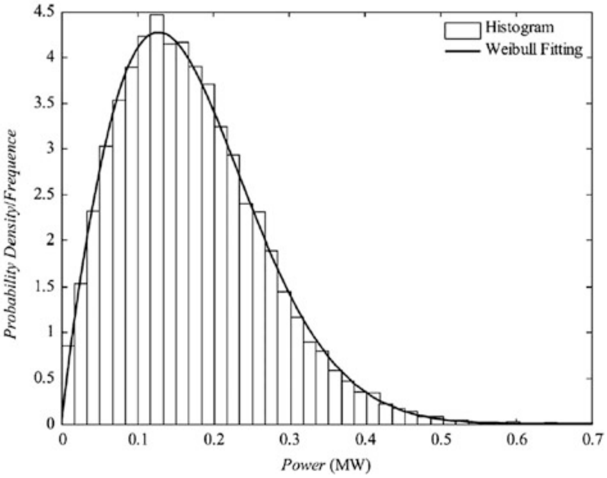


Fig. 2 Probabilistic distribution of power demand for an electric vehicle in a charging station

4 Simulation

Four different strategies have been proposed and simulated in solving the virtual plant planning problem that are presented as follows:

4.1 Solving the problem definitively
4.2 Using actual distribution functions and solving the problem by the Monte Carlo method
4.3 Using normal distribution function and solving the problem by PEM method
4.5 Using normal distribution function and solving the problem by Monte Carlo method

The output power of each of the DG networks is shown in the following Figs. 3, 4, 5 and 6.

5 Conclusion

In this paper, a methodology is presented based on randomized planning for charging the parking of electric vehicles with V2G capability. Planning specifically for charging PHEV vehicles has taken place in the parking lot with the goal of minimizing the cost of parking charge. A parking space with a capacity of 1000 PHEV was considered for the purpose of investigating the corresponding strategy and the random probability functions of each vehicle were generated using the appropriate probability distribution functions. The results of numerical studies

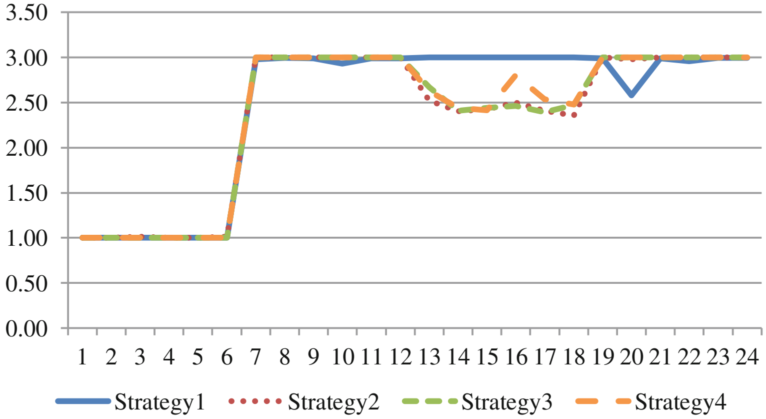


Fig. 3 DG1 output power in different strategies

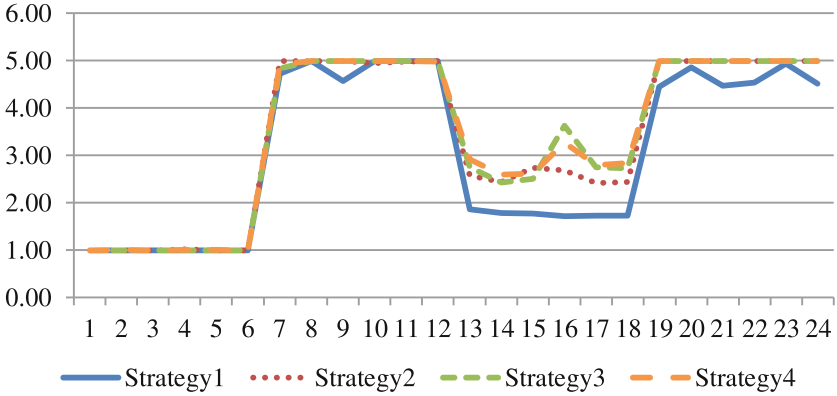


Fig. 4 DG2 output power in different strategies

showed that if there is no charge control on the PHEV, the cost of parking as well as the peak load of the network will be at the highest level. With a controllable charge without V2G capability, both peak load and the cost are significantly reduced. When there is V2G capability in addition to charging control, the cost of peak load is reduced in addition to the reduced cost of the network. According to the results, the necessity of controlling the charge of the PHEV vehicles and the V2G capability is confirmed.

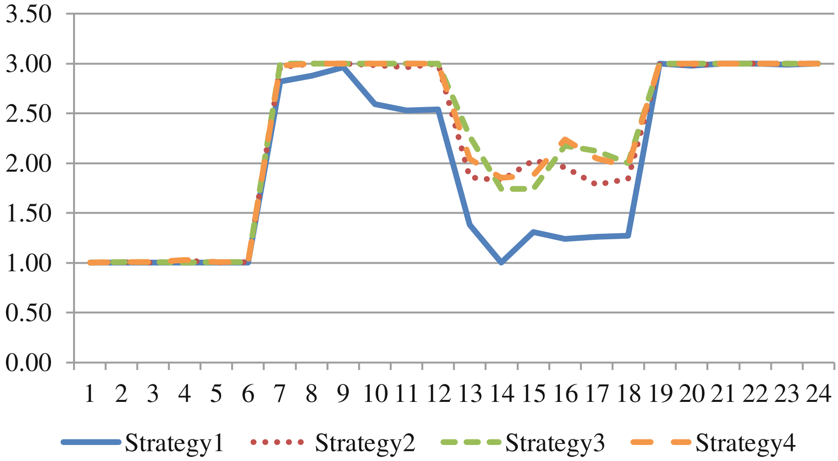


Fig. 5 DG3 output power in different strategies

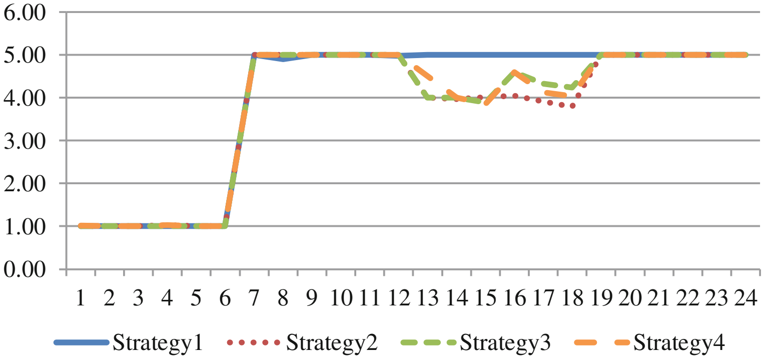


Fig. 6 DG4 output power in different strategies

References

1. Rajakaruna S, Shahnia F, Ghosh A (2016) Plug in electric vehicles in smart grids. Springer, Singapore
2. Falahi M et al (2013) Potential power quality benefits of electric vehicles. IEEE Trans Sustain Energy 4(4):1016–1023
3. Li Y, Crossley PA (2014) Monte Carlo study on impact of electric vehicles and heat pumps on LV feeder voltages 12–13
4. Bahrami S, Parniani M (2014) Game theoretic based charging strategy for plug-in hybrid electric vehicles. IEEE Trans Smart Grid 5(5):2368–2375
5. James JQ, Li VOK, Albert YS (2013) Lam Optimal V2G scheduling of electric vehicles and unit commitment using chemical reaction optimization. In: 2013 IEEE congress on evolutionary computation (CEC). IEEE, New York

6. Celli G, Pilo F, Pisano G, Soma GG (2005) Optimal participation of a microgrid to the energy market with an intelligent EMS. In: The 7th international power engineering conference. IPEC 2005, pp 663–668
7. Kamel RM, Chaouachi A, Nagasaka K (2010) Carbon emissions reduction and power losses saving besides voltage profiles improvement using micro grids. Low Carbon Econ 1(1):1–7
8. Akhtar Z, Chaudhuri B, Hui SYR (2017) Smart loads for voltage control in distribution networks. IEEE Trans Smart Grid 8(2):937–946
9. Wu X et al (2016) Stochastic control of smart home energy management with plug-in electric vehicle battery energy storage and photovoltaic array. J Power Sources 333:203–212
10. Li G, Zhang X-P (2012) Modeling of plug-in hybrid electric vehicle charging demand in probabilistic power flow calculations. IEEE Trans Smart Grid 3(1):492–499

An Improved Harmony Search Algorithm to Solve Dynamic Economic Load Dispatch Problem in Presence of FACTS Devices

Panteha Hashemi and Navid Eghtedarpour

Abstract One of the important optimization problems in power systems operation is Dynamic Economic Load Dispatch (DELD). Economic load dispatch for a time interval of few hours is done considering the constraints related to maximum rate of change of active power generated by units and other system constraints such as prohibited operating zones and the valve effect. These constraints cause the optimization problem to be non-smooth and non-convex. An improved Harmony Search Algorithm (HSA) is presented in this paper to solve the DELD problem in the presence of Flexible AC Transmission System (FACTS) devices considering the above mentioned constraints. The FACTS devices considered in this paper are Static VAR Compensator (SVC) and Thyristor Controlled Series Compensator (TCSC). The proposed algorithm is evaluated on IEEE 30-Bus Test System. Results show high strength of the proposed method for solving the DELD problem.

Keywords Dynamic economic load dispatch · FACTS devices
Harmony search algorithm · The roulette wheel method · Power system constraints

1 Introduction

Dynamic economic dispatch (DELD) is one of the main optimization problems in power system operation. The main objective is to determine how optimally the load can be dispatched between the generating units in order to minimize the total

P. Hashemi (✉)
Department of Electrical and Computer Engineering, Babol Noshirvani
University of Technology, Babol, Iran
e-mail: panteha.hashemi@yahoo.com

N. Eghtedarpour
School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran
e-mail: eghtedarpour@shirazu.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_50

system cost over the entire scheduling periods while satisfying the system constraints [1]. The DELD problem along with different constraints presents a complicated non-smooth and non-convex optimization problem [2] which has usually large dimension. On the other hand for the better utilization of existing power systems and restricting the system expansion via flexible ac transmission system (FACTS) devices are immensely used in today power systems [3, 4]. While FACTS devices are used today, the utilization of these devices add the complexity in the power flow problem and the optimizations along with them. Different methods have been proposed to solve the DELD problem. Conventional optimization methods include quadratic programming [5], dynamic programming [6] linear programming [7], and the interior point method [8, 9]. These techniques are usually computationally efficient, but they suffer mainly from non-convex cost functions [1]. Besides, to solve DELD using these numerical methods, the problem should be continuous; which is not in case of considering the ramp rate limit, prohibited operating zones and valve point loading effect that are exist in DELD problem [10–12].

The computational drawbacks of these traditional methods have forced the researchers to develop various meta-heuristic algorithms [13–26]. The operation of meta-heuristic algorithms does not depend greatly on the objective function structure and the problem constraints. The strength of these algorithms in finding the zones of the solution space which are likely to contain the global optimum solution is mentionable. In [13], DELD problem has been solved using Genetic Algorithm (GA). To compensate the weaknesses of GA, such as premature convergence, in [14–16] Particle swarm and wolf optimization has been used to solve the DELD problem. In [17] prohibited operating zones of generating units have been considered in DELD problem. Other optimization methods like Imperialist Competitive Algorithm [18], Bacterial Foraging [19], and Tabu Search Algorithm [20], Simulated Annealing technique [21] and also combined optimization techniques i.e. Bee Swarm Optimization [22] have also been reported in literature to solve DELD problem. However, in these evolutionary based algorithms the process involved in choosing the control parameters is not straightforward [23]. Optimization based on Harmony Search (HS) was firstly introduced in [24] which is used for many physical optimization problems [25–28]. The HS presents a derivative-free, meta-heuristic algorithm, reveals the improvisation process of music players. Since its development, HS has been widely used in different power system optimization problems [1, 2, 29]. However, its performance on complex and high-dimensional problems is not well proved [2, 30, 31]. In order to overcome these drawbacks, several modifications are presented in recent years. An improved HSA based on the dynamic change of some HSA parameters has been given in [1]. Another version of improving the HSA by changing its parameters is given in [2]. In this paper the parameters of the HSA are changed linearly or exponentially. The amount of performance improvement of the algorithm is illustrated by examining and comparing the obtained solutions using this algorithm in sample optimization problems. In many cases, hybrid algorithms are created from HSA. For example, in [28] good results are achieved by combining HSA with particle swarm algorithm.

In fact, because of the high ability of particle swarm algorithm in finding the exact numeric value of the optimum solution and the high capability of HSA in finding the regions with high potential to contain the global optimum solution, the combination of these two algorithms can lead to creating a powerful algorithm for finding the global optimum solution of the optimization problem.

In this paper, we address the HSA weakness in finding the exact global optimum point (the numeric value), and also the unclearness of optimum settings of its parameters (like all other meta-heuristic algorithms) especially in the presence of FACTS devices; and present a new method for setting the algorithm parameters which improves the performance of the algorithm in finding the exact numeric value of the global optimum solution in the solution space of the problem.

The proposed HSA is used to solve the DELD problem considering all common power system constraints in presence of FACTS devices. The proposed algorithm is tested on different test functions, and the results are shown in tables; these solutions show the improvement of the algorithm performance in finding the exact value of the optimum solution. The rest of the paper is organized as follows. DELD formulation is presented in Sect. 2. This part contains the objective function and constraints. The proposed improved HSA is introduced in part III. Simulation results for three different scenarios of 30-bus IEEE system are presented and discussed in part IV. Part V concludes the whole paper.

2 Dynamic Economic Load Dispatch

One of the most important power system operating problems is DELD. The purpose is to find the set of active power generation setting of generating units such that the total operating cost of the system be minimized while operating constraints are met.

Different power generation cost of generating units, many practical constraints regarding the operating of these units, security constraints of power system and the existence of tap changer and FACTS has made DELD a complex problem. Many papers have considered some of these constraints to solve DELD problem but just few of them have considered all of the constraints. This causes their results impractical. For instance, FACTS parameters optimizing are considered in some papers [19–21] to solve economic load dispatch. But less attention is paid to DELD problem.

In this paper, decision variables are defined as:

$$X = [P_{m,t}, V_{m,t}, T_{k,t}, B_{n,t}, QS_{w,t}, R_{z,t}]. \tag{1}$$

Decision variables defined in vector X are active power generation of unit m, bus voltage of generating unit m, tap position of transformer k, susceptance of capacitor unit n, reactive power injected by SVC w, and reactance of TCSC z, respectively, all at time t.

The optimization problem of DELD with related constraints is defined in (2)–(15). Objective function of DELD problem considering valve point effect is defined in (2) [18].

$$\begin{aligned}\min F_G &= \sum_{t=1}^{T}\sum_{m=1}^{M} F_{m,t}(P_{m,t}) \\ &= \sum_{t=1}^{T}\sum_{m=1}^{M} a_m P_{m,t}^2 + b_m P_{m,t} + c_m + \left| d_m \sin(e_m(P_m^{\min} - P_{m,t}))\right| \end{aligned} \quad (2)$$

In this relation, $F_{m,t}$ defines cost function of unit m, T is the time span of the DELD problem, M is the total number of generating units, $P_{m,t}$ is the active power generated by unit m, constants a_m to d_m are multipliers modeling the cost function of unit m, and $P_m^{\min}$ is lower bound of active power generation of unit m, all at time t.

Equality and inequality constraints considered in DELD problem consist of power balance equality constraints, security constraints, ramp rate constraints, generated active power constraint, prohibited operating zones constraint, reactive power generation constraint, tap position constraint and capacitor susceptance constraint. These constraints are:

$$P_{i,t}^{gen} - P_{i,t}^{dem} = \sum_{j\in\Omega_{i,j}} V_{i,t}V_{j,t}Y_{ij,t}\cos(\delta_{i,t} - \delta_{j,t} - \theta_{ij,t}), \quad \forall i,t \quad (3)$$

$$Q_{i,t}^{gen} - Q_{i,t}^{dem} = \sum_{j\in\Omega_{i,j}} V_{i,t}V_{j,t}Y_{ij,t}\cos(\delta_{i,t} - \delta_{j,t} - \theta_{ij,t}), \quad \forall i,t \quad (4)$$

$$V_i^{\min} \le V_{i,t} \le V_i^{\max}, \quad \forall i,t \quad (5)$$

$$\left|S_{l,t}\right| \le S_l^{\max}, \quad \forall l,t \quad (6)$$

$$P_{m,t} - P_{m,(t-1)} \le UR_m, \quad \forall m,t \quad (7)$$

$$P_{m,(t-1)} - P_{m,t} \le DR_m, \quad \forall m,t \quad (8)$$

$$\max\left(P_m^{\min}, P_{m,(t-1)} - DR_m\right) \le P_{m,t} \le \min\left(P_m^{\max}, P_{m,(t-1)} + UR_m\right), \quad \forall m,t \quad (9)$$

$$\begin{aligned} &P_m^{\min} \le P_{m,t} \le P_{m,1}^L \\ &\quad\vdots \\ &P_{m,(q-1)}^U \le P_{m,t} \le P_{m,q}^L \\ &\quad\vdots \\ &P_{m,Npoz}^U \le P_{m,t} \le P_m^{\max}, \quad \forall m,t \end{aligned} \quad (10)$$

$$Q_m^{\min} \le Q_{m,t} \le Q_m^{\max}, \quad \forall m, t \tag{11}$$

$$T_k^{\min} \le T_{k,t} \le T_k^{\max}, \quad \forall k, t \tag{12}$$

$$B_n^{\min} \le B_{n,t} \le B_n^{\max}, \quad \forall n, t \tag{13}$$

The concept of prohibited operating zone imposed to the generating unit's cost function is depicted in Fig. 1.

In constraints (3)–(10), $V_{i,t}\angle\delta_{i,t}$ and $V_{j,t}\angle\delta_{j,t}$ are voltage of buses i and j respectively, $Y_{ij,t}\angle\theta_{ij,t}$ is admittance of the line connecting buses i and j, $S_{l,t}$ is the apparent power flowing through line l, and maximum increasing and decreasing ramp rates of generating unit m are defined with UR_m and DR_m respectively, all at time t. N_{POZ} is the total number of prohibited operating zone limits, $P_{m,q}^{L}$ and $P_{m,q}^{U}$ are lower and upper limits of prohibited operating zone q. In (11) $Q_{m,t}$ is the reactive power generated by unit m at time t. Inequality constraints regarding tap position of transformers and reactive power injection of capacitors are shown in (12) and (13), respectively.

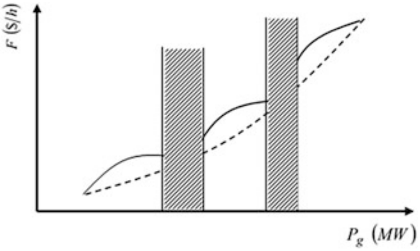
In this paper SVC and TCSC are used to help power system compensate reactive power and reduce power loss. In steady state power system analysis, SVC is modeled as constant VAR source like what depicted in Fig. 2. Based on the variable susceptance of the SVC, it can inject/absorb reactive power in a predefined range shown in (14).

$$QS_w^{\min} \le QS_{w,t} \le QS_w^{\max}, \quad \forall w, t \tag{14}$$

TCSC is simply modeled with a negative reactance. Because of the stability concerns, usually the maximum allowed reactance of the TCSC should be less than 10% of the line reactance. This criteria is modeled as a constraint like in (15).

$$R_z^{\min} \le R_{z,t} \le R_z^{\max}, \quad \forall z, t. \tag{15}$$

Fig. 1 Nonlinear cost function of a generating unit considering prohibited operating zone



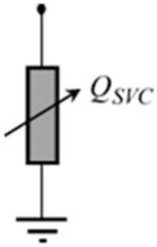


Fig. 2 SVC model used in steady state power system analysis

3 Harmony Search Algorithm

HSA is one of these methods which mimics the process of improvisation of a harmony by musicians which is aesthetically pleasing. Each decision variable in this method is modeled with a musician who plays different notes to find the best one. This best note is the optimal value of the decision variable that generates the best objective function value. A review on steps of standard HSA is as follows.

3.1 Parameters Initialization

To start the process, parameters of HSA should be defined. Main parameters of this algorithm are Harmony Memory Size (HMS), Harmony Memory Consideration Rate (HMCR), Pitch Adjustment Rate (PAR), Band Width (BW) and maximum number of iterations.

3.2 Calculating Initial Objective Function Value of Each Harmony

Initial harmony memory is filled with random notes within their lower and upper bounds. Defined structure of the harmony memory is as (16).

$$HM = \begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_{N-1}^1 & x_N^1 \\ x_1^2 & x_2^2 & \cdots & x_{N-1}^2 & x_N^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1^{HMS-1} & x_2^{HMS-1} & \cdots & x_{N-1}^{HMS-1} & x_N^{HMS-1} \\ x_1^{HMS} & x_2^{HMS} & \cdots & x_{N-1}^{HMS} & x_N^{HMS} \end{bmatrix}. \tag{16}$$

In (16) *HM* is the harmony memory.

Each row in this matric represents a harmony. Each column represents a decision variable. After generating harmony memory, objective function value (fitness value) of each harmony is calculated.

3.3 Improvising a New Harmony

Most important section of each meta-heuristic optimization method is to generate new answers. Let call the new harmony X′ = [x_1', x_2',…, x_N']. With a probability of HMCR, the new note is selected from one of the notes in its corresponding column of harmony memory. It means that $x_i' \in \{x_i^1, x_i^2, \ldots, x_i^{HMS}\}^T$ in order to escape from local optimums, new note may be selected randomly (not from harmony memory) within its lower and upper bounds with a probability of 1-HMCR.

$$x_i' = \begin{cases} x_i' \in \left\{x_i^1, x_i^2, \ldots, x_i^{HMS}\right\} & with\,probability\,HMCR \\ x_i' \in X_i & with\,probability\,(1 - HMCR) \end{cases}. \tag{17}$$

HMCR is usually selected in the range of (0.95–0.99) to prevent premature convergence [30].

If a note is selected from harmony memory, it is imposed to another operator called pitch adjustment. With the probability of PAR, the selected note would experience a small change in a predefined bandwidth around its current value. No changes will be made to the selected note with the probability of 1-PAR.

$$x_i' = \begin{cases} x_i' \pm rand.BW & with\,probability\,PAR \\ x_i' & with\,probability(1 - PAR) \end{cases}. \tag{18}$$

3.4 Updating Harmony Memory

After generating new set of decision variables (improvising a new harmony), corresponding fitness function evaluation is done. If the fitness function of the new harmony is better than the worst fitness function of a harmony in the harmony memory, the new harmony will replace the worst harmony. The worst harmony in a minimization problem is the one with higher fitness value.

3.5 Checking Stop Condition

At each iteration, steps 3–5 are repeated until the number of the iterations reaches the maximum number of iterations. If the maximum number of iterations is reached, the best harmony in the harmony memory is selected as the final optimum point.

3.6 Proposed Improved HSA

There are two main factors that have great importance in optimization algorithms. One is exploration ability and other is local search ability. Two contributions are introduced in this paper to improve these characteristics of HSA. Details of these two contributions are discussed in the following subsections.

3.6.1 Roulette Wheel

The probability of selecting any part of this wheel is proportional to its covered area. Because of the stochastic characteristic of the Roulette Wheel, it can help the optimization algorithms to result in better answers.

To implement this concept to the process of note selection of HSA, the area covered by each harmony in the harmony memory should be proportional to its fitness value. Harmonies with better fitness value should occupy more area of the Roulette Wheel than the ones with worse fitness value and hence have more probability of being selected by new harmony. This is graphically depicted in Fig. 3.

In a minimization problem, the probability of selecting a harmony is proportional to the inverse of its fitness value. This probability is mathematically calculated by (19).

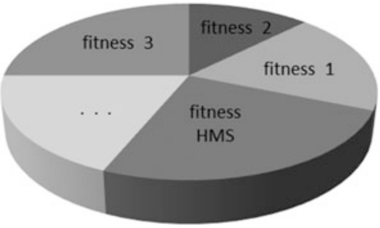
$$p(X_i) = \frac{\frac{1}{fitness_i}}{\sum_{j=1}^{HMS} \frac{1}{fitness_j}}. \tag{19}$$

In (19) $fitness_i$ is the fitness value of harmony i, and $p(X_i)$ is the probability of selecting this harmony.

3.6.2 Parallel Search in Dynamic Subgroups

In many other optimization algorithms like genetic algorithm or PSO, all or some parts of the whole population are changed. But in HSA, just one new harmony is improvised in each iteration. In this paper, the initial harmony memory size is

Fig. 3 Roulette wheel used in HSA



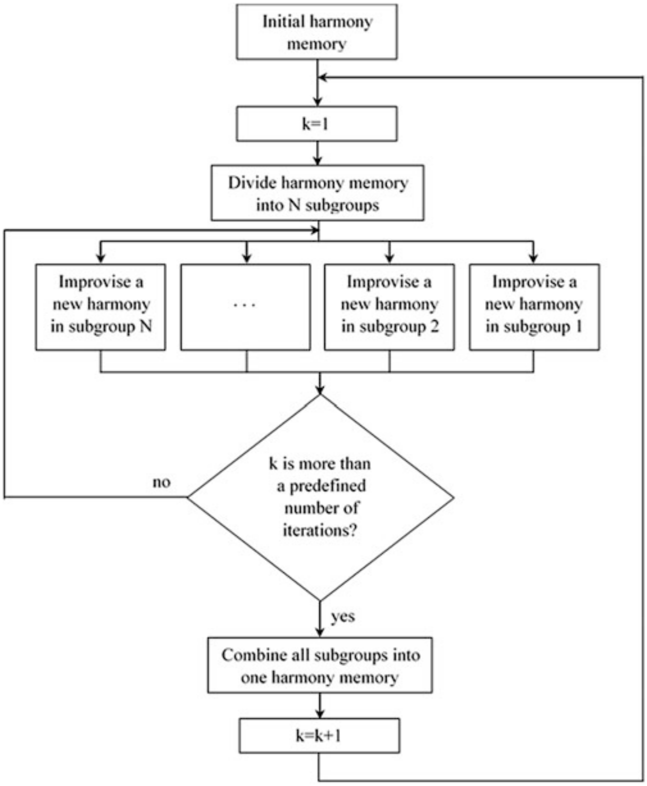

Initial harmony memory
k=1
Divide harmony memory into N subgroups
Improvise a new harmony in subgroup N
. . .
Improvise a new harmony in subgroup 2
Improvise a new harmony in subgroup 1
k is more than a predefined number of iterations?
no
yes
Combine all subgroups into one harmony memory
k=k+1

Fig. 4 Flowchart of the parallel search in dynamic subgroups process

supposed to have many harmonies (big HMS). Then this harmony memory is randomly divided into several subgroups. Each of these subgroups generate a new harmony in each iteration. This procedure takes place in parallel. Then to prevent algorithm to get trapped in local optimums, after a predefined number of iterations, all the harmonies in all subgroups are placed in the initial harmony memory and divided randomly into subgroups again. In this way, members of each subgroup change dynamically and each harmony can use the good data achieved by harmonies from other subgroups. This concept is represented in Fig. 4.

3.7 Handling Constraints in DELD in the Proposed Method

One main item in solving optimization problems using meta-heuristic algorithms is the way constraints are handled. In this paper an improved version of handling constraints using penalty factors called adoptive penalty factor is used to cope with constraints. In this method, two kind of penalty factors are added to the objective function; one based on the amount of deviation from constraints and other based on the number of these unsatisfied constraints.

$$\min f(x) + PF_1 . \sum_{i \in \Omega_h} \Delta h_i(x) + PF_2 . N_h . \tag{20}$$

In which PF_1 and PF_2 are predefined penalty factors, Δh_i is the amount of deviation of constraint i, and N_h is the total number of unsatisfied constraints.

There are two unique kinds of constraints used in DELD problem. The first one is the constraint corresponding to the prohibited operating zones which divide the search space of the problem into several discrete parts. In this paper a new approach is proposed to handle these constraints using a second order penalty function which make the search space continuous.

Considering the prohibited operating zones depicted in Fig. 1, the penalty function for qth zone of decision variable x is defined as in (21).

$$PF_q = -(x - a)^2 + b. \tag{21}$$

To make the search space continuous, parameters a and b are chosen as:

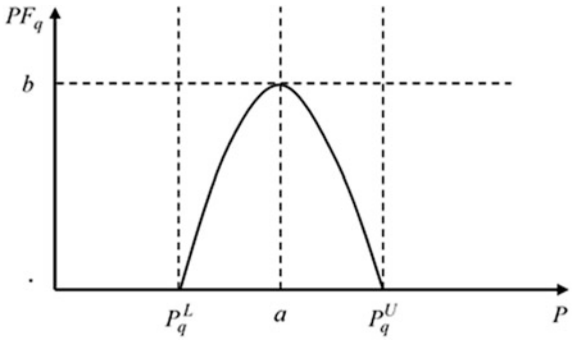
$$a = \frac{P_q^L + P_q^U}{2}. \tag{22}$$

$$b = \left(\frac{P_q^U - P_q^L}{2} \right)^2. \tag{23}$$

The second order penalty function defined in this paper is depicted in Fig. 5.

Other unique kind of constraints used in DELD problem is ramp rate constraint of the generating units. One specific characteristic of these constraints is their vertical (hierarchical) nature. This means that the permissible range of generating active power of a unit at time t is dependent on the amount of active power generated in time t − 1. To handle these constraints, a linearly decreasing penalty

Fig. 5 Penalty function defined for qth prohibited operating zone



factor along the study time horizon is defined. Thus the penalty factor for ramp rate constraint at time t is less than that of time t − 1. Interdependency of ramp rate constraints is relaxed in this method.

Mathematical expression for this penalty factor strategy is described in (24).

$$PF_t = \left(\frac{PF_T - PF_1}{T}\right) \times t + PF_1. \tag{24}$$

In (24) PF_t is the penalty factor defined for time t.

It should be mentioned that in both these methods, a static penalty factor proportional to the number of unsatisfied constraints is also added to the objective function.

4 Simulation Results

In this section, the proposed HSA is used to find the optimal decision variables of DELD problem. As in [32], IEEE 30 bus test system is used to validate the results. A span of 24 h load data is adopted from [32] to solve and compare the results.

Because of the weak VAR areas at the middle of the system, two SVCs are added to the system at buses 17 and 21 each with a capacity of 40 KVAR. Two TCSCs are also considered in this system. The first one is installed between buses 12 and 14 and has a reactance of 0.1 per unit. The other one is placed in line number 38 (between buses 27 and 30) and has a maximum reactance of 0.2 per unit.

Three different scenarios are considered in this chapter. To show the effectiveness and exactness of the proposed algorithm, second and third scenarios are also solved using HSA (HSA), improved HSA (IHSA), global best HSA (GBHSA) and adoptive global best HSA (AGBHSA). Results are compared in terms of the best answer and the convergence speed.

4.1 First Scenario

In this scenario economic load dispatch problem is solved using the proposed HSA. Economic load dispatch is independent of time and is solve for just one hour. So the ramp rate constraints are not considered. In order to demonstrate the effectiveness of the proposed algorithm, results are compared with the ones reported in other papers. To that purpose, all the conditions are considered the same as in [32]. This includes the objective function and constraints. So prohibited operating zones and valve point effect are not considered and there is no FACTS in the system. The comparison is demonstrated in Table 1.

As shown in Table 1, the proposed algorithm in this paper has found better optimal answer compared to the methods reported in other state of the art papers.

Table 1 Optimal results reported in the papers [17]

Optimization method	Cost ($)
Nonlinear programming	802.4
Evolutionary programming	802.62
Genetic algorithm	805.94
Simulated annealing	804.1072
Ant colony optimization	802.578
Improved evolutionary programming	802.465
Enhanced genetic algorithm	802.06
Honey bee mating optimization	802.211
Improved honey bee mating optimization	801.985
Proposed method	800.1146

4.2 Second Scenario

In this scenario, DELD problem is solved and valve point effect is also considered which makes the objective function non-convex. Moreover, prohibited operating zones are also added to the problem to make the search space discrete. Table 2 shows the optimal results found by considered optimization algorithms and the convergence speed diagram is also depicted in Fig. 6.

It is found from Table 2 that the optimal answer found by the proposed algorithm is way better than the ones found by other algorithms. The main reason for such a big difference is the discrete search space of the problem caused by prohibited operating zones. These zones make optimization algorithms trap in local minimums. But the proposed algorithm uses subgroup searching which is a powerful tool to look in all parts of the search space. Another important result is that by adding extra constraint (i.e. prohibited operating zones constraints) to the main problem and taking valve point effect into account, operating cost of the system increases rapidly.

4.3 Third Scenario

As the title of the paper suggests, the main contribution of this paper is to analyze the effect of optimally setting the parameters of FACTS in operating cost of the

Table 2 Optimal results found by different HSA variants in the second scenario

Optimization method	Cost ($)
HSA	27,859
Global best HAS	19,891
Adoptive global best HSA	25,802
Improved HSA	19,972
Proposed method	17,559

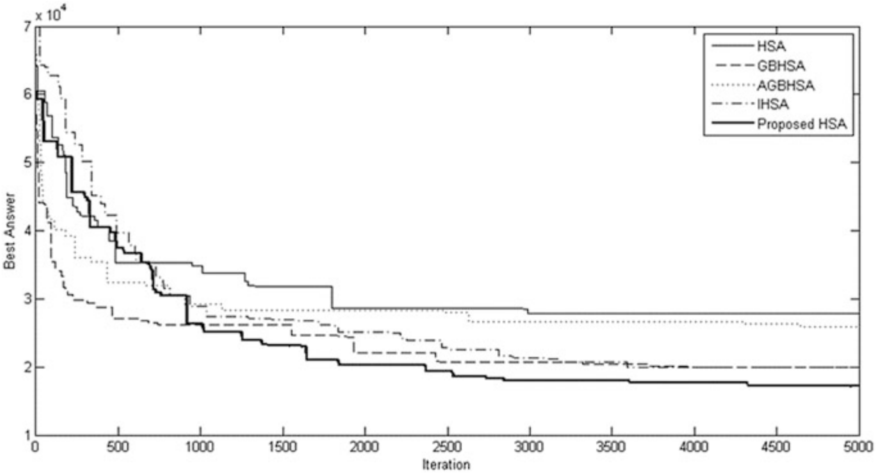


Fig. 6 Convergence diagram of different HSA variants in second scenario

Table 3 Optimal results found by different HSA variants in the third scenario

Optimization method	Cost ($)
HSA	27,151
Global best HAS	20,675
Adoptive global best HSA	26,646
Improved HSA	21,216
Proposed method	17,067

system. Apart from the conditions mentioned in the second scenario, in the third scenario SVC and TCSC devices are also considered in the system. The optimal capacitance of SVCs and reactance of TCSCs are to be found bye solving DELD problem. Optimal results are mentioned in Table 3.

Comparing Tables 2 and 3 may rise such a question: why do some results found by optimization algorithms are better when there is no FACTS in the system? The answer lies in the fact that when we add four FACTS to the system, in fact we have added four more decision variables along with eight inequality constraints to the DELD problem. These make the problem bigger in size and the probability of optimization algorithms get trapped in local minimums increases. But the proposed algorithm has found better optimal answer compared to the on found in the second scenario. This is because FACTS have decreased the operating cost of the system and the algorithm also did not get trapped in local minimums because of its ability to search in parallel subgroups. This result demonstrates the importance of using effective optimization algorithms when dealing with such big problems. The convergence diagram of these algorithms is depicted in Fig. 7.

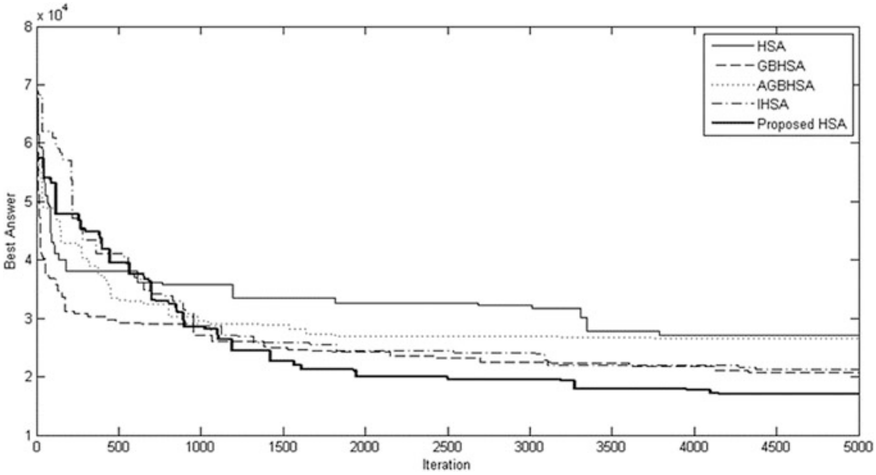


Fig. 7 Convergence diagram of different HSA variants in second scenario

5 Conclusion

As the economic aspects of the energy becomes more important, it is more essential to operate power systems in an optimal manner. DELD is one of the operational optimization problems of power system. Because of the high number of decision variables of this optimization problem and the nonlinear and non-convex structure of the corresponding objective function, it is essential to use a proper optimization algorithm to find the optimal set of decision variables that minimize total operating cost of the system. One of these decision variables is FACTS devices. If optimally set, FACTS could decrease power loss, regulate voltage and decrease operating cost of the system. In this paper, DELD is solve considering FACTS devices like SVC and TCSC.

To optimally set the operating parameters of the system a new version of HSA is proposed in this paper. To improve the performance of this algorithm, two novelties are introduced in this paper. These contributions are the use of Roulette Wheel concept to improvise new harmony and parallel search in subgroups to help the algorithm search more parts of the search space and thus find better optimums. Moreover, a new method is proposed to handle ramp rate constraints and prohibited operating zone constraints.

Finally, the proposed algorithm is tested on IEEE 30 bus test system. Results were compared to the ones reported in the literature and other versions of HSA. It was concluded that the proposed method has better performance and can find better optimums.

References

1. Abu-Mouti FS (2011) Optimal economic and environmental operation of electric power systems via modern meta-heuristic optimization algorithms. PhD Thesis, Dalhousie University
2. Xia X, Elaiw AM (2010) Optimal dynamic economic dispatch of generation: a review. Electr Power Syst Res 80(8):975–986
3. Taranto GN, Pinto LMVG, Pereira MVF (1992) Representation of FACTS devices in power system economic dispatch. IEEE Trans Power Syst 7(2):192–201
4. Sadegh MO, Lo KL (2005) Decentralised coordination of FACTS devices for power system stability enhancement using intelligent programming. COMPEL – Int J Comput Math Electr Electron Eng 24(1):46–60
5. Ekwue AO (1996) Decoupled secure economic dispatch algorithm for power systems operation. COMPEL – Int J Comput Math Electr Electron Eng 5(2):132–142
6. Bechert TE, Kwanty HG (1983) On the optimal dynamic dispatch of real power. IEEE Trans Power Appar Syst 91(3):889–898
7. Van den Bosch PPJ (1985) Optimal dynamic dispatch owing to spinning-reserve and power-rate limits. IEEE Trans Power Appar Syst 104(12):3395–3401
8. Barcelo WR, Rastgoufard P (1997) Dynamic economic dispatch using the extended security constrained economic dispatch algorithm. IEEE Trans Power Syst 12(1):120–128
9. Song YH, Yu I (1997) Dynamic load dispatch with voltage security and environmental constraints. Electr Power Syst Res 43(1):53–60
10. Mahdad B, Srairi K, Bouktir T (2010) Optimal power flow for large-scale power system with shunt FACTS using efficient parallel GA. Int J Electr Power Energy Syst 32(5):507–517
11. Basu M (2011) Multi-objective optimal power flow with FACTS devices. Energy Convers Manag 52(2):903–910
12. Edward JB, Rajasekar N, Sathiyasekar K, Senthilnathan N, Sarjila R (2013) An enhanced bacterial foraging algorithm approach for optimal power flow problem including FACTS devices considering system loadability. ISA Trans 52(5):622–628
13. Ongsakul W, Tippayachai J (2002) Parallel micro Genetic algorithm based on merit order loading solutions for constrained dynamic economic dispatch. Electr Power Syst Res 61 (2):77–88
14. Zhao B, Guo C, Cao Y (2004) Dynamic economic dispatch in electricity market using Particle Swarm Optimization Algorithm. In: 5th World Congress on intelligent control and automation, Vol 6, pp 5050–5054
15. Chen G, Liu L, Guo Y, Huang S (2016) Multi-objective enhanced PSO algorithm for optimizing power losses and voltage deviation in power systems. COMPEL – Int J Comput Math Electr Electron Eng 35(1):471–483
16. Jayakumar N, Subramanian S, Ganesan S, Elanchezhian EB (2015) Combined heat and power dispatch by grey wolf optimization. Int J Energy Sect Manage 9(4)
17. Niknam T, Abarghooee RA, Narimani MR (2012) Reserve constrained dynamic optimal power flow subject to valve-point effects, prohibited zones and multi-fuel constraints. Energy 47(1):451–464
18. Ivatloo BM, Rabiee A, Soroudi A, Ehsan M (2012) Imperialist competitive algorithm for solving non-convex dynamic economic power dispatch. Energy 44(1):228–240
19. Vaisakh K, Praveena P, Rao SRM, Meah K (2012) Solving dynamic economic dispatch problem with security constraints using bacterial foraging PSO-DE algorithm. Int J Electr Power Energy Syst 39(1):56–67
20. Pothiya S, Ngamroo I, Kongprawechnon W (2008) Application of multiple tabu search algorithm to solve dynamic economic dispatch considering generator constraints. Energy Convers Manage 49(4):506–516
21. Panigrahi CK, Chattopadhyay PK, Chakrabarti RN, Basu M (2006) Simulated annealing technique for dynamic economic dispatch. Electr Power Compon Syst 34(5):577–587

22. Niknam T, Golestaneh M (2013) Enhanced bee swarm optimization algorithm for dynamic economic dispatch. IEEE Syst J 7(4):412–423
23. Khorsandi A, Hosseinian SH, Ghazanfari A (2013) Modified artificial bee colony algorithm based on fuzzy multi-objective technique for optimal power flow problem. Electr Power Syst Res 95(1):206–213
24. Geem ZW, Kim JH, Loganathan GV (2001) A new heuristic optimization algorithm: harmony search. J Simul 76(2):60–68
25. Mahdavi M, Fesanghary M, Damangir E (2007) An improved HSA for solving optimization problems. Appl Math Comput 188(2):1567–1579
26. Chen J, Pan QK, Li JQ (2012) HSA with dynamic control parameters. Appl Math Comput 219(2):592–604
27. Pan QK, Suganthan PN, Tasgetiren MF, Liang JJ (2010) A self-adaptive global best HSA for continuous optimization problems. Appl Math Comput 216(3):830–848
28. Wang X, Yan X (2013) Global best HSA with control parameters co-evolution based on PSO and its application to constrained optimal problems. Appl Math Comput 29(19):1059–1072
29. Pandi VR, Panigrahi BK (2011) Dynamic economic load dispatch using hybrid swarm intelligence based HSA. Expert Syst Appl 38(7):8509–8514
30. Ahangaran M, Ramezani P (2013) HSA: strengths and weaknesses. J Eng Inf Technol 1(2):1–7
31. Geem ZW (2012) Effects of initial memory and identical harmony in global optimization using HSA. Appl Math Comput 28(22):1337–1343
32. Pothiya S, Ngamroo I, Kongprawechnon W (2008) Application of multiple tabu search algorithm to solve dynamic economic dispatch considering generator constraint. Energy Convers Manag 49(4):506–516

Coordinated Operation of Wind Farm, Pumped-Storage Power Stations, and Combined Heat and Power Considering Uncertainties

Hamid Jafari, Ehsan Jafari and Reza Sharifian

Abstract One approach to increase the economic efficiency of renewable power generation units is to use combined projects. The supplied energy from these renewable resources is unpredictable because the wind speed and the amount of production in wind farms is not certain. Thus, to increase the reliability, maximize the profits, and supply the load demands of these wind farms and storage systems they must be considered much more than the supply amount. By combining two or more resources, in combined systems, with the capability of predicting production, the controllable production increases; in fact, these resources cover each other`s shortcomings. Here the wind farms, pumped storage, and combined heat and power (CHP) are used in a coordinated way to supply the load demands. Determining independent and coordinated operation strategy of power generation units in the IEEE 30-bus standard System shows that the proposed method is efficient and appropriate.

Keywords Wind · Combined heat and power · Pumped-storage
Uncertainty · Electricity market · Coordinated operation

1 Introduction

The ever-increasing development of the communities and industrial development in different countries has increased the demands of energy resources. These resources are limited in size and amount and bring about pollutions for the environment. The

H. Jafari · E. Jafari (✉) · R. Sharifian
Department of Electrical Engineering, Islamic Azad University,
Lenjan Branch, Isfahan, Iran
e-mail: jafari@iauln.ac.ir

H. Jafari
e-mail: hamid.jafary62@yahoo.com

R. Sharifian
e-mail: sharifian@iauln.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_51

mortality of fossil fuels, diversifying the energy recourses, stable development and creating energy safety, the environmental problems resulted from fossil fuels consumption from one hand and renewability and cleanliness of the alternative energies such as sun, wind, biomass, etc. from the other hand, have drawn serious attention to the development of renewable energy and increasing the share of the alternative energy in the energy basket of the world. Nowadays, the activities and budgets of the governments and companies in the research, development, and supplying renewable energy systems are significantly increasing and these activities and spending the mentioned budget finally cause reduction of renewable energy costs and competition with the present conventional energies. This has fulfilled about wind energy and some applications of biomass energy; and the fast process of cost reduction about other renewable energy resources is on operation now. In recent years, politicians and researchers started to pay attention to the renewable energy.

Using the wind energy has many advantages such as:

1. It is free
2. It is clean and less harmful to the environment
3. It is available

However, there are some shortcoming in these power stations such as low reliability, random production of these power stations, and lack of precise predation. In other words, since the output from these kinds of power stations and the degree of unbalance costs are not defined, applying such power stations are risky. One problem about applying such energy is the uncertainty of wind generator productions. Therefore, there have been numerous studies on the fluctuation effect of wind farms power on the power systems. Edward et al. [1] reviewed some previous works on planning wind farms and pumping-storage in the electricity market. This could be one of the best references. Bruno et al. [2] analyzed the reliability of power systems including huge water-storage stations in the electricity market. Zhang et al. [3] studies the harmony between wind farms and pumped-storage in the electric energy market so that the optimizing problem formulized as a random program but ignored the reliability of the stations.

Amjady and Vahidinasab [4] studied the optimized utilization of wind farm in the pumped-storage power stations so that the worst production scenario was wind unit. In [4], reliability was not taken into accounts as well. Conejo et al. [5] modeled a wind farm to study the reliability evaluation. Conejo et al. [5] used sequential and multiphase Monto Carlo simulation method to display wind farm. Wei et al. [6] examined the reliability of wind farm with accordance to pumped-storage power station in the power system, to create harmony between wind farm and pumped-storage power station the Monto Carlo simulation method was applied. The target was to calculate the sufficient indexes for the period of one year.

In [2–8] the optimization methods and methods to improve the reliability in a fundamental manner were elaborated.

The present study examined the 30 Bus IEEE power system in which the energy is supplied through wind production, pumped-storage and CHP units. The main objectives are a harmonic planning and optimizing these three units for one day (24 h) enjoying PSO Algorithm.

2 The Electricity Market

The present study assumes that the electricity market is single rate and the structure of local limit prices are taken into account for pricing. Producers and consumers presents their production and consumption proposals for each hour to the day-ahead energy market, respectively [9–20]. After receiving presented proposal to the market, the market beneficiary performs the market balance algorithm so that the production proposal and accepted or non-accepted consumption are defined in the market and limit price of the system are defined per hour. The process of market balance algorithm is in a way that the winner producer in the market receives the costs in the extent of accepted production, per hour, in the market multiplied by limit price system in that hour; if the producer could not produce as much as hour production proposal in the accepted market he must pay the cost of non-balance. The cost of considered non-balance in the present study is a percentage of lime price of the system multiplied by absolute value of the created unbalanced extent (the difference between the real production and pre-accepted value in the market in that hour) by that producer. Of course, any other algorithm is applicable to the random programming model in the present study.

3 Wind Power Uncertainty and Analysis Modeling

Time sequential methods are the simplest and the most inexpensive methods for predicting wind speed. Thus, in most references most methods are used in the wind speed modeling. In the following, the ARMA time sequential method which has been applied in most articles for wind modeling is used. This algorithm is simple in the concept so that the wind speed per hour depends on wind speed in the previous hour. In this method, producing wind speed is done through a random number (because the climates are random). This method predicts the wind speed by statistical concepts and is known as ARMA. The wind speed model in each scenario and hour can be calculated through $SW_t = \mu_t + \sigma_t y_t$ relation [21]. Since the power-speed property output of these turbine is nonlinear, the nonlinear math functions could be used to describe them, so that the coefficients of these two relations depends on turbine properties such as minimum permitted speed, rated speed, and maximum permitted speed [6, 11].

The target function as the maximizing the profit in the period of one day is based on relation.

$$\begin{aligned}
\mathrm{R_W} = \sum_{t=1}^{24} \{ & (\mathrm{MCP_t(ti)\,.\,P_{Wb,t}(Bi, Wi, ti)}) \\
& + (1 - \mathrm{bi,\ t(Bi,\ Wi,\ Si,\ ti))\,.\,MCP_t^{up}(ti)\,.} \sum_{\mathrm{Si}} (\mathrm{P_{Wi,t}(Bi, Wi, Si, ti)} \\
& - \mathrm{P_{Wb,t}(Bi, Wi, ti))\,.\,P_i(Bi, Wi, Si, ti))} \\
& - \mathrm{b_{i,t}(Bi, Wi, Si, ti)\,.\,MCP_t^{down}(ti)\,.} \sum_{\mathrm{Si}} ((\mathrm{P_{Wb,t}(Bi, Wi, ti)} \\
& - \mathrm{P_{Wi,t}(Bi, Wi, Si, ti))\,.\,P_i(Bi, Wi, Si, ti)}
\end{aligned} \tag{1}$$

4 Pumped-Storage Power Stations Analysis Modeling

The purpose of this section is the analytical proposing modeling of pumped-storage power station in the common market of day-ahead energy. In the present modeling, the programming period for the pumped-storage power station is for one day. In fact, the stored energy in the higher tank of the pumped-storage power station at the late hours of the day equals to stored energy in the early hours of the same day. The target function as the maximizing the profit in the period of one day is based on relation [14].

$$\begin{aligned}
\mathrm{R_{PS}} = & \sum_{t=1}^{24} \mathrm{MCP_t(ti)\,.\,P_{gp,t}(Bi, PSHPPi, ti)} \\
& - \sum_{t=1}^{24} \mathrm{C(Bi, PSHPPi)\,.\,(P_{gp,t}(Bi, PSHPPi, ti) + P_{pP,t}(Bi, PSHPPi, ti))} \\
& - \sum_{t=1}^{24} \mathrm{MCP_t(ti)\,.\,P_{pP,t}(Bi, PSHPPi, ti)}
\end{aligned} \tag{2}$$

$$\begin{aligned}
\mathrm{E_{u,t}(Bi, PSHPPi, ti)} = & \ \mathrm{E_{u,t-1}(Bi, PSHPPi, ti)} \\
& + \eta\mathrm{(Bi, PSHPPi)\,.\,P_{pP,t}(Bi, PSHPPi, ti)} \\
& - \mathrm{P_{gP,t}(Bi, PSHPPi, ti)}
\end{aligned} \tag{3}$$

$$\begin{aligned}
\mathrm{E_u^{min}(Bi, PSHPPi)} \le & \ \mathrm{E_{u,t}(Bi, PSHPPi, ti)} \\
& \le \mathrm{E_u^{max}(Bi, PSHPPi)}
\end{aligned} \tag{4}$$

$$\begin{aligned}
& \mathrm{P_{pP}^{min}(Bi, PSHPPi)\,.\,n_{p,t}(Bi, PSHPPi, ti)} \le \mathrm{P_{pP,t}(Bi, PSHPPi, ti)} \\
& \quad \le \mathrm{P_{pP}^{max}(Bi, PSHPPi)\,.\,n_{p,t}(Bi, PSHPPi, ti)}
\end{aligned} \tag{5}$$

$$\begin{aligned} & \mathrm{m}_{\mathrm{g,t}}(\mathrm{Bi, PSHPPi, ti})\left(\frac{1}{\mathrm{N}}\right). + \mathrm{n}_{\mathrm{p,t}}(\mathrm{Bi, PSHPPi, ti}) \leq 1 \\ & \left(\mathrm{m}_{\mathrm{gP,t-1}}(\mathrm{Bi, PSHPPi, ti} - 1) + \left(\frac{1}{\mathrm{N}}\right). \mathrm{n}_{\mathrm{p,t}}(\mathrm{Bi, PSHPPi, ti})\right) \leq 1 \end{aligned} \tag{6}$$

$$\left(\mathrm{m}_{\mathrm{gP,t}}(\mathrm{Bi, PSHPPi, ti}) + \left(\frac{1}{\mathrm{N}}\right).\mathrm{n}_{\mathrm{p,t-1}}(\mathrm{Bi, PSHPPi, ti} - 1)\right) \leq 1 \tag{7}$$

$$\mathrm{E}_{\mathrm{u}}^{0}(\mathrm{Bi, PSHPPi, ti}) = \mathrm{E}_{\mathrm{u}}^{\mathrm{end}}(\mathrm{Bi, PSHPPi, ti}) \tag{8}$$

$$\begin{aligned} & \mathrm{P}_{\mathrm{gP,t}}(\mathrm{Bi, PSHPPi, ti}) - \mathrm{P}_{\mathrm{pP,t}}(\mathrm{Bi, PSHPPi, ti}) \\ & \quad = \mathrm{P}_{\mathrm{Pb,t}}(\mathrm{Bi, PSHPPi, ti}) \end{aligned} \tag{9}$$

$$\begin{aligned} & \mathrm{P}_{\mathrm{gP}}^{\mathrm{min}}(\mathrm{Bi, PSHPPi}).\, \mathrm{m}_{\mathrm{g,t}}(\mathrm{Bi, PSHPPi, ti}) \leq \mathrm{P}_{\mathrm{gP,t}}(\mathrm{Bi, PSHPPi, ti}) \\ & \quad \leq \mathrm{P}_{\mathrm{gP}}^{\mathrm{max}}(\mathrm{Bi, PSHPPi}).\, \mathrm{N}.\mathrm{m}_{\mathrm{g,t}}(\mathrm{Bi, PSHPPi, ti}) \end{aligned} \tag{10}$$

N is the number of similar units in the pumped-storage power station,

$\mathrm{P}_{\mathrm{gp,t}}(\mathrm{Bi, PSHPPi, ti})$ is the amount of hour-production of pumped-storage power station,

$\mathrm{P}_{\mathrm{pP,t}}(\mathrm{Bi, PSHPPi, ti})$ is the amount of proposed hour energy of the pumped-storage power station based on daily market,

$\mathrm{P}_{\mathrm{Pb,t}}(\mathrm{Bi, PSHPPi, ti})$ The amount of hourly energy supply offered by pumped storage plant to the daily market.

$\mathrm{E}_{\mathrm{u,t}}(\mathrm{Bi, PSHPPi, ti})$ is the stored energy in the in the higher tank of pumped-storage power station in hour.

$\mathrm{m}_{\mathrm{g,t}}(\mathrm{Bi, PSHPPi, ti})$ is the binary variable that its being 1 shows that the manner of power station production is in *ti* hour.

$\mathrm{n}_{\mathrm{p,t}}(\mathrm{Bi, PSHPPi, ti})$ is the correct variable displaying the number of units that are in hour of consumption state (pumping).
The range of these variable is between 0 and N. The $\eta(\mathrm{Bi, PSHPPi})$ is the pumped-storage power station efficiency.

$\mathrm{E}_{\mathrm{u}}^{\mathrm{max}}(\mathrm{Bi, PSHPPi})$ is the maximum storable energy in the higher tank of pumped-storage power station,

$\mathrm{E_u^{min}(Bi, PSHPPi)}$ is the minimum stored energy in the higher tank of pumped-storage power station,

$\mathrm{E_u^0(Bi, PSHPPi, ti)}$ is the degree of stored energy in the higher tank of pumped-storage power station in the early hours of the day,

$\mathrm{E_u^{end}(Bi, PSHPPi, ti)}$ is the degree of stored energy in the higher tank of pumped-storage power station at the late hours of the day.

$\mathrm{P_{gP}^{max}(Bi, PSHPPi)}$ is the maximum production power of pumped-storage power station,

$\mathrm{P_{gP}^{min}(Bi, PSHPPi)}$ is the minimum production power of pumped-storage power station,

.$\mathrm{P_{PP}^{max}(Bi, PSHPPi)}$ is the maximum consumption power of pumped-storage power station,

$\mathrm{P_{PP}^{min}(Bi, PSHPPi)}$ is the minimum power of pumped-storage power station,

C is the cost of pumped-storage power station operation,

RP is the profit of pumped-storage power station by participating in daily market.

Relation (3) shows the expected stored energy in the higher tank of pumped-storage power station per hour and Relation (4) shows the permitted energy range in this tank. Also, in practice the lower tank of the pumped-storage power station is bigger than the higher one so that mentioning the permitted stored energy range is not taken into account in the lower tank. Relations (5) and (6) show permitted range of energy production in the production hours and the energy consumption range in the consumption hours of pumped-storage power station. Relation (7) grantees that whenever one of the unit of the pumped-storage power station be in pumping position, the production cannot be done. Relations (8) and (9) model the changing time of pumped-storage power station manners so that in a regular base changing the power station manner from production to consumption and vice versa demands passing the time for some minutes; this, in the one-hour-based market causes losing the opportunity of power station cooperation in the market. Relation (10) shows the balance of stored energy in the higher tank [21].

5 Combined Heat and Power Station Modeling

When an industrial process demands huge amount of heat supplied from non-electric resources such as fossil fuels or biomass, using one combined production factory is economical. Advances in the electronics created simplifications of access to security and quality affairs of electric companies. With the advent of power electronic equipment with high reliability, installing combined production equipment even in domestic level has become economical and safe. These installations can produce household hot water, electricity, household heating and sell the extra energy to the electricity company [22].

Since the objective is examining the simultaneous production of electricity and heat, here the energy function CHP power station is introduced.

$$P_{GCHP}(t) - P_{GCHP}(\alpha) - \frac{P_{GCHP}(\alpha) - P_{GCHP}(\beta)}{H_{GCHP}(\alpha) - H_{GCHP}(\beta)}(H_{GCHP}(t) - H_{GCHP}(\alpha)) \leq 0 \quad (11)$$

$$P_{GCHP}(t) - P_{GCHP}(\beta)\frac{P_{GCHP}(\beta) - P_{GCHP}(\theta)}{H_{GCHP}(\beta) - H_{GCHP}(\theta)}(H_{GCHP}(t) - H_{GCHP}(\beta)) \geq -(1 - M(CHP,)) * Y \quad (12)$$

$$P_{GCHP}(t) - P_{GCHP}(\theta)\frac{P_{GCHP}(\theta) - P_{GCHP}(\lambda)}{H_{GCHP}(\theta) - H_{GCHP}(\lambda)}(H_{GCHP}(t) - H_{GCHP}(\theta)) \geq -(1 - M(CHP,)) * Y \quad (13)$$

The target function as the maximizing the profit in the period of one day is based on relation.

$$\mathrm{R_{CHP}} = \sum_{\mathrm{t=1}}^{24}((\mathrm{MCP_t(ti)\,.\,PG_{CHP}(Bi, Gi, ti)}) + \mathrm{SR_{income-CHP}(ti)}) - \mathrm{OC_{uc}(Bi, Gi, ti)} \quad (14)$$

6 Defining the Variables

The present study examines two different structures; the first structure is the manner in which the pumped-storage power station and wind farm and CHP are utilized separately; in the second structure the three power stations operate next to each other as the complementary. In this section, the descriptive math relations of the problem are introduced for each of the mentioned structures. In addition, in the present study the absolute value function is used in the target function relation, because this function is easily able to be linear so that problem solving or linear methods would be possible.

To have a harmonic performance in all three power stations, optimizing profits and reducing the wastes, the cost function in all three power stations must be optimized through Particle Swarm Optimization Algorithm. Thus, the target function by which the optimization is defined is as the following:

$$\mathrm{MAX}\ (\mathrm{Rtotal}\ ) = \mathrm{R_{CHP}} + \mathrm{R_{Ws}} + \mathrm{R_{ps}} \tag{15}$$

The users' demand must be supplied through different available energy producers in the common market. So, in relation (14) we have:

$$\begin{aligned} &\mathrm{PG_{CHP}(ti)} + \mathrm{PG_{Wi,t}(ti)} + \mathrm{PG_{gP,t}(ti)} \\ &= \mathrm{P_{D,t}(ti)} + \mathrm{P_{DpP,t}(ti)} \forall \mathrm{ti} \in \mathrm{T} \end{aligned} \tag{16}$$

7 Particle Swarm Optimization Algorithm

Particle Swarm Optimization (PSO) method is a state minimization method by which some problems whose answer is one point or level in the next n space are solved. In such space, there are some assumptions and one primitive speed is allocated to them. Moreover, the connection channels among the particles are taken into accounts. Then, these particles move in the response space and the results are computed based on a "qualified criterion" after each time interval. After passing the time, the particles accelerate toward the particles with higher qualified criterion and are in the same connection group. Despite working in well in some range of problems this method proved to be very successful in connected optimization problems.

The PSO Algorithm was first introduced in 1998 and in the optimization problems particularly in the electricity-power engineering parameters have had many applications. The PSO Algorithm is a social search algorithm modeled from the birds' social behaviors. At first, this algorithm was used to explore the governing pattern on the simultaneous flight of the birds and their sudden direction change to the optimized position. In PSO, the particles are flowing in the search space. The position change of the particles in the search space is influenced by their experience and knowledge and those of their neighbors. Thus, the other position of the particle mass effect on how a particle searches. The modeling results of this social behavior is a searching process that moves the particles toward successful regions. Particles learn from one another that based on the gained knowledge move toward to their best neighbors. The base principle of PSO is that each particle in every moment arranges its position in the search space according to the best position in its entire neighborhood. This method has been used for optimization of combined cost function and in the following the results is presented. The modeling result of this behavior is a search process in which the particles move towards the successful regions. The simulation was done in MATLAB software and in mfile environment.

8 Simulation Results

According to the description in the previous section, the simulation was done into manners as independent and harmonic for the pumped-storage and CHP and the results are as the following:

Figure 1—Uncoordinated operation of wind farm for supplying power in a day.

Figure 2 shows the Coordinated operation of wind farm for supplying load in a day. Comparing these two figure shows the effect of harmonic performance of the power stations and optimization algorithm (PSO) on reducing the produced power.

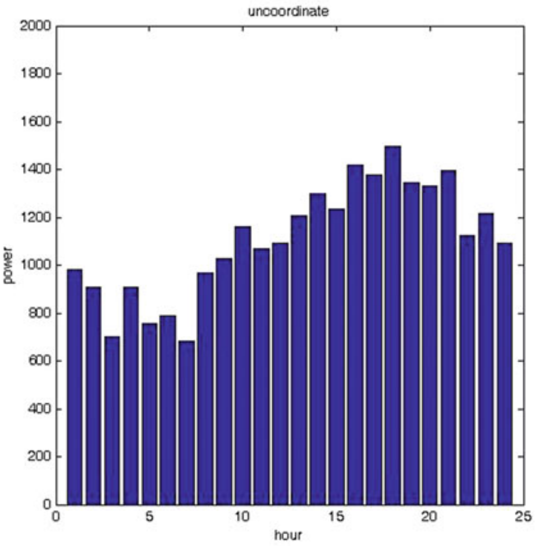
Figures (2, 3 and 4) shows the Coordinated and Uncoordinated operation of pumped-storage power station.

Optimized production of this unit effects Coordinated operation proportionate to Uncoordinated operation conditions. These conditions are observed in power unit diagram of CHP shown in Figs. 5 and 6 as the harmonic and independent performance.

Figure 7 (PSO Algorithm) shows the optimized system by the particle Algorithm and the convergence of this algorithm are shown on Fig. 8.

Table 1 shows the costs and profits in the Uncoordinated and Coordinated operation for the units. According to the results, the costs and profits of the power stations in independent and harmonic performance can be compared.

Fig. 1 Uncoordinated operation of wind farm



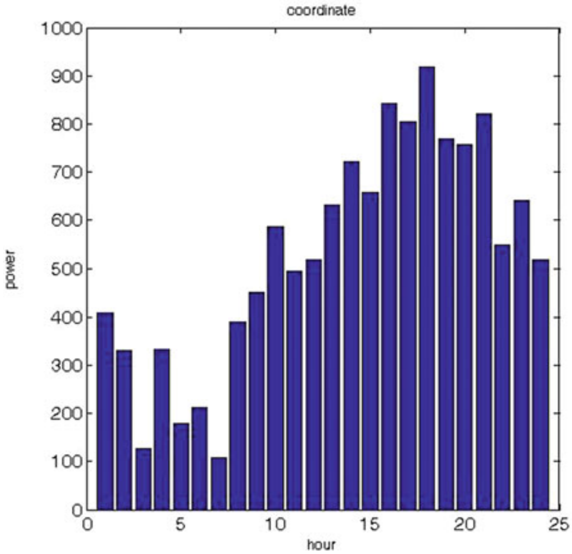


Fig. 2 Coordinated operation of wind farm

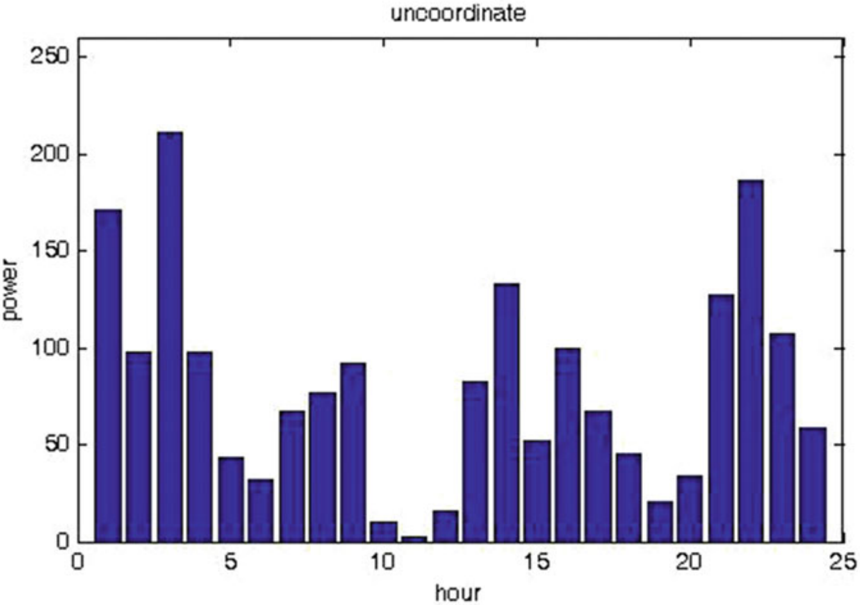


Fig. 3 Uncoordinated operation of pumped-storage power station

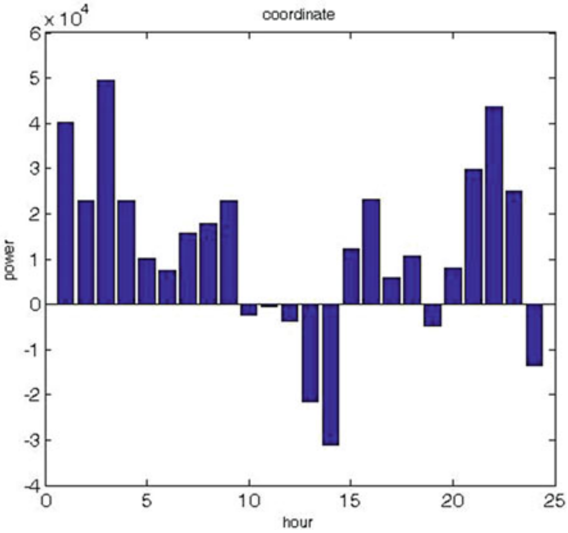


Fig. 4 Coordinated operation of pumped-storage power station

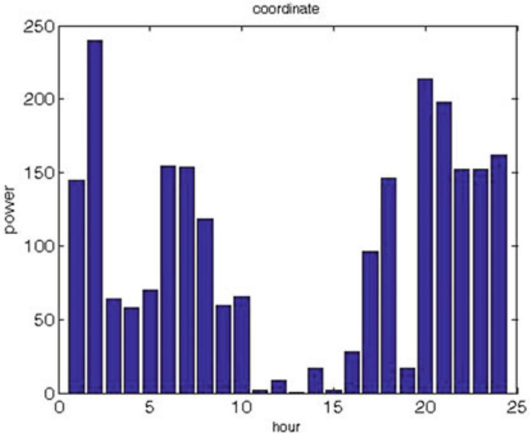


Fig. 5 Coordinated operation of CHP power station

9 Conclusion

The present study examined the Coordinated operation of wind farm, pumped-storage and CHP power stations and the simulation results were compared with of the independent performance of each units. First the independent cost function was presented for each of the power stations and then for their harmonic performance the sum of cost functions of three units were used. The cost function for these three power station units in the harmonic state was through PSO algorithm and the degree of production in each unit for optimizing the costs and increasing the profit was done by PSO algorithm. The simulation results proved that this method was efficient.

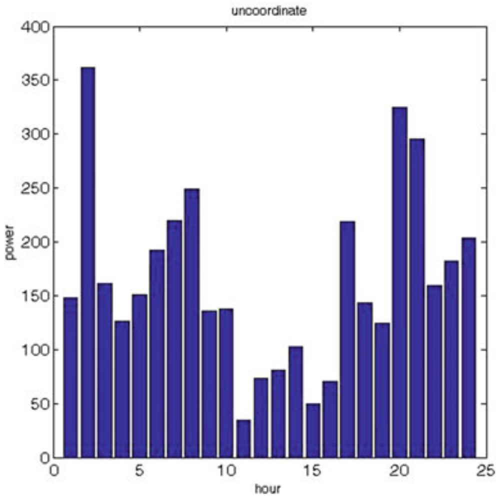


Fig. 6 Uncoordinated operation of CHP power station

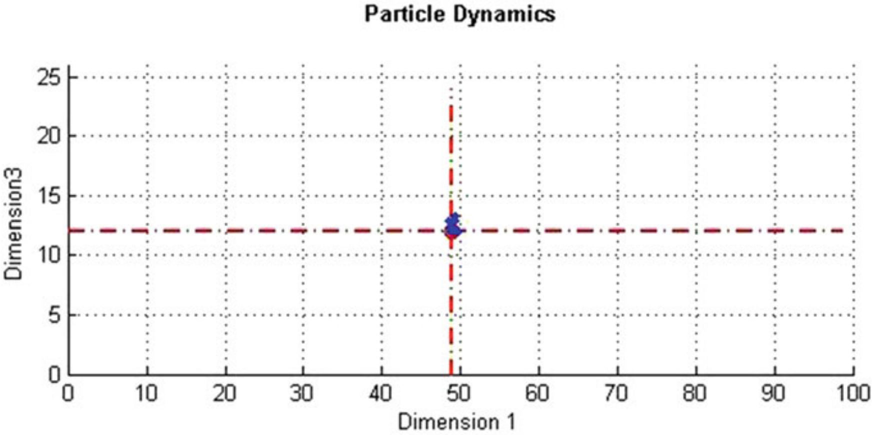


Fig. 7 The optimized cost function by particle algorithm

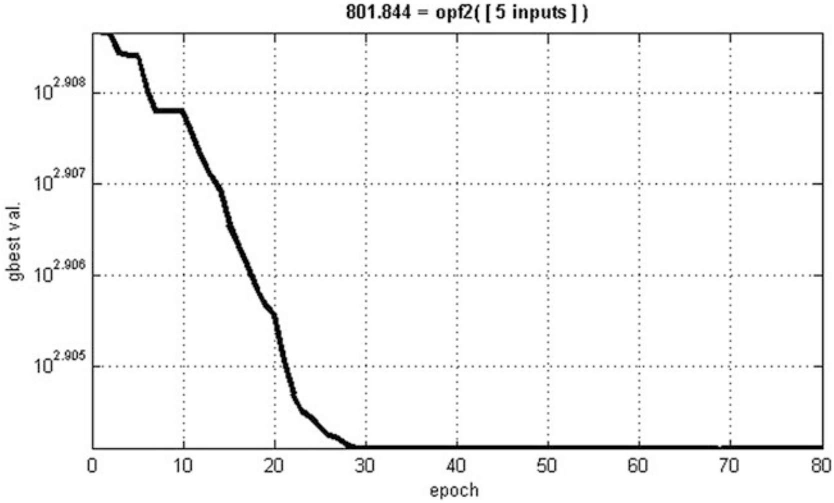


Fig. 8 Converging the PSO Algorithm

Table 1 Cost and profits of the units

Profit ($)	Sell (KW)	Power(W)	
18,910	136,962	68,481	Uncoordinated operation
23.21	148.76	79.32	Coordinated operation

References

1. Edward B et al (2016) A review of pumped hydro energy storage development in significant international electricity markets. Renew Sustain Energy Rev 61:421–432
2. Bruno V et al (2016) A multiple criteria utility-based approach for unit commitment with wind power and pumped storage hydro. Electr Power Syst Res 131:244–254
3. Zhang L et al (2017) A multiobjective robust scheduling optimization mode for multienergy hybrid system integrated by wind power, solar photovoltaic power, and pumped storage power. Math Probl Eng
4. Amjady N, Vahidinasab V (2013) Security-constrained self-scheduling of generation companies in day-ahead electricity markets considering financial risk. Energy Convers Manag 65:164–172
5. Conejo AJ, Carrión M, Morales JM (2010) Decision making under uncertainty in electricity markets. Springer, New York
6. Wei W, Liu F, Wang J, Chen L, Mei S, Yuan T (2016) Robust environmental—economic dispatch incorporating wind power generation and carbon capture plants. Appl Energy 183:674–684
7. Jiang R, Wang J, Guan Y (2012) Robust unit commitment with wind power and pumped storage hydro. IEEE Trans Power Syst 27(2):800–810
8. González C, Juan J, Mira J, Prieto FJ, Sánchez MJ (2005) Reliability analysis for systems with large hydro resources in a deregulated electric power market. IEEE Trans Power Syst 20:90–95

9. Zhao J, Lo KL, Lu J (2016) Variously worldwide types of deregulated electricity markets and their respective transmission congestion management schemes. In: 51st international universities' power engineering conference
10. Alismail F, Xiong P, Singh C (2017) Optimal wind farm allocation in multi-area power systems using distributionally robust optimization approach. IEEE Trans Power Syst
11. González JG, de la Muela RMR, Santos LM, González AM (2008) Stochastic joint optimization of wind generation and pumped-storage units in an electricity market. IEEE Trans Power Syst 23:456–461
12. Hu P (2009) Reliability evaluation of electric power systems including wind power and energy storage. Ph.D. Thesis, Department of Electrical and Computer Engineering, University of Saskatchewan Saskatchewan, Saskatoon
13. Karki R, Hu P, Billinton R (2006) A simplified wind power generation model for reliability evaluation. IEEE Trans Power Syst 20:533–540
14. MacKay DJC (2007) Enhancing electrical supply by pumped storage in tidal lagoons. Cavendish Laboratory, University of Cambridge
15. Matevosyan J, Soder L (2006) Minimization of imbalance cost trading wind power on the short-term power market. IEEE Trans Power Syst 21:1396–1404
16. Shahidehpour M, Yamin H, Li Z (2002) Market operations in electric power systems: forecasting, scheduling, and risk management. Wiley, New York
17. Vahidinasab V, Jadid S (2010) Stochastic multiobjective self scheduling of a power producer in joint energy and reserves markets. Electr Power Syst Res 80:760–769
18. Shi N, Luo Y (2017) Energy storage system sizing based on a reliability assessment of power systems integrated with wind power. Sustainability 9:395
19. Adineh B, Mashhadi HR, Hajiabadi ME (2014) Determining appropriate buses and networks for applying demand side management programs by structural analysis of EENS. Iran J Electr Electron Eng 10
20. Afshar K et al (2007) A new approach for reserve market clearing and cost allocating in a pool model. Iran J Sci Technol 31(B6):593
21. Jiang R, Wang J, Guan Y (2011) Robust unit commitment with wind power and pumped storage hydro. IEEE Trans Power Syst 23:1–11
22. Jafari E, Soleymani S, Mozafari B, Amraee T (2018) Optimal operation of a micro-grid containing energy resources and demand response program. Int J Environ Sci Technol https://doi.org/10.1007/s13762-017-1525-6

Optimization of Exponential Double-Diode Model for Photovoltaic Solar Cells Using GA-PSO Algorithm

Vahdat Nazerian and Sogand Babaei

Abstract In this paper, an equivalent electrical circuit based on the photovoltaic effect (PV) is presented with studies on the simulation of the solar energy system. This model consists of exponential double diodes illustrates how solar cells behave in order to generate electricity. By using the MATLAB software, we performed simulations. Our goal is to calculate the minimum error value for the unknown parameters of the model, which is attained by using root mean square of errors (RMSE). Regarding to the offered model, which we intend to investigate with the suggested GA-PSO algorithm, we obtain the minimum error value (RMSE) after achieving unknown parameters and then we will compare the results with other methods. Therefore, it can be shown that the proposed algorithm with a RMSE value of 2.02 provides an optimal result. According to the computed calculations, the runtime of this algorithm for each calculation is approximately 1 min and 30 s, while the total time of the algorithm will be figured according to the parameter values and the frequency of repetition. With the progress of the calculation process this time comes out to 3 min and 30 s.

Keywords Photovoltaic effect · Solar cells · Exponential Double-diode model
GA-PSO algorithm · Root mean square of errors (RMSE)

V. Nazerian (✉)
Department of Electrical Engineering,
University of Mazandaran, Babolsar, Iran
e-mail: v.nazerian@umz.ac.ir

S. Babaei
Department of Software Engineering,
University of Mazandaran, Babolsar, Iran

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_52

1 Introduction

The sun is the biggest renewable energy source on the earth. If only 1% of the world's deserts are equipped with thermal power plants, that is enough to produce world-wide annual demand for electricity. For every 1 kw of solar energy, the production of 6 kg of pollutant is prevented, which is equivalent to an average of 5270 kg annually, resulting from burning 21,222 L of diesel fuel. The efficiency of these cells has been improved from 6 to 22% over the past 60 years. With the exordium of nanotechnology in the production of solar cells, the efficiency of these cells has increased drastically [1]. There are two ways to use solar energy: direct use of sunlight and convert it to electricity through photovoltaic cells and direct use of solar energy and convert it to other energies (power plant applications).

The economic benefits of using solar energy include no need for fuel, easy to install and connected at any location, low cost utilization, long-term economic savings, free energy source and creation of culture. Recently, photovoltaic arrays have been used in many applications such as battery chargers, solar water pumping systems, network connected PV systems, solar hybrid vehicles, and satellite systems. In all solar energy systems, efficient simulations of photovoltaic panels are required before any testing approvals [2–4].

2 The Electrical Circuit Model and Related Equations

2.1 Exponential Double-Diode Circuit

The equivalent circuit based on the double-diode model of solar cell is shown in Fig. 1.

This circuit model has a light-dependent current source, two exponential diodes, series resistance and parallel resistance [5, 6]. The amount of current source is directly proportional to the light emitted on the photovoltaic cell, which changes as a linear coefficient with light intensity. Due to the semiconductor junction in photovoltaic cells, this circuit also uses two diodes and two resistors for the modeling.

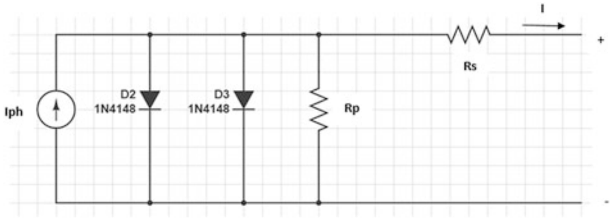


Fig. 1 Double-diode circuit model

2.2 Mathematical Equations

According to the model presented in Fig. 1, the current of the first and second diodes, parallel resistance, photovoltaic current source and the output current of the array are given by the Eqs. (1)–(5) below, respectively.

$$\mathrm{I_{D1}} = \mathrm{I_{S1}} * \left[\left(e^{\frac{\frac{V}{N_S}+R_S*I}{A_1*V_T}}\right) - 1\right] \tag{1}$$

$$\mathrm{I_{D2}} = \mathrm{I_{S2}} * \left[\left(e^{\frac{\frac{V}{N_S}+R_S*I}{A_2*V_T}}\right) - 1\right] \tag{2}$$

$$\mathrm{I_{sh}} = \frac{\frac{V}{N_S} + R_S * I}{R_P} \tag{3}$$

$$\mathrm{I_{ph}} = \mathrm{I_{D1}} + \mathrm{I_{D2}} + \mathrm{I_p} + \mathrm{I} \tag{4}$$

$$\mathrm{I} + \mathrm{I_{S1}}* \left[\left(e^{\frac{\frac{V}{N_S}+R_S*I}{A_1*V_T}}\right) - 1\right] + \mathrm{I_{S2}}* \left[\left(e^{\frac{\frac{V}{N_S}+R_S*I}{A_2*V_T}}\right) - 1\right] + \frac{\frac{V}{N_S} + R_S * \mathrm{I}}{R_P} - \mathrm{I_{ph}} = 0 \tag{5}$$

where $\mathrm{I_D}$ is diode current, $\mathrm{I_S}$ is reverse saturation current, V is output voltage of solar array, $\mathrm{V_T}$ is thermal voltage, A is ideality factor of diode, $\mathrm{I_{ph}}$ is photovoltaic current source, I is output current of solar array, $\mathrm{R_s}$ is series resistance, $\mathrm{R_{sh}}$ is parallel resistance [7, 8].

Using Eq. (5), we are supposed to find a set of values for unknown model parameters $\mathrm{A_1}$, $\mathrm{A_2}$, $\mathrm{I_{S1}}$, $\mathrm{I_{S1}}$, $\mathrm{I_{ph}}$, $\mathrm{R_P}$ and $\mathrm{R_S}$ in such a way that the equation is just a function of I and V, so that the I-V characteristic fits on the experimental I-V curve of the photovoltaic array with measured values of Table 1.

In this case, by altering V from 0 to $\mathrm{V_{oc}}$ in Eq. (5), we will have different current (I), so that the results obtained from the circuit model matches properly with

Table 1 Measured values

$\mathrm{I_{sc}}$	$\mathrm{V_{oc}}$	$\mathrm{I_{mp}}$	$\mathrm{V_{mp}}$	$\mathrm{N_s}$
3.11	21.3	2.88	17	36

Where the measured values of the PV array are defined as follows

$\mathrm{N_s}$: The number of photovoltaic cells in series

$\mathrm{V_{oc}}$: Open circuit voltage

$\mathrm{I_{sc}}$: Short circuit current

$\mathrm{I_{mp}}$: Current of maximum power

$\mathrm{V_{mp}}$: Voltage of maximum power experimental values of Table 1. The maximum conformity of the values is our goal in this paper. So, the error function of RMSE is defined below as Eq. (6).

$$\mathrm{RMSE} = \sqrt[2]{\frac{1}{5} * (I_{sca} - I_{sc})^2 + (V_{oca} - V_{oc})^2 + \left(I_{mpa} - I_{mp}\right)^2 + \left(V_{mpa} - V_{mp}\right)^2 + \left(P_{mpa} - P_{mp}\right)^2} \tag{6}$$

where the RMSE is the root mean square of errors and we should minimize the error value to obtain the unknown parameters of the model.

Ideally, the 7 model parameters are defined using the proposed optimization algorithm so that the same values of Table 1 could be extracted from the Eq. (5) to obtain the optimal possible state.

3 Proposed GA-PSO Algorithm

This algorithm combines the genetic algorithm and the PSO. The genetic algorithm uses Darwinian selective principles to find the effective formula to forecast or pattern matching. The genetic algorithms are often a good option for regression-based prediction techniques. The flowchart of GA-PSO algorithm is shown in Fig. 2, which was used in this paper.

4 Results and Discussion

By specifying the range of each model parameter, GA-PSO algorithm gives the optimal output using MATLAB programming as shown in Table 2.

According to the results of the parameters from Table 2, the minimum error value (RMSE) of the algorithm is 2.02. Figures 3 and 4 show the I-V and P-V characteristics of the solar array using GA-PSO algorithm, respectively.

The calculation time for each round of this algorithm at the beginning of the program is 1 min and 30 s. By program progressing and aggravating, the calculation time also reaches to 3 min and 30 s. However, it is necessary to expand the domain of the parameters in order to obtain better result, the frequency of repetition and the duration of the entire program will increase. For the values rather reasonable in this algorithm, it takes about 12 h on the PC systems. Finally, the RMSE value of this algorithm is equal to 2.02. Due to the calculation time, the error value obtained from GA-PSO algorithm is appropriate, which seems to be used in optimization of model parameters as well.

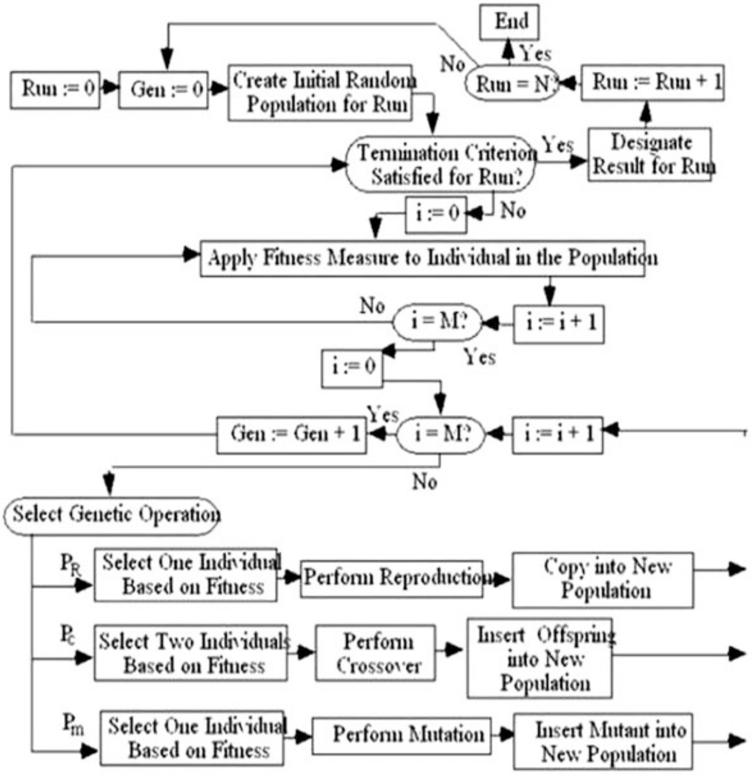

End
Yes
No
Run := 0
Gen := 0
Create Initial Random Population for Run
Run = N?
Run := Run + 1
Termination Criterion Satisfied for Run?
Yes
Designate Result for Run
i := 0
No
Apply Fitness Measure to Individual in the Population
No
i = M?
i := i + 1
Yes
i := 0
Yes
Gen := Gen + 1
i = M?
i := i + 1
No
Select Genetic Operation
P_R
Select One Individual Based on Fitness
Perform Reproduction
Copy into New Population
P_c
Select Two Individuals Based on Fitness
Perform Crossover
Insert Offspring into New Population
P_m
Select One Individual Based on Fitness
Perform Mutation
Insert Mutant into New Population

Fig. 2 The flowchart of GA-PSO algorithm

Table 2 The results of missing parameters with GA-PSO algorithm

R_s	R_p	I_{ph}	I_{s1}	I_{s2}	A1	A2	
0.1–0.001	1000–100	1.5Is–Is	10^{-7}–10^{-10}	10^{-7} –10^{-10}	2.5–1	2.5–1	Range of missing parameters
0.04	916.07	4.45	1.01×10^{-8}	6.82×10^{-8}	1.13	1.54	Result

5 Conclusion

In this paper, exponential double-diode model of PV solar panels was presented to optimize the electrical equivalent circuit parameters. Regarding to the model examined by the proposed GA-PSO algorithm, the minimum error value (RMSE) for the unknown parameters is 2.02, which its calculated time takes along about 1 min and 30 s. The results of the optimization can forecast the I-V and P-V

Fig. 3 I-V characteristic using GA-PSO algorithm

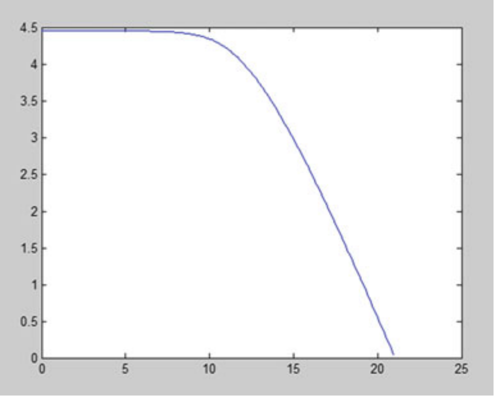
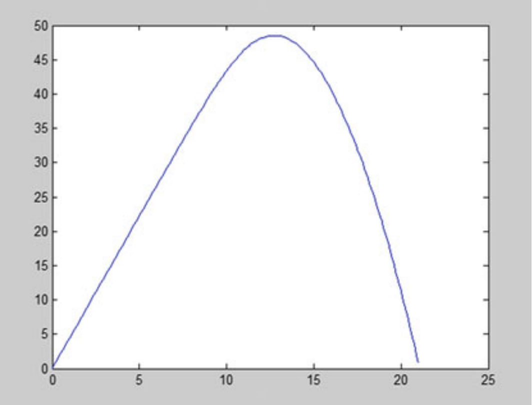


Fig. 4 P-V characteristic using GA-PSO algorithm



characteristics of the PV arrays obtained from experimental measurements as well. Comparing the results of the GA-PSO algorithm shows the accuracy of the proposed model and algorithm for optimizing PV in various conditions.

References

1. Hyvarinen J, Karila J (2003) New analysis method for crystalline silicon cells. In: Proceedings of 3rd world conference on photovoltaic energy conversion, pp 1521–1524
2. Dondi D, Bertacchini A, Brunelli D, Larcher L, Benini L (2008) Modeling and optimization of a solar energy harvester system for self-powered wireless sensor networks. IEEE Trans Ind Electron 55(7):2759–2766
3. Chatterjee A, Keyhani A, Kapoor D (2011) Identification of photovoltaic source models. IEEE Trans Energy Convers 26(3):883–889
4. Campbell RC (2007) A circuit-based photovoltaic array model for power system studies. In: Proceedings of 39th North American power symposium, pp 97–101

5. Nazerian V, Alahgholi M (January 2017) Modelling of photovoltaic solar panels using linear-piece multi-diode equivalent circuit model. In: 4th international conference on electrical & computer engineering, Tehran, Iran, 12 Jan 2017
6. Nazerian V, Firuzjah KhG, Gholamzadeh S, Dizaj MH (March 2017) Simulation of photovoltaic solar cells using MATLAB/Simulink. In: International conference on the new horizons in the basic and technical sciences and engineering, Tehran, Iran, 4 Mar 2017
7. Salam Z, Ishaque K, Taheri H (2010) An improved two-diode Photovoltaic (PV) model for PV system. In: Proceedings on joint international conference, PEDES, pp 1–5
8. Villalva MG, Gazoli JR, Filho ER (2009) Comprehensive approach to modeling and simulation of photovoltaic arrays. IEEE Trans Power Electron 24(5):1198–1208

Part V
Telecommunication Engineering

Hierarchical Routing in Large Wireless Sensor Networks Using a Combination of LPA * and Fuzzy Algorithms

Farhad Mousazadeh and Sayyed Majid Mazinani

Abstract This paper presents a new routing method that increases the network life by combining the fuzzy approach with the A-star algorithm. This algorithm determines the optimal path from source to destination based on maximum battery energy, minimum number of jumps and minimum traffic loads. Due to the limitations of network-aware algorithms for storing the entire grid data in each node's memory, a new clustering strategy has been used. This clustering method identifies the paths that have more densities of the nodes, and we consider them as spinal cords, and so we call it the backbone of the network, and we select the cluster selection based on its proximity to the spine. For comparison, the LPA algorithm without clustering and Patil (a clustering method based on the weight distribution criterion that includes node-level parameters, distance to node neighbors, node speed, and time spent) and Mounir (a new clustering combination with Using LEACH and MTE protocols). The simulation results show that the shelf life of the network created by the proposed method can be increased by extending the network and increasing the number of node.

Keywords Wireless sensor network · Routing · Clustering · LPA*
Fuzzy

1 Introduction

A wireless sensor network is a finite energy network that includes sensor nodes and base stations. The sensors are designed with small batteries, which are often non-replaceable or recharged. Therefore, saving energy is essential. The data must

F. Mousazadeh
Software Engineering, Azad University, Neyshabur, Mashhad, Iran
e-mail: Farhad.mousazadeh@gmail.com

S. M. Mazinani (✉)
Faculty of Engineering, Imam Reza International University, Mashhad, Iran
e-mail: smajidmazinani@imamreza.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_53

be transferred from the nodes to the station using some intermediaries, because the direct transmission of data from each node to the base station requires a great deal of energy. Therefore, the selection of one or more nodes as an interface (cluster head) helps reduce the amount of energy loss.

Because of this, the selected routes from the source node to the base station by clustering play a key role in these networks.

Clustering in sensor networks has the following advantages:

(1) Reduce intra-group communication.
(2) Using cluster heads, load balances can run across the network.
(3) Reduces the update, while limiting most of these messages to intra-group communications.
(4) Increase scalability.

Selecting the best path from each node to the base station is possible when each node has information about all the nodes of the route and, accordingly, chooses the best route. This awareness includes the remaining battery life, the amount of information to be sent in the queue, the distance from the base station, the awareness of the position relative to the neighbors, and …. The algorithms that are based on these notions can certainly have much more efficiency than algorithms that exploit explorations and random methods and increase network lifetime. Due to the limited amount of sensor memory, large networks cannot store all information and the entire network map in their memory [1]. Therefore, the sensor connection to the base station is sent in a multistage. As the grid is divided into hypothetical classes, each node stores the map of the grid in which it is located, and the selected nodes are known as the head of the neighboring grid position.

The proposed method is to investigate the issues of energy-balanced consumption and maximize the life of wireless sensor networks. A new approach combining fuzzy approach and the A-star algorithm to select the best path from source to destination, taking into account the maximum energy of the waste battery, minimum number of mutations and minimum Provides traffic load. The algorithm A * ensures that the path chosen by this algorithm will be the best path [2]. But there are two basic problems in this solution, which seems to be superior to all of the other solutions presented in all tests, which is related to the inherent characteristics of the A * algorithm. The first problem is that if the network nodes work or the new node is added to the set, all nodes must update the tables for their A * algorithm, which imposes one-time processing and transmission of heavy data across the network. In order to solve this problem, the LPA-Star algorithm, which supplied the A * algorithm with the formulation suggestions that partly overcome this problem, but the second problem relates to the amount of memory per node. Given that each node holds all the nodes inside its network to use the A * algorithm, the size of the network will be limited to the amount of memory of each node [3].

To solve this problem in this research, with the clustering of nodes and the use of the backbone in the network due to the density of nodes, we have solved the problem of extensibility in this solution.

1. The main challenge

Currently, many routing protocols are specially designed for data-centric routing protocols, location-based routing, and hierarchical routing. The most important factor in the development of this research is the development of an expandable algorithm by ensuring that the best route is selected by selecting the nodes with the highest remaining energy, the shortest path to the base station and the least traffic on the route.

One of the newest algorithms is a combination of the A * algorithm and fuzzy logic, which according to the existing fuzzy table and its rules, the flexibility of the algorithm for the various applications is very high and the optimal path is guaranteed by the A * algorithm, but at the same time due to the use From the A * algorithm, the problems that are inherent in this algorithm are extensively discussed in this section. In this study, we decided to cluster the nodes and decided to increase the network magnitude and density of the nodes without worrying about the memory capacity of each node to provide broader and wider uses of this algorithm.

2. Research records

In recent years, with the development of computational intelligence (CI), routing protocols based on such intelligence algorithms as RL, ACO, Fuzzy Logic (FL), Genetic Algorithm (GA) and Neural Networks (NNS) Is provided to help improve the performance of the wireless sensor network. These algorithms have been shown to work well under certain conditions such as interruptions of communications, topology changes, and moving nodes. These intelligent algorithms have different properties and requirements, and require a specific software depending on the scenario. In the past, the focus has been on traditional routing protocols and research on routing protocols based on the ACO Ant colony. However, in addition to ACO, many other intelligent algorithms such as RL, FL, GA and NNS are also used to optimize the routing problem for the wireless sensor network. In the following, we will discuss the following:

1. Looking at the optimization of network lifetime in a wireless sensor network, this article has selected some of the routine routing protocols discussed. We are going to present new ideas for the wireless sensor network and provide a way to collaborate between wireless sensor network and artificial intelligence.
2. To evaluate the performance of these algorithms, they are used to increase the network lifetime. Many of the routing algorithms that are proposed for the wireless sensor network in order to increase network lifetime have defined the network lifetime as network startup until the first sensor node loses its energy.

The efficiency of simulation programs and network lifetimes is defined by the following three criteria:

- Start the network as long as the first node of your battery power is lost.
- The first time a live node fails to find a path to transmit data to Sink.
- The amount of data delivered when no data can be sent to the sink.

The results of the measures taken to optimize the network lifetime are as follows: reduction and balance of energy consumption, the use of different paths, reducing the delay time (traffic load) and improving the safe transferability (successful transmission). In the following, we consider the above considerations to investigate routing algorithms for the wireless sensor network.

In [4], an algorithm is used to find k the shortest path in the network using a genetic algorithm, in which a node of node connectivity is first obtained, and then a primary population is constructed with the condition of connecting the path nodes. In this algorithm, bandwidth has a significant effect on the merit function.

Dr. Rucket Kumar used genetic algorithm to optimize the shortest path in [5]. The proposed algorithm performs crossover operations with a probability of 0.7 and a mutation operation with a probability of 10%. Selection of roulette wheels is also used to select chromosomes. To implement the mutation operation, a node is inserted randomly in a random position on the chromosome.

In [6], the genetic algorithm is implemented using C++. The mutation operation is considered to be 0.01–0.08%. The roulette wheel has also been used to select chromosomes for Crossover operations. In order to construct the population, the next stages of the chromosomes are sorted according to the fit of the fitness function, and half of the population that is suitable fit will be transferred to the population of the next stage and the rest of the population will be through the Crossover.

The EAGER protocol of the protocol, which is expressed in [7], divides the area into the gravity, and in each grid, a node is randomly selected using a time counter headed for the grid to receive and transmit sensed data.

The MAXEW mission in this quest proposed in [8] has benefited from a factor called maximum energy welfare in order to reduce energy consumption. Each node, together with its neighbors, forms a local community, and the protocol seeks to maintain energy in this local community.

The RGM's RGM routing method, introduced in [9], uses a spatial location to create multiple paths between the source and the base station, which increases the reliability of data transmission.

The ALURP algorithm for this algorithm proposed in [10] has the ability to support the mobility of the base station. In fact, the sink movement will split the traffic load on all nodes, which will result in a longer lifetime for the sensor network. This protocol, with the formation of circular areas around the mobile sink, will cause the new station to be released in the same areas, which will reduce energy consumption.

The DDRP's DIRP methodology, introduced in [11], supports the mobility of the base station to reduce traffic and reduce the energy consumption of the nodes. Moving the sink by sending messages to the nodes identifies its first, second and third level neighbors, and thus enough sensed data from any place in the network reaches the neighbors of the base station and the neighbors are transmitted to the neighboring lower level to reach the sink.

In the paper [12], the LEACH method delivers sensor nodes in a local cluster, in which a node acts as a cluster head.

The TL-LEACH protocol is presented in [9]. This protocol is developed by the LEACH algorithm. TL-LEACH uses two levels of cluster (primary and secondary) plus other simple nodes. In this algorithm, in each cluster, the primary cluster communicates with the head of the secondary clusters, and the secondary clusters of each cluster

$$C = \frac{\sum_{i}^{n} d_i}{n}$$

The lower the C value the node has a better centrality. The value of the distance of the two nodes relative to each other is calculated from Eq. (2).

$$d = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2}$$

To calculate the central amount of 0.25 we divide it, the lower the C value the smaller V, the C is the central value.

$$V = \frac{0.25}{C}$$

Relationship 2 is used to calculate the distance criterion from the main spine.

Which is considered the smallest distance of one meter and we use Eq. 4 to convert the distance to a number from 0 to 0.25.

$$F = \frac{0.25}{d}$$

Given the above relations, the power of each node is calculated in Fig. 5.

$$\mathrm{P = E + (n * (0.25/N)) + V + F}$$

Select the cluster

After calculating the power of each node, it must be determined which node is selected as the header. To do this, the nodes whose power is greater than zero will compete, and each node will wait as long as T, as T will be calculated from Eq. 6, and after that time, the node sends its message (M) as a cluster head to all its neighbors.

$$\mathrm{T = t_m/P}$$

P is the power of each node, and tm is a constant value for the entire network, which is based on the network structure. The larger the p is, the longer it will wait.

The time that the node waits for one of the following scenarios may occur:

The node receives the message for another node: this state indicates that there is another node that has a higher power and less time to wait, and the node that sent the message is selected as the node and the nodes that this They have received the message and they will not send their message again.

- The node will not receive any messages during this period so the node sends its message to all its neighbors and is selected as the header.

The advantage of this random time is that the nodes enter into controversy with each other and do not calculate the amount of power that they consume this energy operation, a higher-power node will wait for less time and then send its message.

1.1 Selection of Cluster Members

Once the nodes are identified, other nodes must choose their own cluster, if a node has received a message from several clusters, it selects the lava that has more power and sends the connection message to the cluster header and the connection is established.

1.2 Proposed routing method

In this method, the topology of a wireless sensor network is modeled in the direct graph G (N, A), in which N is a set of nodes and A is a set of direct links between nodes. A well node is responsible for collecting data for other nodes in its own transmission range (5, 9, 10, and 26). The routing program is calculated by the main station.

The proposed method assumes:

(1) All sensor nodes are randomly distributed in this area, and each sensing node assumes that its position and the adjacent nodes and wells are known.
(2) All sensor nodes have the maximum transmission range and have the same primary energy value.
(3) Each node has a certain amount of waiting traffic in the node's queue, this queue includes traffic as well as the traffic of packet nodes.

The primary goal of this paper is to design a protocol that maximizes the life of wireless sensor networks by limiting the cost of energy in reciprocating and aggregating data and sharing the same energy consumption. To achieve this goal, we used both the fuzzy approach and the A-star algorithm. The new method uses 3 routing criteria (i.e., maximum waste energy, minimum leakage and minimum traffic load) to select the next optimal jump to reach the base station.

1.2.1 Applying the A-Star Algorithm

In the new routing method, the main station prepares the routing program and distributes it to each node on the network. The A-star algorithm, which is used to find the optimal path from the node to the main station, will apply to all nodes. The tree-like A-star algorithm searches for an optimal routing path from an arbitrary node to the main station. The tree node is discovered on the basis of its evaluation function (f (n)). The function we used is as follows:

$$\mathbf{f}(\mathbf{n}) = \mathbf{NC}(\mathbf{n}) + (\mathbf{1}/\mathbf{MH}(\mathbf{n}))$$

where NC (n) is the node cost for the node n, which is initialized by the output of the fuzzy algorithm and has values from 0 to 1. The fuzzy approach for the waste energy and the traffic load of the node n is considered in order to calculate the optimal cost of the node n. MH (n is the short distance from the node n to the main station). As a result, the node n, which has the largest value f (n), will be selected as the optimal node.

1.2.2 Applying the Fuzzy Approach

The goal of the fuzzy section of the proposed protocol is to determine the optimal cost of the node (NC) of the node n, which depends on the residual energy (RE) and load traffic (TL) of the n node. Figure 1 is a fuzzy approach with two variable inputs RE (n) and TL (n) and one output NC (n) with reference sets [5 … 0], [10 … 0] and [10 … 0] respectively. The new method includes 5 membership functions for each variable.

The input and output are as shown in Fig. 1.

The fuzzy algorithm contains a series of rules if-then that interconnects the input fuzzy variables and the output variables using descriptive variables, each with fuzzy sets and the fuzzy argument operator AND. Table 1 shows the rules if-then used in the proposed method, with a total number of 25 lines for fuzzy rules. For example, if RE (n) is very high and TL (n) is very small, then NC (n) will be very high. Finally, non-phaseizing will result in a definite output value of the fuzzy response space.

The fuzzy algorithm contains a series of rules if-then that interconnects the input fuzzy variables and the output variables using descriptive variables, each with fuzzy sets and the fuzzy argument operator AND. Table 1 shows the rules if-then used in the proposed method, with a total number of 25 lines for fuzzy rules. For example, if RE (n) is very high and TL (n) is very small, then NC (n) will be very high. Finally, non-phaseizing will result in a definite output value of the fuzzy response space.

1.3 Assured Aggregation by Cluster Head

One of our proposed solutions is to aggregate data by the cluster so that it prevents duplicate data transmission, which reduces energy consumption. Each cluster is tasked with collecting cluster data that acts as follows.

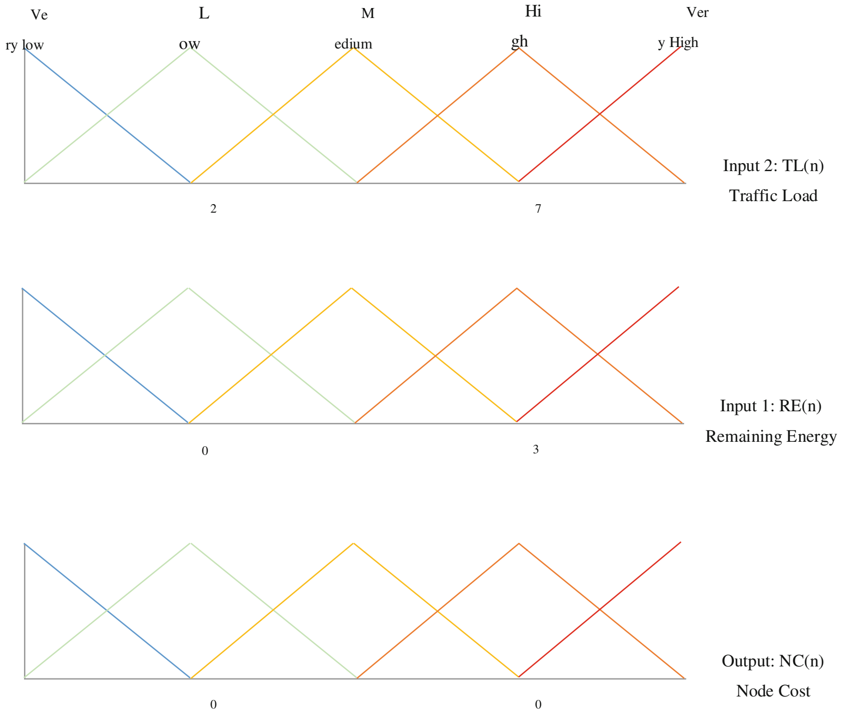


Fig. 1 Fuzzy logic inputs

- According to a timetable determined by the base station, the cluster sends the collected data at specific timescales. This timing makes the cluster to send their data in turn to prevent high traffic congestion in the main path.

As long as the data transmission time is not in the cluster header, all sensors send their data to the cluster header. At this time, the cluster collects data and has the opportunity to filter them

- The header does not consider duplicate and incomplete data and prevents them from sending to the base station.
- Each cluster encodes and frames data that consumes less energy at the time of sending
- Useful and collected data sent by the cluster to the spine at post time.
- The pseudo-code for data aggregation is as follows.

```
for (all ch node)
   [1] {
         o   CH specifies time
         o   all node send data for CH
         o   CH filtering data and coding
         o   CH send all data to sink
     }
```

Table 1 Fuzzy logic output

R	Reference		Result
	RE(n)	TL(n)	NC(n)
1	VL	VL	L
2	VL	L	VL
3	VL	M	VL
4	VL	H	VL
5	VL	VH	VL
6	L	VL	M
7	L	L	M
8	L	M	L
9	L	H	L
10	L	VH	VL
11	M	VL	H
12	M	L	M
13	M	M	M
14	M	H	L
15	M	VH	L
16	H	VL	VH
17	H	L	H
18	H	M	H
19	H	H	M
20	H	VH	M
21	VH	VL	VH
22	VH	L	VH
23	VH	M	VH
24	VH	H	H
25	VH	VH	H

In this step, the cluster uses a scheduler to delete duplicate data as shown below.

After clustering to extend the network life and overlapping the environment, sensors that are close to each other or have a high environmental contribution can be active, so that, according to a scheduling algorithm at any one moment, one of them is active. The advantage of this method is this the sensors that are close to each other and transmit the same information do not work at the same time and will disappear later, which will increase the quality of coverage. The process of deactivating adjacent nodes is as follows

- Each cluster is divided into four parts (four quadrants of trigonometry)
- If the size of each quarter of the size of the environment to be covered by each sensor is v, then at least n = s/v in each quadrant is required. Now, if the number of sensors is greater than n, the sensors are turned on, each sensor turns off after sending each message and the next sensor is turned on.

1.4 Send Data

Once the clustering and backbone of the network are formed and the data is gathered in a network, each cluster saves its data in the spine so that the information reaches the main server through the spine to allow traffic to be controlled. And no congestion. A scheduler has been used as follows

- Blocks are allowed to send data at certain times
- This timing is based on the cluster's distance from the main station because the cluster is closer to the base station sooner its data.
- As far away as possible, the cluster head sends data at shorter intervals. Closer clusters send late to late delivery.
- This timing is fixed and will be performed by the main server and will cause decent traffic on the main routes.

The pseudocode of this section is as follows.

```
for (all node)
           {   T=Time calculated based on the distance
               if T=live time node can send data
           }
```

2 Simulation

The method presented in this study for simulating clustering in this section is simulated with MATLAB software and is compared by three different methods according to several important criteria, the results of which are described below.

2.1 Comparative Methods

To compare the proposed method of method [13], this paper is used to rout and maintain energy from LPA-based routing.

The second method is to compare a cluster-based method for clustering and determining the cluster head dynamically based on a weight distribution criterion, which includes one or more parameters such as node degree, distance to node neighbors, node speed, and time spent as a The cluster head uses [14].

The third comparison method for routing uses a new clustering approach based on the combination of LEACH and MTE protocols [15].

2.2 Simulation Parameters

Matlab software is used to implement proposed algorithms. Simulations are carried out in an environment of 100 m at 100 m. In order to investigate the efficiency of the algorithm, the coverage parameters, the number of live nodes and the remaining energy of the proposed algorithm are compared with the other three methods. As discussed in Chap. 3, the optimal algorithm should have more live nodes and more energy remaining in different periods of implementation and high packet delivery rates. Since the early pattern of initial dispersion of the sensors has a great influence on the validation of these parameters in each round in which a simulation is being implemented, the original pattern is applied to all algorithms.

2.3 Network Lifetime

One of the most important evaluation criteria for a wireless sensor network is the lifetime of a network. The network lifetime is actually equal to the length of time that the network can continue to operate until all sensors lose energy, in order to assess this important criterion, two scenarios opinions have been asked.

The first mode

- Fixed network size
- The number of variable network nodes

In this scenario, the network size is fixed, but in each step the number of nodes is increased. The nodes are randomly distributed in the network. For the exact criterion to be calculated, for each step, 10 runs are performed, and the average value of these 10 steps is considered as the output. In Fig. 2, this criterion is apparent for the scenario described.

The condition of the second scenario is that the number of distributed nodes in the network is constant, but the network size is increased in each step. The result is shown in Fig. 3.

In this scenario, given the fact that the grid is larger at each stage, the number of messages is greater, and due to this, the network life span is descending, but in both scenarios the proposed method is better.

2.4 Number of Live Nodes

One of the most important criteria is the clustering of the number of live nodes in the network over time. To simulate this section, Fig. 4 shows the network size and number of nodes, and the number of live nodes is displayed after a time lapse.

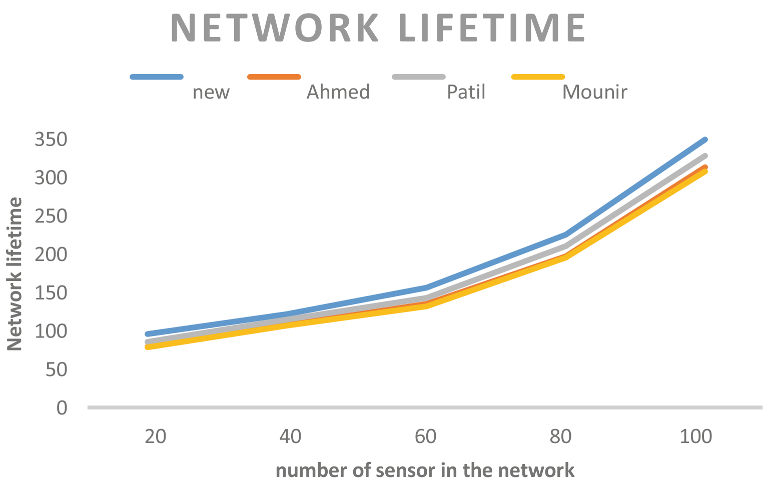


Fig. 2 Network lifecycle with variable numbers

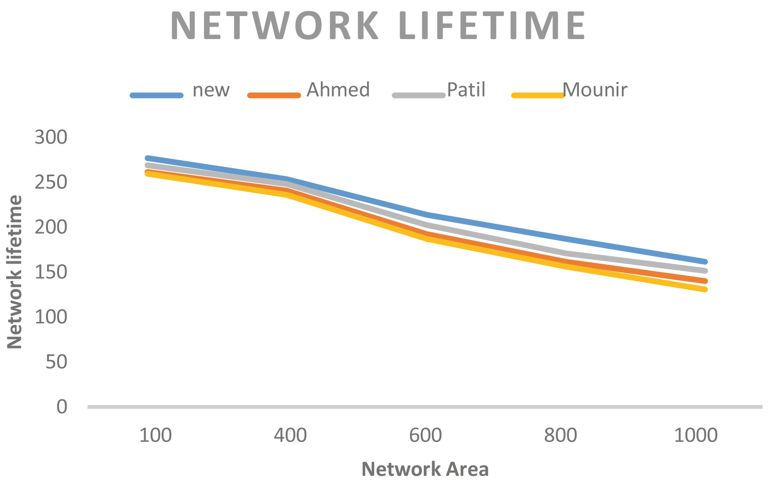


Fig. 3 Network lifecycle with variable grid size

According to the results of the simulation, the proposed method has better performance. The reason for improving the proposed method is because of the main criteria for route and clustering.

2.5 Number of Received Data

This criterion shows how much of the data sent by the node reaches the base station. In this scenario, the network size and sensor number are fixed, and the result is shown in Fig. 5.

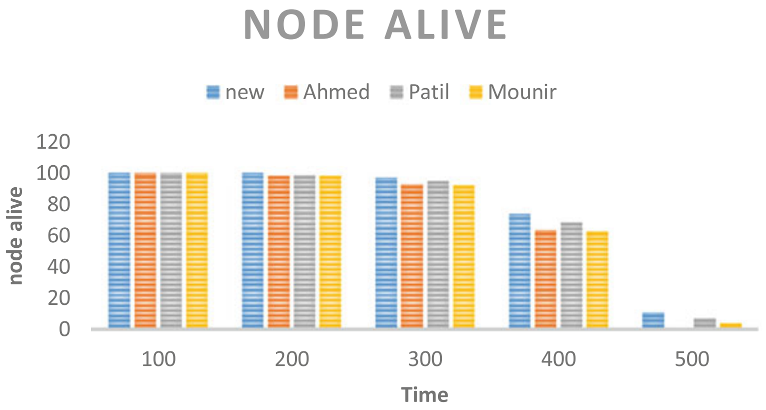


Fig. 4 Number of live nodes

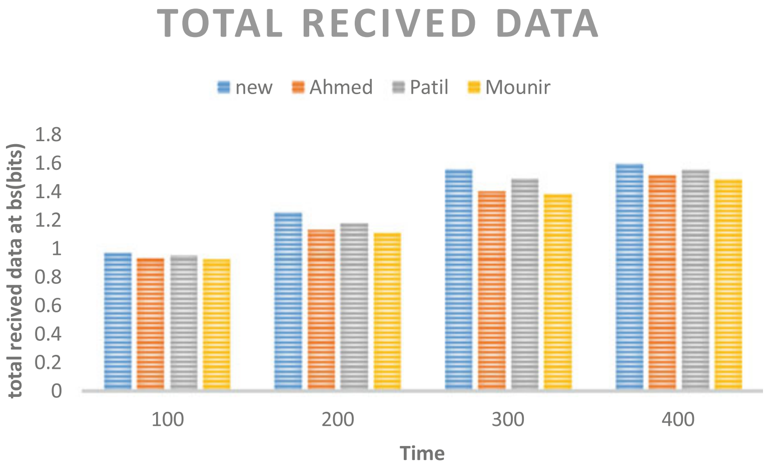


Fig. 5 Number of received data

In the proposed method, due to the use of scheduling, the correct sending rate has increased.

2.6 Numbers

In Fig. 6, the number of clusters in the network is displayed in different periods. In this simulation, the network size and the number of nodes are constant.

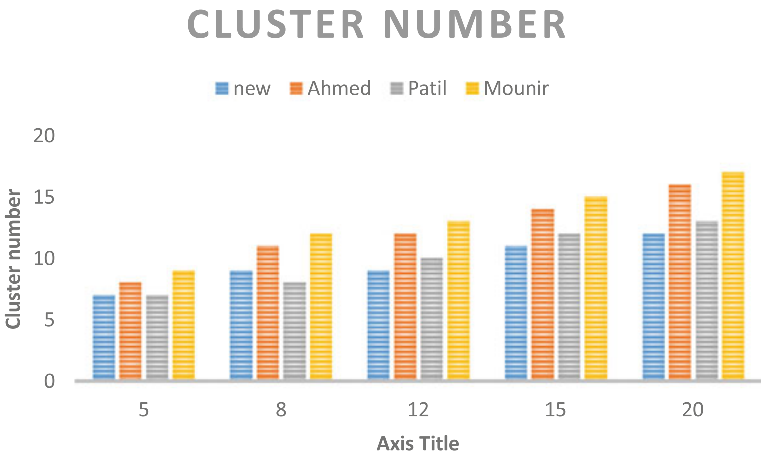


Fig. 6 Number of headroom

In the proposed method, the number of clusters is lower, and the yen makes energy savings. The reason for the decrease in the number of clusters is that it is considered in the clustering of centrality and number of neighbors.

2.7 *The Amount of Remaining Energy*

The amount of energy remaining in the entire network in different periods of execution of the algorithm is shown in Fig. 7.

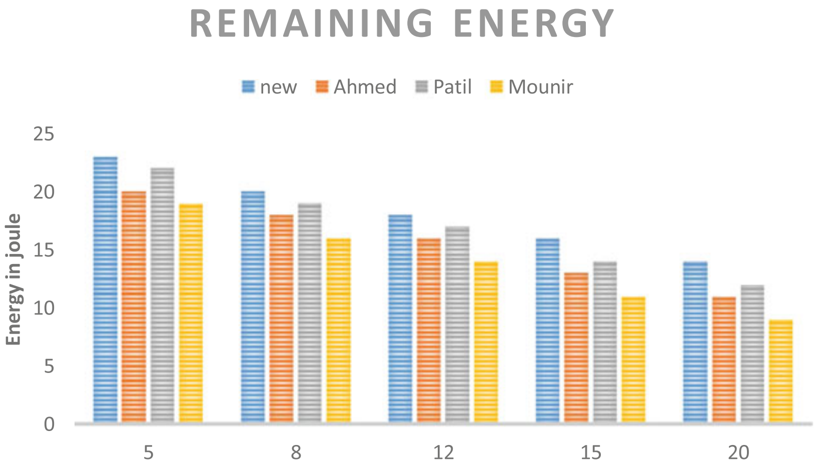


Fig. 7 The amount of energy remaining

3 Conclusion

The simulation results show that the proposed method is very good in energy consumption, which reduces consumption by increasing the life of the network. Also, by choosing the optimal cluster and transmitting data completely aware of the general state of the grid and a very good balance in energy consumption, with increasing number of nodes in the grid, this algorithm uses the other algorithms compared.

4 Future Work

Considering the clustering criterion that is considered to be the density of nodes in a network, in networks with nodes being distributed uniformly and uniformly, this algorithm lost its clustering performance and results are somewhat weaker than the other Methods are obtained. Considering that the fuzzy algorithm used for routing in this paper can use different inputs for specific methods, which makes it possible to customize this algorithm for specific applications and easily modify the creation of Made.

References

1. Ammar AB (2016) Multi-hop LEACH based cross-layer design for large scale wireless sensor networks. In: Wireless communications and mobile computing conference (IWCMC), 2016 International, 29 Sept 2016
2. Magno M, Boyle D, Brunelli D, Popovici E, Benini L (2014) Ensuring survivability of resource intensive sensor networks through ultra-low power overlays. IEEE Trans Industr Inform 10(2):946–956
3. Belabed F (2016) An optimized weight-based clustering algorithm in wireless sensor networks. In: Wireless communications and mobile computing conference (IWCMC), 2016 International
4. Kumrawat M, Dhawan M (2015) Survey on clustering algorithms of wireless sensor network. (IJCSIT) Int J Comput Sci Inf Technol 6(3):2046–2049
5. Bhat S, Pai V, Kallapur PV (2015) Energy efficient clustering routing protocol based on LEACH for WSN. Int J Comput Appl 120(13):0975–8887
6. Desai K, Rana K (2015) Clustering technique for wireless sensor network. In: International conference on next generation computing technologies (NGCT), pp 223–227, Sept 2015
7. Ye X, Li J, Chen WT, Tang F (2015) LT codes based distributed coding for efficient distributed storage in wireless sensor networks. In: IFIP networking conference (IFIP networking), pp 1–9, May 2015
8. Alaouil N, Cances JP, Meghdadi V (2015) Energy consumption in wireless sensor networks for network coding structure and ARQ protocol. In: 1st international conference on electrical and information technologies ICEIT', pp 317–321, March 2015
9. Priti KH, Manali S (2015) Improved LEACH protocol using vice cluster in wireless sensor networks. Int J Innovative Comput Sci Eng 2(3):30–34

10. Ajay S, Sushil K (2015) Energy efficient clustering in heterogeneous wireless sensor networks using degree of connectivity. Int J Comput Netw Commun 7(2):19–31
11. Abdi A, Zakerolhosseini A (2014) CFMTL: clustering wireless sensor network using fuzzy logic and mobile sink in three-level. ACSIJ Adv Comput Sci: Int J 3(12):23–28
12. Gajjar S, Talati A (2015) FUCP: fuzzy based unequal clustering protocol for wireless sensor networks. In: Systems conference (NSC), 2015 39th National
13. Alkadhmawee AA, Lu S (2016) Prolonging the network lifetime based on LPA-star algorithm and fuzzy logic in wireless sensor network. In: 2016 12th World Congress on intelligent control and automation (WCICA) June 12–15, 2016, Guilin, China
14. Patil M, Biradar RC (2017) Energy efficient weighted clustering algorithm in wireless sensor networks. Global J Comput Sci Technol XVII(II) Version Year (E) 2017
15. Arioua M, el Assari Y, Ez-zazi I, el Oualkadi A (2016) Multi-hop cluster based routing approach for wireless sensor networks. In: The 7th international conference on ambient systems, networks and technologies (ANT 2016)

Improving Security Using Blow Fish Algorithm on Deduplication Cloud Storage

Hamed Aghili

Abstract Nowadays, most of the commercial processes have been digitized. Data mostly is of great value, thus any damage or loss of it can be a great disaster for its owner. Large enterprises want to store information in a place with maximum security and low cost. One of the information-storing place is cloud that the world moving toward it. The providers of storage space are trying to improve security in the cloud. To decrypt of data in cloud storage server 64-bits secret-key block cipher called Blowfish algorithm is used. The blowfish algorithm has improved in point of security and performance comparing DES, 3DES, AES. In this paper we assume a deduplication storage server and set the blowfish. At 20 Mb block size the time of blowfish algorithm was (1.7 t) comparing with other algorithms. Also, failure time of blowfish was 60 t that it was less than DES, AES, and 3DES failure time. The results are shown with blowfish algorithm the security of storage data improved. In addition, this algorithm was implemented on the deduplication server. The Winhex output has shown all data in the encrypted format that confirmed the attacker couldn't access to the original data.

Keywords Blowfish · Encryption · Security · Reduplication server

1 Introduction

In this modern world, data, images, documents are stored more in computers, hard disk, compact disks. That in virtual cloud that it called cloud storage. These data might be stored in unsecured cloud while some attacker agrees to storage thus we need images be in secure at cloud.

Data security is an essential part of an organization; it can be achieved by the using various methods. In order to maintain and upgrade the model still efforts are

H. Aghili (✉)
Department of Electrical Engineering (Robotic engineering),
Payame Noor University (PNU), Tehran, Iran
e-mail: engineer.aghili@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_54

required and increase the marginally overheads. The encrypted data is safe for some time but never think it is permanently safe [1].

Cryptography provides a solution for this problem. Cryptography can be defined as the art of safeguarding images and it makes sure that only the intended people are able to visualize its content. Every security system must provide a bundle of security functions that can assure the secrecy of the system. These functions are usually referred to as the goals of the security system. The five main goals behind using Cryptography include Confidentiality, Authentication, Integrity, Non-Repudiation, Service Reliability, and Availability. These objectives ensures that the private data remains private, the data is not altered illegally and assures against a party denying a data or a communication that was initiated by them. There are two types of cryptography namely secret key cryptography and public key cryptography. In secret key cryptography, both the sender and receiver know the same secret code called key. Images are encrypted by the sender using the key and the receiver decrypts it using the same key. In public key cryptography, sender and receiver uses different key for encryption and decryption. The sender encrypts the data using a public key and this key will be known by all the parties included in the communication. The receiver decrypts the data using a private key and it should be kept as a secret. The word key is unavoidable term, which can be a word, number or a phrase. Knowing the algorithm without the key does not help the hacker untangle the information. Encryption is the process by which information is transformed to an unreadable form called cipher text where its contents are hidden to eavesdroppers. Any person who sees the cipher text will not be able to determine anything about the original message. An encryption scheme usually needs a key generation algorithm to randomly produce keys. Decryption is the process of retrieving the original data back from the cipher text. A decryption process is generally the reverse of encryption process. Both processes make use of corresponding key. Longer the encryption key is, the more difficult it is to decode [2].

2 Cryptography

When using symmetric algorithms, both parties share the same key for en- and decryption. To provide privacy, this key needs to be kept secret. Once somebody else gets to know the key, it is not safe anymore. Symmetric algorithms have the advantage of not consuming too much computing power. A few well-known examples are DES, Triple-DES (3DES), IDEA, CAST5, BLOWFISH, and TWOFISH [3].

Asymmetric algorithms use pairs of keys. One is used for encryption and the other one for decryption. The decryption key is typically kept secretly, therefore called “private key” or “secret key”, while the encryption key is spread to all who might want to send encrypted messages, therefore called “public key”. Everybody having the public key is able to send encrypted messages to the owner of the secret key. The secret key cannot be reconstructed from the public key.

Asymmetric algorithms seem to be ideally suited for real-world use: As the secret key does not have to be shared, the risk of being known is much smaller. Every user only needs to keep one secret key in secrecy and a collection of public keys that only need to be protected against being changed. With symmetric keys, every pair of users would need to have an own shared secret key. Well-known asymmetric algorithms are RSA, DSA, and ELGAMAL [3].

However, asymmetric algorithms are much slower than symmetric ones. Therefore, in many applications, a combination of both is being used. The asymmetric keys are used for authentication and after this have been successfully done; one or more symmetric keys are generated and exchanged using the asymmetric encryption. This way the advantages of both algorithms can be used. Typical examples of this procedure are the RSA/IDEA combination of PGP2 or the DSA/ BLOWFISH used by GnuPG [3].

As a mention below, we need a high-speed algorithm with advance performance in few times, so in this article we choose asymmetric security algorithm.

3 Asymmetric Algorithms

To encrypt the data various cryptographic algorithms such DES, 3DES, blowfish, AES, etc. are used. First of all, we compare these algorithm in security, architecture, limitation and flexibility.

3.1 DES

Data Encryption Standard (1974), designed by IBM [4] based on their Lucifer cipher was the first encryption standard to be published by NIST (National Institute of Standards and Technology). The DES was initially considered as a strong algorithm, but today the large amount of data and short key length of DES limits its use [5].

In architecture, DES is symmetric key algorithm based on the backbone concept of Feistel Structure. The DES is a block cipher that uses a 64 bit plain text with 16 rounds and a Key Length of 56-bit, originally the key is of 64 bits (same as the block size), but in every byte 1 bit in has been selected as a ‘parity’ bit, and is not used for encryption mechanism. The 56 bit is permuted into 16 sub- keys each of 48- bit length. It also contains 8 S- boxes and same algorithm is used in reversed for decryption [5].

The security strength of DES depend on its 56 bit key size generating 7.2 × 1016 possible keys, making it extremely difficult to originate a particular key in typical threat environments. Moreover, if the key is changed frequently, the risk of unauthorized computation or acquisition can be greatly moderated. Moreover DES exhibits a strong avalanche effect i.e. a miniature modification in the plaintext or key, might change the cipher text noticeably. Initially DES was considered secure and was difficult to crack; Brute-force attacks became a subject of speculation immediately after the release of algorithm is in public domain, although DES survives different linear and differential attacks but in 1998 Electronic Frontier Foundation (EFF) designed a special-purpose machine for decrypting DES. In one demonstration, it achieves the key of an encrypted message [6] in less than a day in combination with an alliance of computer users all around the world. In general, DES was proved insecure for large corporations or governments and it is simpler not to use DES algorithm. However for backward compatibility, and cost of upgrading, DES should still be preferred, outweighing the risk of exposure.

In limitation, DES is highly vulnerable to linear cryptanalysis attacks, Weak keys is also a great issue. DES is also exposed to brute force attack [7].

3.2 Triple-DES

DES was superseded by triple DES (3DES) in November 1998, concentrating on the noticeable imperfections in DES without changing the original structure of DES algorithm. TDES was much more complicated version of DES achieving high level of security by encrypting the data using DES three times using with three different unrelated keys. 3DES [8] is still approved for use by US governmental systems, but has been replaced by the advanced encryption standard (AES) Sub subsections [9].

In architecture, 3DES is exactly what it is named–it performs 3 iterations of DES encryption on each block. As it is, an enhanced version of DES so is based on the concept of Feistel Structure. The 3DES uses a 64-bit plain text with 48 rounds and a Key Length of 168- bits permuted into 16 sub- keys each of 48- bit length. It also contains 8 S- boxes and same algorithm is used in reversed for decryption [9].

In security, TDES is an enhanced version of DES; 3DES use a larger size of key (i.e. 168-bits) to encrypt than that of DES. DES operations (encrypt-decrypt-encrypt) are performed 3 times in 3DES with 2-3 different keys, offering 112 bits of security, avoiding so called meet-in-the-middle attack [10]. TDES offers high level of security in comparison with DES and still in use by the US government.

In limitation, 3 DES is exposed to differential and related-key attacks. It is also susceptible to certain variation of meet-in- in-the-middle attack [10].

3.3 AES

Rijndael developed by Joan Daemen and Vincent Rijmen, becomes U.S.'s new Advanced Encryption Standard in October 2000 declared by the National Institute of Standards and Technology. Rijndael using variable key size is extremely fast and compact cipher. Its symmetric and parallel structure provides great flexibility for implementers, with effective resistance against cryptanalytic attacks [8]. AES can be well adapted to a wide range of modern processors such as Pentium, RISC and parallel processors. In general, AES is the name of the standard, and Rijndael is the algorithm described however, in practice the algorithm is also referred to as AES.

In architecture, AES is also a symmetric key algorithm based on the Feistel Structure. The AES is a block cipher that uses a 128-bit plain text with variable 10, 12, or 14 rounds (Rijndael's Default of Rounds is dependent on key size. Default of Rounds = key length/32 + 6) and a variable Key Length of 128, 192, 256 bit permuted into 10 sub- keys each of 128, 192, 256 bit length respectively. It only contains a single S- box and same algorithm is used in reversed for decryption.

In limitation, AES(R) has no serious weakness; although it was observed that a mathematical property (not an attack) of the cipher might be vulnerable into an attack. Further in AES (Rijndael) the inverse cipher implementation is inappropriate on a smart card than the cipher itself [11].

3.4 Blowfish

Blowfish by Bruce Schneider, author of Applied Cryptography, is considered as a highly rated encryption algorithm in terms of security, with different structure and Functionality than the other mentioned encryption algorithms.

In architecture, Blowfish is also a symmetric key Feistel Structured algorithm consisting of 2 parts: key expansion part and data-encryption part. Blowfish is a block cipher that uses a 64 bit plain text with 16 rounds, allowing a variable key length, up to 448 bits, permuted into 18 sub- keys each of 32- bit length and can be implemented on 32- or 64-bit processors. It also contains 4 S- boxes and same algorithm is used in reversed for decryption [12].

In security, Blowfish's security lies in its variable key size (128–448 bits).

Providing high level of security, Attempts to cryptanalysis Blowfish started soon after its publication however less cryptanalysis attempts were made on Blowfish than other algorithms. Blowfish is invulnerable against differential related key attacks, since every bit of the master key involves many round keys that are very much independent, making such attacks very complicated or infeasible. Such autonomy is highly enviable.

In limitation, Blowfish has some classes of weak keys. Four rounds of blowfish are exposed to second order differential attacks. Therefore, reliability of Blowfish is questionable due to the large no. of weak keys [11, 13].

3.5 Flexibility of All Algorithms

The status of DES, AES, 3DES and Blowfish. Algorithms are compared in Table 1.

Table 1 Comparing of algorithms in flexibility

Algorithms	Flexible	Modification	Commands
DES	NO	None	The structure of DES doesn't support any modifications
Triple-DES	Yes	168	The structure of 3DES is same as DES, it doesn't support any changes but as it iterates DES 3 times so the key size is extended to 168 bits
AES	Yes	128.192.256	The strucftire of AES(R) was extendable to the multiple of 64 bits, have same sub key size as the size of the key
Blowfish	Yes	64-448	Blowfish key length must be multiples of 32 bits

4 Choose Appropriate Algorithm

According to algorithm comparison, we need an appropriate algorithm which making most of the encrypted output in specific time. All of algorithms were coded in Python, compiled with C++ and simulation with Johnny the Ripper. Figure 1 show the result of advantage of Blowfish algorithm over algorithm in term of processing data time it show Blowfish need less time than another algorithms and it can encrypted more data in identical time.

Figure show Blowfish need less time than another algorithms and it can encrypted more data in identical time.

For comparing algorithm in security, we use failure time parameter, this mean we encrypted same data with each algorithm then output checked by Johnny the ripper tools and then consider every algorithm would be broken in several cases and different time that result show algorithm successful in protection data and the result is shown in Fig. 2.

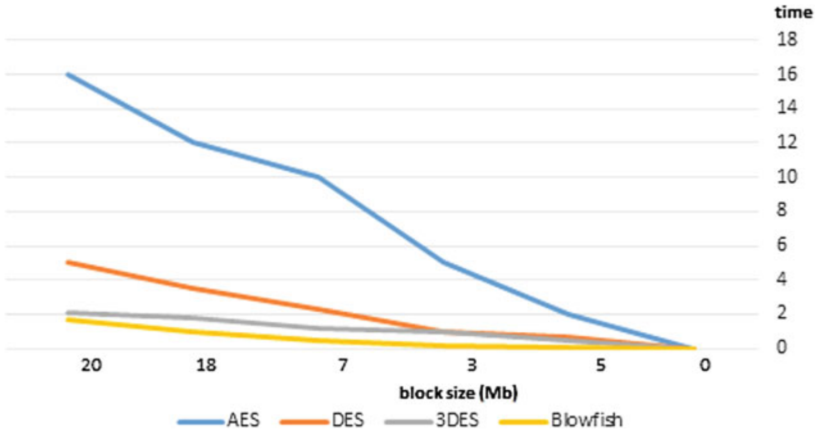


Fig. 1 Performance result of speed

5 Implementing

In this paper, we implemented the Blowfish algorithm using Python and set it on deduplication entrance, which it is a simulation of IBM Deduplication service. Each sever can have a different type of entrance so we make a filtering on server that received just image with JPG format. First, data checked base on filtering then store on server after that, encryption algorithm applied on the data and we can see result of this output by WinHex (Fig. 4). With compare, original image (Fig. 3) with

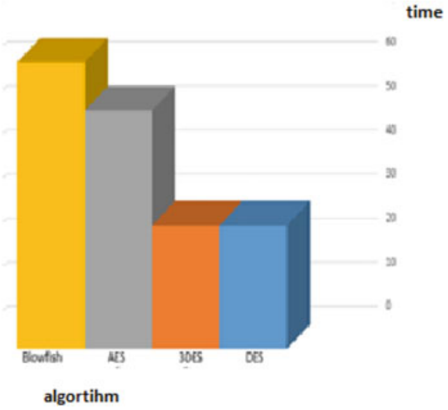
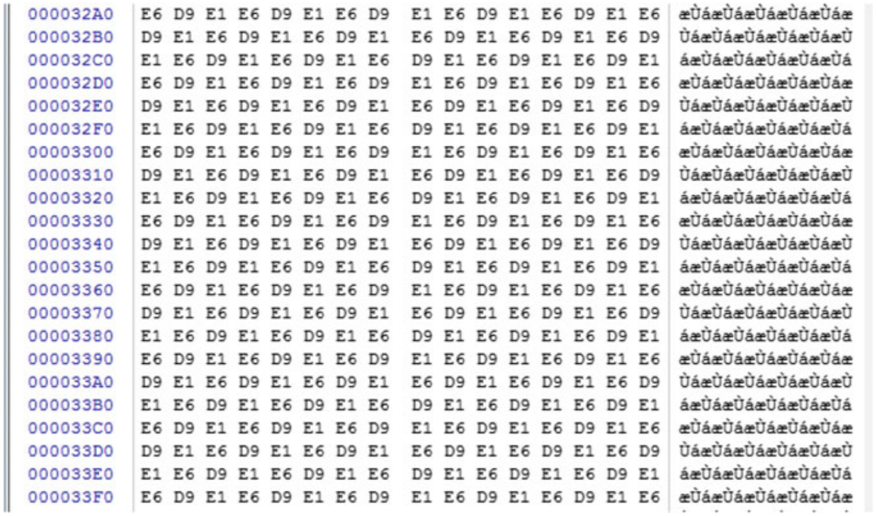
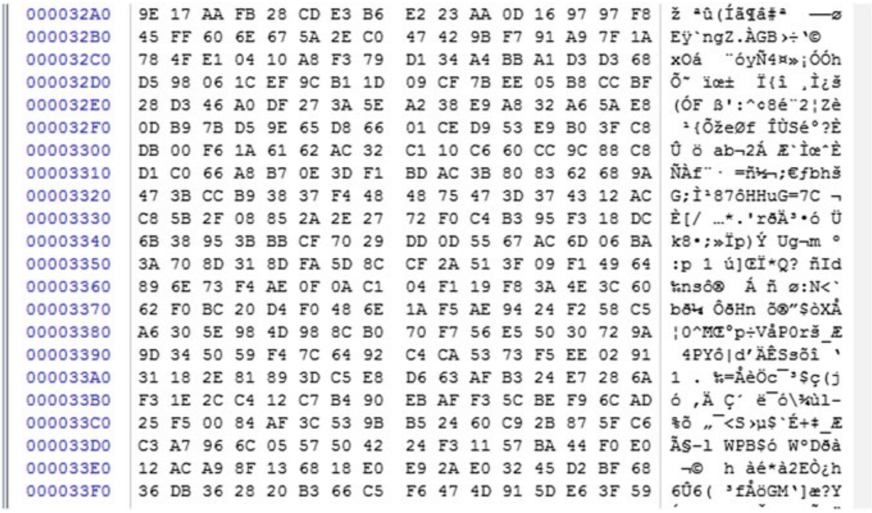


Fig. 2 Failure result



```
000032A0  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
000032B0  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
000032C0  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
000032D0  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
000032E0  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
000032F0  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
00003300  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
00003310  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
00003320  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
00003330  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
00003340  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
00003350  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
00003360  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
00003370  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
00003380  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
00003390  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
000033A0  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
000033B0  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
000033C0  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
000033D0  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
000033E0  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
000033F0  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
```

Fig. 3 Original image of input data

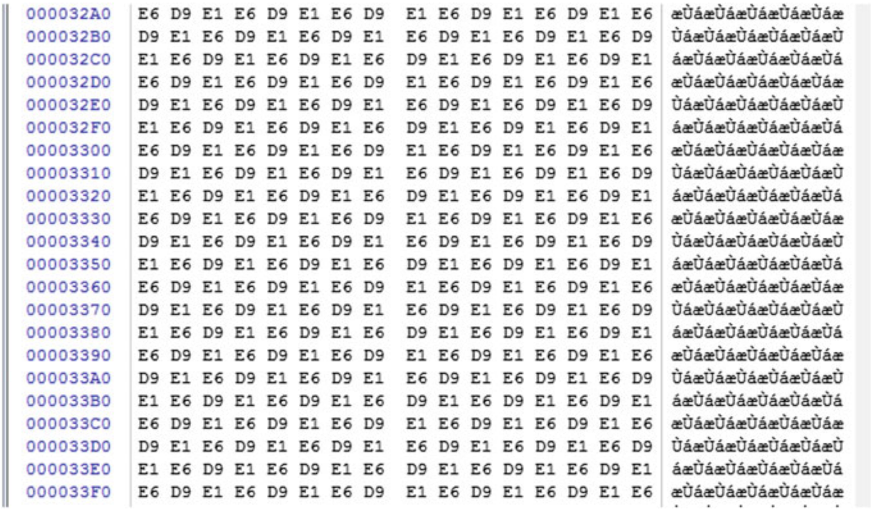


```
000032A0  9E 17 AA FB 28 CD E3 B6  E2 23 AA 0D 16 97 97 F8  ž ªû(Íã¶â#ª  —ø
000032B0  45 FF 60 6E 67 5A 2E C0  47 42 9B F7 91 A9 7F 1A  Eÿ`ngZ.ÀGB›÷'©
000032C0  78 4F E1 04 10 A8 F3 79  D1 34 A4 BB A1 D3 D3 68  xOá  ¨óyÑ4¤»¡ÓÓh
000032D0  D5 98 06 1C EF 9C B1 1D  09 CF 7B EE 05 B8 CC BF  Õ˜  ïœ±  Ï{î ¸Ì¿
000032E0  28 D3 46 A0 DF 27 3A 5E  A2 38 E9 A8 32 A6 5A E8  (ÓF ß':^¢8é¨2¦Zè
000032F0  0D B9 7B D5 9E 65 D8 66  01 CE D9 53 E9 B0 3F C8   ¹{ÕžeØf ÎÙSé°?È
00003300  DB 00 F6 1A 61 62 AC 32  C1 10 C6 60 CC 9C 88 C8  Û ö ab¬2Á Æ`Ìœˆ È
00003310  D1 C0 66 A8 B7 0E 3D F1  BD AC 3B 80 83 62 68 9A  ÑÀf¨· =ñ½¬;€ƒbhš
00003320  47 3B CC B9 38 37 F4 48  48 75 47 3D 37 43 12 AC  G;Ì¹87ôHHuG=7C ¬
00003330  C8 5B 2F 08 85 2A 2E 27  72 F0 C4 B3 95 F3 18 DC  È[/ …*.'rðÄ³•ó Ü
00003340  6B 38 95 3B BB CF 70 29  DD 0D 55 67 AC 6D 06 BA  k8•;»Ïp)Ý Ug¬m º
00003350  3A 70 8D 31 8D FA 5D 8C  CF 2A 51 3F 09 F1 49 64  :p 1 ú]ŒÏ*Q? ñId
00003360  89 6E 73 F4 AE 0F 0A C1  04 F1 19 F8 3A 4E 3C 60  ‰nsô®  Á ñ ø:N<`
00003370  62 F0 BC 20 D4 F0 48 6E  1A F5 AE 94 24 F2 58 C5  bð¼ ÔðHn õ®”$òXÅ
00003380  A6 30 5E 98 4D 98 8C B0  70 F7 56 E5 50 30 72 9A  ¦0^˜M˜Œ°p÷VåP0rš
00003390  9D 34 50 59 F4 7C 64 92  C4 CA 53 73 F5 EE 02 91   4PYô|d'ÄÊSsõî '
000033A0  31 18 2E 81 89 3D C5 E8  D6 63 AF B3 24 E7 28 6A  1 . ‰=ÅèÖc¯³$ç(j
000033B0  F3 1E 2C C4 12 C7 B4 90  EB AF F3 5C BE F9 6C AD  ó ,Ä Ç´ ë¯ó\¾ùl-
000033C0  25 F5 00 84 AF 3C 53 9B  B5 24 60 C9 2B 87 5F C6  %õ „¯<S›µ$`É+‡_Æ
000033D0  C3 A7 96 6C 05 57 50 42  24 F3 11 57 BA 44 F0 E0  Ã§–l WPB$ó WºDðà
000033E0  12 AC A9 8F 13 68 18 E0  E9 2A E0 32 45 D2 BF 68   ¬©  h àé*à2EÒ¿h
000033F0  36 DB 36 28 20 B3 66 C5  F6 47 4D 91 5D E6 3F 59  6Û6( ³fÅöGM']æ?Y
```

Fig. 4 Encrypted image of input data

encrypted image (Fig. 4) see all bits of output changed and unrecognizable that prevent of penetration attackers.

If user needs own data, server authentication user then decrypted data provided for user. As we can see, result of the decrypted data with original data is same (Fig. 5). Sometimes, we have a little different between decrypted file and original



```
000032A0  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
000032B0  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
000032C0  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
000032D0  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
000032E0  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
000032F0  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
00003300  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
00003310  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
00003320  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
00003330  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
00003340  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
00003350  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
00003360  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
00003370  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
00003380  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
00003390  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
000033A0  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
000033B0  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
000033C0  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
000033D0  D9 E1 E6 D9 E1 E6 D9 E1  E6 D9 E1 E6 D9 E1 E6 D9  ÙáæÙáæÙáæÙáæÙáæÙ
000033E0  E1 E6 D9 E1 E6 D9 E1 E6  D9 E1 E6 D9 E1 E6 D9 E1  áæÙáæÙáæÙáæÙáæÙá
000033F0  E6 D9 E1 E6 D9 E1 E6 D9  E1 E6 D9 E1 E6 D9 E1 E6  æÙáæÙáæÙáæÙáæÙáæ
```

Fig. 5 Decrypted image of input data

file, which this changing is below the value threshold, the result of this changing cannot seen.

6 Conclusion

In this paper, first of all we prepared a detailed analysis of symmetric block encryption algorithms is presented on the basis of different parameters and we choose a blowfish algorithm by conclusion and comparative of all algorithms for improving security in deduplication server. In this way, the images stored on server, from which the largest number of bits in the encryption algorithm used blowfish compare with other algorithms can be encrypted in a very short time and does not change the original file.

References

1. Scholar VBR (2012) Implementation of new advance image encryption algorithm to enhance security of multimedia component. Int J Adv Technol Eng Res 2(4)
2. Abdul Jaleel J (2008) Guarding images using a symmetric key cryptographic technique: blowfish algorithm. JEIT 3(2) (Certified)
3. Garloff K (2000) Symmetric vs. asymmetric algorithms. http://users.suse.com/~garloff/Writings/mutt_gpg/node3.html. 28 Aug 2000 [Online]
4. Hansche (2003) Cryptography, (ISC) 2 Press
5. Data Encryption Standard, Federal Information Processing Standard (FIPS) Publication 46, National Bureau of Standards, U.S. Department of Commerce, Washington, DC Jan 1977
6. W. P. a. C. D. E. F. F. Cracking DES: secrets of encryption research
7. Elbaz L, Bar-El H (2000) Strength assessment of encryption algorithms. website: http://www.discretix.com/PDF/Strength%20Assessment%20of%20Encryption%20Algorithms.pdf, Oct 2000
8. Zhou L, Zhou L, Bhuyan L, Xie H Architectural analysis of cryptographic applications for network processors. Department of Computer Science & Engineering, University of California, Riverside
9. Stalling W (2006) Cryptography and network security, principles and practices. Retrieved on 8 Dec 2006
10. W Comparison of ciphers, summary of algorithms [Online]
11. Elbaz L, Bar-El H (2000) Strength assessment of encryption algorithms, website: http://www.discretix.com/PDF/Strength%20Assessment%20of%20Encryption%20Algorithms.pdf, Oct 2000
12. Schneier B (1994) The blowfish encryption algorithm. Dr. Dobbs J Softw Tools 19(4):38, 40, 98, 99
13. Mansoor Ebrahim SKBK (2013) Symmetric algorithm survey: a comparative analysis. Int J Comput Appl

Increased Rate of Packets in Cognitive Radio Wireless ad hoc Network with Considering Link Capacity

Seyedeh Rezvan Sajadi

Abstract Cognitive radio is a new method helps to utilize a very valuable and natural limited resource which is named frequency spectrum. This method can learn it's surrounding, make a decision and to adapt itself with environmental conditions. One of the most important goals of the cognitive radio is to use the spectrum efficiently. In cognitive radio networks, spectrum bands between primary users (licensed users) and cognitive radio users (secondary or unlicensed users) are shared prioritized. By studying the frequency allocation in these networks, you can understand that almost all of the usable parts of this frequency bands are allocated to the primary users and apparently we are faced with a lack of bandwidth, whereas if we review the used frequency spectrum, we realize that some parts of the spectrum are unused mostly. In this article, a determination of link capacity method for intelligent radio networks is used. The proposed method helps to improve the acceptance capacity of intelligent radio networks by channel determination. The mechanism of proposed channel determination reduces the number of spectrum hand-offs which significantly improves the efficiency and the capacity of the network.

Keywords Cognitive radio (CR) · Spectrum frequency · Spectrum sensing
Spectrum hole · Primary user (PU) · Secondary user (SU)

1 Introduction

Cognitive Radio was first introduced by James Mitola (2000) in his doctoral dissertation as a promising solution to the application of spectrum dynamics [1]. Cognitive Radio can be considered as a point for the development of wireless digital personalized (PAD) devices and related networks of radio resources and

S. R. Sajadi (✉)
Master of Computer Engineering, Science and Research,
University of Fars, Shiraz, Iran
e-mail: sajadi.sh80@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_55

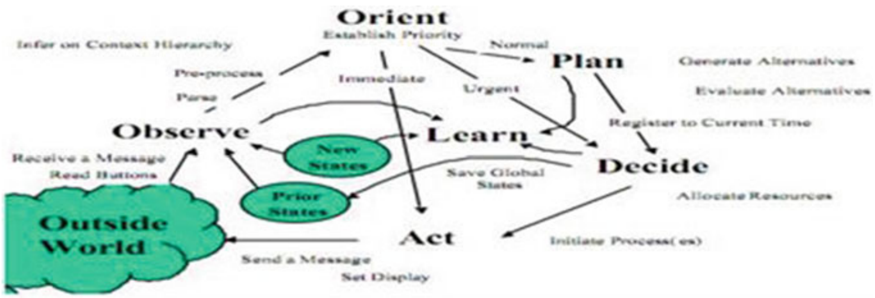


Fig. 1 Cognitive radio recognition cycle [1]

computer communications to the corresponding computer to detect user communications in order to be able to Identify communication needs as a function of field use and providing the most suitable radio and wireless resources to meet these needs. It can be said that the performance of each Cognitive Radio is like a cycle. He called this cycle a cognition cycle. Figure 1 shows the cognition cycle introduced by Mitola. This cycle consists of three main parts.

- Understanding the surrounding radio environment, this recognition includes steps such as detecting frequency cavities, estimating channel state information, predicting channel capacity for sender send.
- Management or decision making that includes spectrum management (controlling spectrum access), routing, traffic shaping, optimized forward control, and service quality review.
- Doing the operation, in this section, based on the decisions taken in the previous section, the operation is performed, for example modulation type, and so on. In fact, Cognitive Radio effectively enables the optimal use of the spectrum. As it appears from this cycle, the most important part of the Cognitive Radio is to detect and use the frequency spectrum. Therefore, using this method, it is possible to coexist the secondary use with the primary use in a spectrum frequency. Cognitive Radio is a key identifier that can improve the use of the radio spectrum through dynamic access to the spectrum without creating a disturbing interference with the original owners of the spectrum.

The existence of vacant holes in the reserved frequency band and the random use of the main users created the idea of smart networks. In Cognitive Radio networks, the user has the right to use the empty sections of the spectrum and frequency of the reservation. In fact, the Cognitive Radio user alternates the frequency spectrum and uses frequency cavities when it comes to it. When the main owners of the frequency band intend to use the reserved band, they will empty it and use another unencrypted frequency band to send their data, which improves the efficiency of the smart network. In these networks, they have provided a method for channeling dynamic channel allocation to cognitive radio users. In this way, users are prioritized, and if users with higher priority apply to the channel and there is no empty channel, the channel is taken from the user with a lower priority and given to the user with a higher priority. One of the most important ways to control interference is to increase network efficiency and provide quality requests for users.

Cognitive radio networks have been considered in recent years because of the importance of their use. These networks are equipped with transmitter and receiver devices that can change their profile and radio parameters. Recently radio cognitive networks are considered as a promising technology for the implementation of remote monitoring and control systems and other telecommunication applications in smart networks. Cognitive radio is a novel way to improve the use of a very valuable natural resource called frequency spectrum. This approach, based on learning the environment, can have an understanding of the surroundings, and one of the most important radio-target goals is spectrum availability. Smart networks are a new and promising technology to provide the connectivity and performance of these types of networks.

Easy-to-access algorithms for the spectrum for dynamic spectrum access networks that are often designed without regard to the interference of the adjacent channel. In this work, we will examine the problem of allocation of channels, capacities, and posts, while we consider this limitation.

In this paper, we will use a method for determining the link capacity for cognitive radio networks. The proposed scheme in this paper will improve the reception capacity of cognitive radio networks through channel determination. The proposed channel determination mechanism reduces the number of spectra which significantly improves the efficiency and reception capacity of the network. The proposed scheme in this study is the first plan to improve the reception capacity of cognitive radio networks in relation to frequency band constraints. It seems that considering the link capacity in allocating the spectrum efficiency in the cognitive radio network seems to be effective.

The innovative topics of the article are as follows:

- Using Link Capacity to reduce the number of spectrum handoffs.
- Using channel allocation aware of link capacity in smart networks.

2 The Issue of Considering Link Capacity

Najatian and colleagues presented an article called PUSH in 2013 [2, 3]. In this paper, they performed integrated handoff, and considered a series of parameters such as primary user, secondary user, channel heterogeneity, secondary user mobility, and primary user activity. The handoff process was performed in such a way that when the spectrum provided to primary users was temporarily not used and used by a secondary user, if the allowed primary user needed that band again, the secondary user will immediately be transferred to another bandwidth, and the band will be freed up for the first user, which is referred to as a handoff. The PUSH algorithm uses the prediction of the availability of cognitive links to estimate the maximum period of availability of links to primary users and to avoid frequency interference. But in this paper, the path that they chose was a hypothetical path in which the link capacity was not considered for channel selection and handoff. We want to consider the link capacity, in addition to the rest of the factors in this study.

2.1 System Structure

Number of channels available to the secondary user at time is C. The number of C channels to T is divided into different categories from the point of view of the transmission sufficiency, which is related to the primary user. These channels are classified into L according to their data transmission range. Each channel group is characterized by T_l that $|T_l| = C_l$ and $C = C_1 + \cdots C_L$. Regarding the primary user activity, any secondary user can access the total number of C channels in any situation. The total number of channels identified in the node is $c = c_1 + c_2 + \cdots + c_{L.}$ The transmission range of the signal of type T_l, is R_l. Consider a pair of receivers of the secondary transmitter, which, for communication, uses a channel of type T_l to exchange data, the distance between them is less than R_l. When the secondary users move and their distance from the R_1 increases, the node must change and select another channel. In this case, the new channel should have a transmission signal greater than R_1. Also, L number of primary user exists, each of which can only work on a channel of type l. As long as a primary user of type l is activated and there are no empty channels of type l, the secondary user occupies the type l channel, the secondary user must empty the channel and deliver it to the initial user.

2.2 Proposed Method

Najatian et al. presented a flowchart In the PUSH article, in which link capacity was not considered [3]. We added the link capacity to this flowchart as shown below (Fig. 2). A link is available if the two nodes communicating with the link are within the range of each other's data transfer and are outside of the interfacing area of each primary user. Also, two different transmission thresholds are defined. The first threshold of the relay to the channel area is called the SHTH[1] bandwidth threshold. The second transmission threshold is the LFH[2] threshold (LHTH),[3] which is related to the LFH area ahead. These thresholds are used to initiate the transfer due to the leaving of the node and reducing the quality of the channel. The proposed algorithm

[1] Spectrum Handoff Threshold.

[2] Local Flow Handoff.

[3] Local Handoff Threshold.

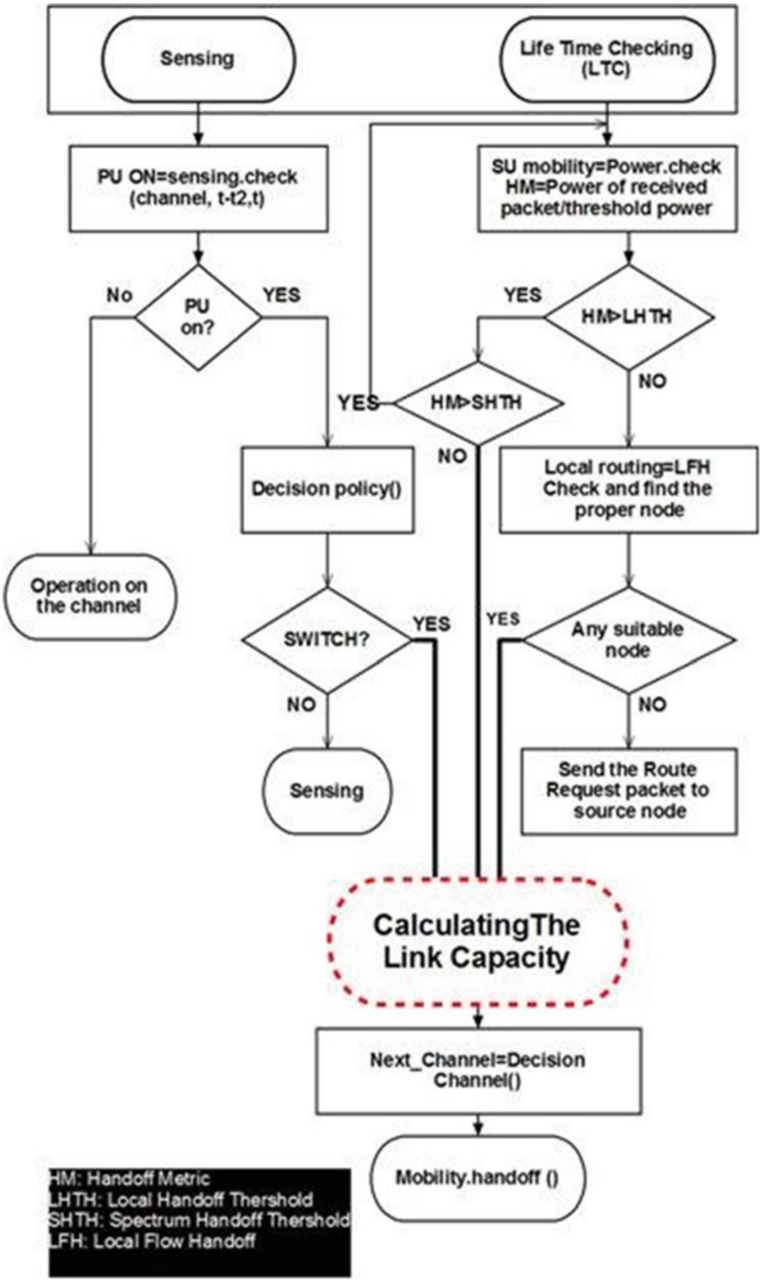

Sensing
Life Time Checking (LTC)
PU ON=sensing.check (channel, t-t2,t)
SU mobility=Power.check HM=Power of received packet/threshold power
No
PU on?
YES
YES
HM>LHTH
NO
YES
HM>SHTH
NO
Decision policy()
Local routing=LFH Check and find the proper node
Operation on the channel
SWITCH?
YES
YES
Any suitable node
NO
NO
Sensing
Send the Route Request packet to source node
CalculatingThe Link Capacity
Next_Channel=Decision Channel()
Mobility.handoff ()
HM: Handoff Metric
LHTH: Local Handoff Thershold
SHTH: Spectrum Handoff Thershold
LFH: Local Flow Handoff

Fig. 2 Integrated mobility and handoff management algorithm [3]

for the PUSH scheme is also shown in the following pseudo-code. The decision-making unit starts the transfer of the band based on the threshold of transfer. When the band transfer cannot be performed in a hop, LFH runs to communicate. Table 1 shows abbreviations and meanings that are used in pseudo-code.

```
1   check the HMF;
2    if HMF := 1
3              go to 46; // for checking the HM
4   end if
5   if there is no data to send
6           go to 2; // for checking the HMF
7   end if
8   start scanning radio;
9   $PU_{z^*}$ ON_ = Sense (channel, $t_s$, switching);
10   if $PU_{z^*}$ ON
11          go to the Decision. Policy;
12    if switching must be done
13          defer for Channel_Switching_Delay;
14          for Next_Channel Decision go to 28;
15          Calculating The Link Capacity;
16          CSF = 1;
17          go back to 5; // to check the availability of data for sending
18    end if
19  end if
20  if (CSF)
21          send notification to upper layers for spectrum handoff;
22          adapt channel parameters;
23          CSF:= 0 and go to 25;
24  end if
25  start operation timer ($T_0$);
26  transmit data till $T_0$ expires;
27  go back to 2;
28  for k := 1, kcandidate intermediate node;
29    if $PU_k$ ≠ ON
30          NAC:= NAC +1;
31          LAC (NAC) := k;
32    end if
```

Table 1 Abbreviations used in the pseudo code and their definitions

Symbol	Definition
CSF	Channel switching flag
LAC	List of available and detected channels
HMF	Handoff metric flag
T_V	Operating time
T_X	Sensing time
NAC	Available channel number (Channel ID)
RERR	Rout error request
CSR	Channel switching request
CSA	Channel switching acknowledgment
HM	Handoff metric

```
33  end for
34  if LAC := ∅
35            send RERR packet to the source node;
36    else if LAC ≠ ∅
37          sending CSR; // sending PU-HREQ
38          go to 40;
39  end if
40  upon receiving CSA then
41  if CSA ≠ ∅
42          switch to the selected channel;
43          HMF:= 0;
44        go to 16;
45  end if
46  calculate the HM;
47  if HM ≤ LHTH
48            Start the LFHREQ timer, broadcast the LFHREQ and go to 52;
49    else if HM ≤ SHTH
50          go to 28; // for making a decision on the next channel
51  end if
52  upon the LFHREQ timer was expired then
53  if LFHREQ:= ∅
54              broadcast the RERR to the source node;
55     else if LFHREQ ≠ ∅
56          find the best candidate with the max TPU;
57          send the local route HR to the local source through the best
58          send the CSR to the intermediate node;
59          go back to 40;
60  end if
```

3 Link Capacity

To enable us to measure link capacity, we use the paper presented by Nejatian and his colleagues entitled Interference Aware Channel Assignment(IACA) for Cognitive Wireless Mesh Networks [4]. In this paper, the concept of awareness of end-to-end frequency interference is presented at the end of the CWMN radio cognitive wireless mesh network. And it has been shown that the end-to-end interference model has the ability to perform better than the SINR basic model. In order to increase the efficiency of the unused spectrum, the channel selection strategy must have a vigilance mechanism to prevent frequency interference.

In this work, we show that the end-to-end interference model delivers the ability to achieve the goal of resource optimization in CWMN. We assume the following scenario: a CWMN is composed of Mesh Router (MR), Mesh Gateway (MG), Mesh Clint (MC)/Secondary User (SU) and Primary User (PU) (Table 2).

- The source node denoted as S, destination node as D.
- Moreover, $S, D \in A, L_{\mathrm{i}} \subseteq L_j$ because j cannot use a channel that is not available to i node.
- Let's $\mathrm{k} \in \mathrm{K} \in \mathrm{L_i} \in \mathrm{L_j}, \mathrm{M} \notin \mathrm{A}, \mathrm{m_j} \in \mathrm{M}, \mathrm{Q} \notin \mathrm{A}, \mathrm{q_j} \in \mathrm{Q}, \mathrm{R_i} = 2\mathrm{R_T}$.
- PU is randomly active at the location of all nodes in the different time slots.
- Each node in this system has more than one neighboring nodes and they are randomly active for using the same channel.

In order to take into account, the interference in the proposed end to end model, CR-MANETS is modeled as a bidirectional graph using graph theory. G = (V, E) where V is the vertices representing a set of nodes and E is edges that represents a set of wireless links. To develop end to end interference model in a system, $\mathrm{V_c}$ is the set of the corresponding node in the network.

$$\mathrm{V_c} = \left\{\mathrm{Link_{ij}} \quad \text{in a communication link}\right\} \tag{1}$$

Table 2 Notations

Symbol	Definition
A	The set of nodes in the system
$\mathrm{Link_{ij}}$	Link exists for any pair of nodes i, j ∈ A
L	The set of available channels in system
$\mathrm{L_i}$	The set of available channel at node i
K	The set of common channels of nodes
M	The set of neighboring SUs from which channel is sharing.
Q	The set of neighboring SUs from which node j can hear (or sense) a packet but not sharing channel
$\mathrm{R_T}$	Transmission distance
$\mathrm{R_i}$	Interference range

where Link_{ij} developed if node i and node j have the one same channel in their channel list. A link is considered available if the two nodes associated with the link are within the transmission range of each other and out of interference region of any PU [5]. The connectivity graph shows conflict in which one vertex is linked to the other vertices having same channel availability, where these two vertices are linked to one vertex. In this scenario the corresponding links are interfaced to each other and present the conflict graph to produce interference in wireless links. If the neighboring node is active at that time slot and starts sharing the same channel, then it will create co-channel interference at particular link.

3.1 End-to-End Interference Model

In this section, we start from a physical interference model and then extended it by adding PU activity and SU activity as additional features. As shown, when the interference on the link is minimum then the capacity of the link is maximized. To account this, we are considering three key parameters. Link interference ratio IR_{T}, PU activity and SU activity. Link between node i and node j is established if the SINR at the receiver node is above the threshold limit. This communication transmission depends on the required communication parameters, which are channel, data rate etc. Furthermore, indicating the transmission strength of a packet from node j at node i referred as P (link_{ij}) signal strength on the link_{ij}. This link is referred as Link_{ij}. The packet from node i to node j is delivered if

$$\frac{P(Link_{ij})}{N+\sum_{m_j\in\int}\tau(m)P_i(m)+\sum_{q_j\in Q}P_j(s)}\geq SINR_T \tag{2}$$

where N represents the background noise, M is the set of neighboring SUs $M\notin A$ and m ∈ M denotes the set of nodes from which node j can hear (or sense) a packet. $\tau(m)$ gives the fraction of time node m occupies the channel, Pi(m) is used to weight the signal strength of interfering node m, S is the set neighboring SUs on same channel but not share the channel, s_{j} is referred as node which is an interference range of node j. The data rate, channel characteristics and modulation scheme are the main parameters for calculating SINR. By using SNR and SINR the Interference ratio (IR) is defined in [6]. Link from node i to node j is denoted Link_{ij}. IR_{ij}(i) for a node i in a Link_{ij} = (i, j).

Where $\left(0<IR_i(Link_{ij})\geq 1\right)$ as follows:

$$IR_i(Link_{ij})=\frac{SINR_i(Link_{ij})}{SNR_i(Link_{ij})} \tag{3}$$

where

$$SINR_i(Link_{ij}) = \frac{P(Link_{ij})}{Noise} \tag{4}$$

$$SINR_i(Link_{ij}) = \frac{P_j(i)}{Noise + \sum_{m\in M(j)-i} \tau(m)P_i(m) + \sum_{sj\in S} P_j(s)} \tag{5}$$

Total interference ratio (IR_T) for first hope (h1) to last hope (hN) defined as

$$IR_T = IR_{ijh1} + IR_{ijh2} + IR_{ijh3} + \cdots IR_{ijhN} \tag{6}$$

$Link_{il}$=(i, l) exists as conflict link availability in communicating nodes, where i, j, l $\in$ A. In the next section we explain the steps to avoid the conflict in start of initial communication process. The Bernoulli random variable with binomial distribution is used for each available channel to obtain the preferred channel where PU activity is least on the $Link_{ij}$.

- Let's assume that attempts of PU are a Bernoulli random variable with "e" defined as an event when PU is active.
- Let's take the assumption that the p_{PU} (probability of active PU) is denoted by p.
- Total number of attempts is n.

$$p_i(e) = \binom{n}{e} p^e (1-p)^{n-e} \tag{7}$$

The event that a PU attempts at a particular node is modeled as a Bernoulli random variable, in which an active PU is defined as a 1, with probability 'p', and a non-active PU is defined as a '0', with probability '1-p '.

3.2 Interference Aware Channel Assignment (IACA) Algorithm

These three parameters are fed into the weightage mechanism, which assigns different weights to each of the parameters. Based on this weightage, the channel is selected. Figure 3 explains in implementation steps. We define $W_{ij}(k)$ as weightage for all available channels at initiating node

$$W_{ij}(k) = p_{PU}(i,j,k) + p_{SU}(i,j,k) + IR_{IJ}(k), \quad \forall i,j \in A \tag{8}$$

by adding different priorities (α, β, γ) to each parameter for achieving optimum results.

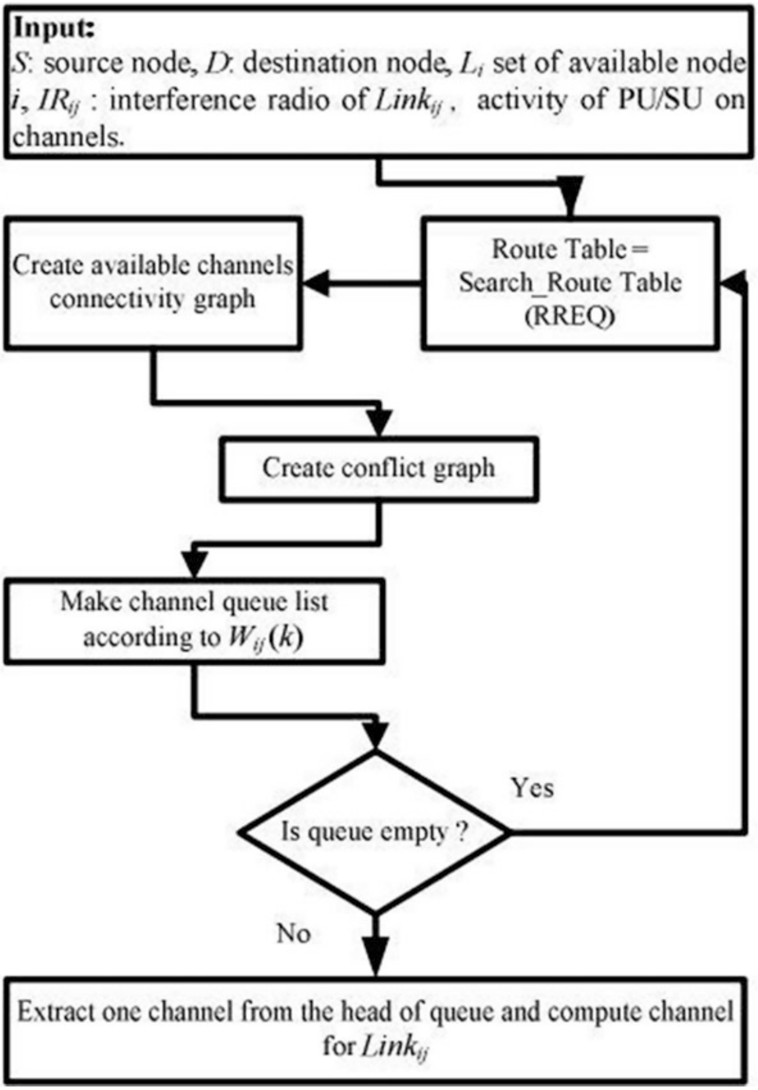

Input:
S: source node, D: destination node, L_i set of available node i, IR_{ij} : interference radio of $Link_{ij}$, activity of PU/SU on channels.
Route Table = Search_Route Table (RREQ)
Create available channels connectivity graph
Create conflict graph
Make channel queue list according to $W_{ij}(k)$
Is queue empty ?
Yes
No
Extract one channel from the head of queue and compute channel for $Link_{ij}$

Fig. 3 IACA algorithm [4]

$$W_{ij}(k) = \alpha p_{PU}(i,j,k) + \beta p_{SU}(i,j,k) + \gamma IR_{IJ}(k), \quad \forall i,j \in A \tag{9}$$

where L_j denotes the list of available channels for node j. Then, each route has a corresponding available channel set, represented by L_1, L_2, … , L_A, where A is the number of nodes on the route. In order for a route to be valid, every pair of neighboring nodes on the route has at least one common channel available to both nodes. Node i select an available channel k_1 with the smallest weightage level value form set K, that is,

$$k_1 = argmin\ W_{ij}(k) \tag{10}$$

The route request propagation process uses the Route Request (RREQ) same as broadcast in AODV. Through all available channels the route discovery request (RREQ) is broadcasted by initiating source node S towards its neighboring nodes. In the initial handshake phase neighbors' operating channel are obtained. S Node receives the channel's parameters from the neighboring nodes that include channels IRij, P_{PU} activity and P_{SU} activity. The minimum weighted channel is selected.

4 Simulation and Results

Here is a path from the origin (S) to the destination (D) that has already been selected to send the data (Fig. 4).

An algorithm operates for identifying the appropriate spectrum bands, based on channel quality, spectrum, and the mobility of secondary users. Therefore, a precise and knowledgeable positioning mechanism is essential. In the decision of the spectrum handoff, the accuracy of the time availability of the link and the channel should be considered.

4.1 Simulation Scenarios

In this section, three different scenarios will be compared using the Network Simulator software (NS-2). Three different scenarios are considered to examine the

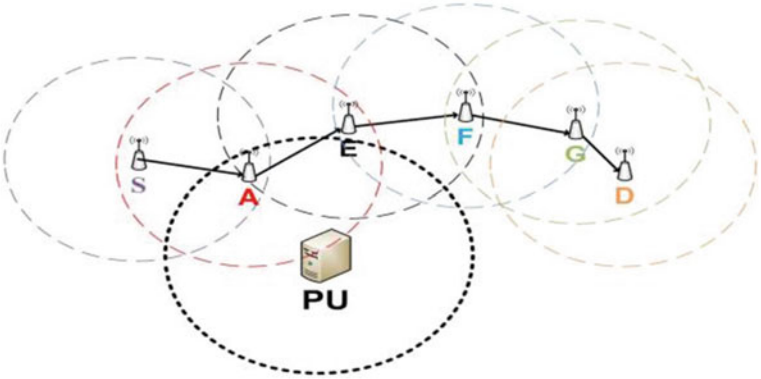


Fig. 4 Data sending path [3]

possibility of blocking the handoff. These three different scenarios of handoff management include: Unified Spectrum Handoff (USH) management system, Proactive Unified Spectrum Handoff (PUSH) management system, and Proactive Unified Spectrum Handoff management system with considering link capacity (LCC-PUSH). The first two scenarios are the scenarios presented by Najatian et al. [2, 4], and the third scenario is related to this research. In all three scenarios, unified spectrum handoff management is considered. In these scenarios, the local handoff flow is considered, if needed, in the management system. The first system does not consider handoff threshold while the PUSH and LCC-PUSH systems consider the handoff stand. In system LCC-PUSH The link capacity is considered based on the method presented in Chap. 3.

The total number of channels is C = 10 channels, which are divided into two categories of channels with different transmission rates C1 = 5 and C2 = 5. The transmission range of these two channel categories is R1 = 75 m and R2 = 125 m. Secondary mobile users in a dimensional network 2000 m × 2000 m are randomly distributed at a speed of 3 m/s. The transmission of the primary users whose stationary location is constant is considered 200 m. The activity of these users is modeled as a two-state turn-on/off-state process. Also, two different threshold values for handoff are defined.

4.2 Handoff Blocking Probability

In this section, the probability of maintaining a secondary user path is studied. The handoff threshold is 6 s and the entrance rate for both primary and secondary users is 0.25. Also, the AODV routing protocol is used to form a path in the cognitive radio.

Figure 5 shows the handoff blocking probability in three different scenarios. Based on the results, the proposed LCC-PUSH method improves the probability of maintaining the path to the first scenario, while it has a greater chance of blocking the handoff than the PUSH. The reason for this is that, considering that the LCC-PUSH takes into account the link capacity, and if the link capacity is low, the handoff is not carried out, and as a result, a number of empty channels are unused as if the number of channels is reduced. And the chance of a successful Handoff will be reduced.

4.3 Throughput

Figure 6 shows the delivery rate of the secondary user package versus the number of secondary users. Here, the user's secondary packet rate is 200 pps. The initial packet send rate is 12 packets per second. It is clear that the packet delivery rate is

Fig. 5 Handoff blocking probability in three different scenarios

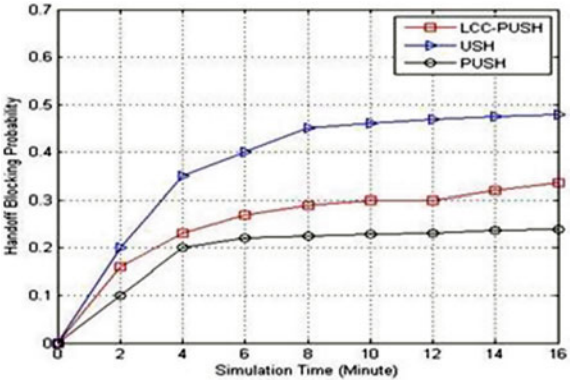
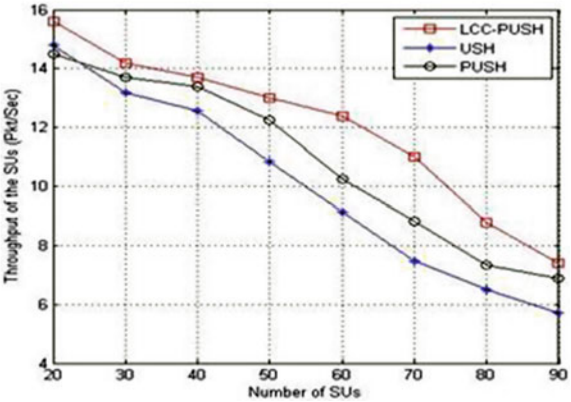


Fig. 6 Secondary packet delivery rate versus second user number



decreasing with the increase in the number of secondary users. The reason is that a larger number of secondary users reduces the chances of access to empty channels. On the other hand, as shown in the figure, the LCC-PUSH method improves packet delivery rates.

Figure 7 shows the delivery rate of the secondary user packet against the number of primary user channels. Here, the primary and secondary user login rates are equal to the values in Fig. 6. The number of primary user channels varies and is divided into two groups with equal numbers and different transmission rates for the secondary user. The number of secondary users is fixed to 50. When the number of primary user channels increases, the delivery rate of the secondary packet will increase, which is because the chances of accessing empty channels to the secondary user increase. In this figure, the LCC-PUSH method improves packet delivery rates.

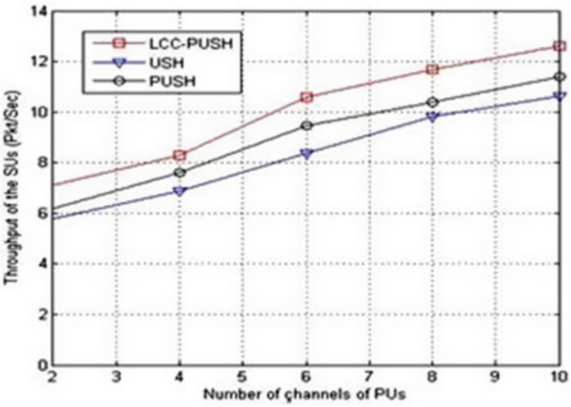


Fig. 7 The delivery rate of the secondary packet versus the number of primary user channels

5 Conclusion

Given the simulation results, we found that the proposed method has a greater chance of blocking the index. This is due to the fact that this method takes into account the capacity of the link, and if the link capacity is low, the handoff is not done and the chances of success are less. So, we conclude that the chance of blocking the handoff against the front is greater, but we have a larger packet rate than PUSH. This is because after the handoff has been done, given that the link capacity is high, we can send packet at a higher rate and there is no barrier. According to the results, we conclude that the proposed LCC-PUSH method works better than other methods in terms of packet sending rate, which is because we consider the link capacity. This taking into account the link capacity leads to a delay, but ultimately, given the fact that we are considering a higher-capacity link, we have an increase in packet sending rates. So the way we've presented a good and effective method.

6 Suggestions

Considering that the proposed method increases the packet rate but, on the other hand, increases the blockade of the handoff, it is suggested to use other methods to combine with this method to solve this problem and simultaneously reduce the delay. Because if the delay is reduced, the packet sending rate will increase much more.

References

1. Mitola J (2000) Cognitive radio: an integrated agent architecture for software defined radio. Ph. D. dissertation, KTH Royal Institute of Technology, Stockholm, Sweden
2. Nejatian S, Syed Yousof SK, Abdol Lattif NM, Asadpour V (2011) Proactive integrated handoff management in CR-MANETS: a conceptual model. In: IEEE symposium on wireless technology and applications (ISWTA), Bandung, 2011, pp 33–3
3. Nejatian S, Syed-Yusof SK, Abdul Latiff NM, Asadpour V, Hosseini H (2013) Proactive integrated handoff management in cognitive radio mobile ad hoc networks. EURASIP J Wirel Commun Netw 2013:224
4. Maqbool W, Syed Yusof SK, Abdul Latiff NM, Hafizah S, Nejatian S, Farzamnia A, Zubair S (2013) Interference aware channel assignment (IACA) for cognitive wireless mesh networks. In: 2013 IEEE 11th Malaysia international conference on communications 26–28 Nov 2013, Kuala Lumpur, Malaysia
5. Chen R, Park M, Reed JH (2008) Defense against primary user emulation attacks in cognitive radio networks. IEEE J Select Areas Commun Spec Issue Cogn Radio Theory Appl 26(1) (Jan 2008)
6. Yucek T, Arsalan H (2009) A survey of spectrum sensing algorithms for cognitive radio application. IEEE Commun Surveys Tutorials 11(1) (First Quarter 2009)

Deadlock Detection in Routing of Interconnection Networks Using Blocked Channel Fuzzy Method and Traffic Average in Input and Output Channels

Maryam Poornajaf

Abstract One of the most important issues in parallel processing is routing message from a source node to destination that is done by routing nodes. Deadlock in routing is a damaging factor that arises because of dependence of buffers and routing channel dependency cycles as insoluble one. Many routing algorithms to confronting deadlocks use prevent methods and deadlock avoidance. In these methods sources aren't used in optimal forms, because in these methods there are some limitations in routing and using sources. But in many algorithms of routing the method of deadlock detection and deadlock recovery is used because of in appropriate use of sources in prevention method. In this paper deadlock detection method is indicated based on traffic average in input and output channels in each node and using fuzzy techniques. And it tries reducing the messages that are introduced wrongly as deadlock involving messages and as much as possible in real deadlock, less messages as involving deadlock been introduced to deadlock recovery.

Keywords Deadlock · Routing · Virtual channel · Threshold

1 Introduction

Two types of parallel computers are multi computer and multiprocessors. Parallel process in multicomputer systems is using passing message, but in multiprocessors systems, processors are communicated by shared memory. In parallel architectures, interconnection networks facilitate the connection between network factors, and it lead to direct messages to destination. Interconnection network was classified based on various architectures. Their topologies are as two classes of dynamic and static

M. Poornajaf (✉)
Technical and Vocational University, Ilam, Iran
e-mail: m_poornajaf1981@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_56

[1–3]. In a static network, its nodes are connected as point to point are fixed in a static topology. Some types of topology in static network are: ring—binary tree—mesh—hyper cube—K—ray n cube and linear [4]. But a dynamic network allowed connections to change in network nodes as dynamic, so switch is used [3]. Some dynamic networks are: buss networks—crossbar and multilayer networks.

One common approach of routing in multicomputer is wormhole routing [4, 5]. In wormhole routing a packet is transferred among nodes called Flit, Flit is the smallest unit of message. Header Flit of message is containing routing information and other Flits that are after it containing data. Since just header flits have routing information, all other Flits are after it and move in pipeline [4, 6, 7]. When header Flit reaches a router, all packet Flits are wait to free an appropriate output channel, if there isn't an appropriate free output channel. In this method, any router needs enough space of buffer in each channel to storing some Flits, and it doesn't need to buffer with high capacity [8, 9]. In this method of storing, routing and forward would do in same times, so time of transfer of packet is decreased, because each Flit doesn't wait to other one that is in its behind [10]. Wormhole routing decreases packet delay in compare to method of store and forward [6, 7].

Routing can be deterministic or adaptive. In deterministic routing, connection route between source and destination nodes is fixed, even there are several routes. But its flexibility is not sufficient to adapt with network conditions such as congestion and destruction. Adaption is an important feature that provides alternative routes to packets that confront to defect hardware, blocked channels and so on [6, 10, 11]. Deterministic routing prevents effective use of physical connections, because physical channels allocated inflexible.

Routing by method of shortest path, direct packet in shortest path and so fault tolerance is low [4]. But in none short routing, chosen route may not be shortest one and it provide better fault tolerance. The best choosing route is dependent on network load. To reach good performance over all traffic loads, routing algorithm must be adaptive, and chooses routes based on network conditions. More flexibility in adaptive routers under same load and none same load can improve the performance.

In wormhole routing it may occur a deadlock event [12]. Indeed because of channels and buffers dependency, deadlock in interconnection networks is determine as dependency cycle, and in this dependency cycle each deadlock message, demands other sources that are in control of other messages of cycle, and this performance leads to an inevitable dependency cycle in routing [6, 12, 13]. Deadlock is different from message blocking. It may some packets been blocked without any deadlock, but if this deadlock occurred it must block all packets. Messages blocking is recovery by slowing traffic, but deadlock cycle is not recovered by slowing traffic, and only way to recovering it is break down the dependency deadlock cycle.

There are some methods to confronting to deadlock such as prevention and avoidance that sources in those are not used as optimal. In these methods, routing was done as none adaptive and may partially adaptive, and creates some limitations in networks to prevent from deadlock [11, 14–16].

But detection methods and recovering deadlock are not in this way, routing in detection methods and recovering deadlock are done as adaptive and it allowed to deadlock [14, 17, 18]. In these methods, involving messages with deadlock are identifying and they reported, and then related cycle to this deadlock is breakdown to deadlock recovering. Most of existence multicomputer networks, using deterministic routing, although there were several routes between source and destination, but to avoidance from deadlock, deterministic routing determines a route from source to destination [17, 19].

2 A Fuzzy Method to Deadlock Detection

There are many methods to deadlock detection in interconnection network. A base method to deadlock detection is such that if a router couldn't rout a message (when to routing message appropriate output channels aren't free), and if all appropriate output channels were blocked more than a determine time out, then that message is reported as an involving deadlock message (Blocked channel method) [1, 6, 11, 14, 20]. But choosing threshold (time out) is an important factor, in that if threshold considered as low, it is possible a message that is not deadlock involve be introduced as involve one. And if threshold be considered as high, it is possible that involve message be detection more lately. Time out of deadlock detection not is constant in any traffic and network load [21, 22]. In this paper threshold of deadlock detection is determined based on network traffic as a fuzzy form. In this paper also traffic is computed based on output channels and input channels, in contrast to other algorithms that calculated traffic based on output channels. In this paper, since busy input channels of each node are effective in traffic increasing, and if deadlock occurred so it can be said that input channel is a part of deadlock dependence cycle, the input channel positions was investigated in traffic computation. By increasing network traffic in various dimensions, the probability of channels blocking and packets blocking is increased, and because of it, threshold must be high and it prevent instead of blocked packets to deadlock involve packets. By decreasing network traffic and reducing packets and blocked channels threshold be low, because probability of block in packets and their mistake identify is low instead of packets involve deadlock.

To fuzzing threshold of deadlock detection, it is necessary that determine high and low level of threshold and it performed to determine amount of threshold using fuzzy method and it leads to this that threshold amount that determine by fuzzy method is not be out of control. Primary amount of threshold must be determined in each node, and its amount must be determined in low and high range and threshold amount be always in the range. In each pulse, virtual output channels and virtual input channels are studied. By studying free channels and busy on in output channels and input channels, we can calculate network traffic as local and approximately in each node. And if this application be implementing in a time range, we can calculate traffic as an average of traffics in pulses of that time range.

An input variable of fuzzy is used to show the different amount of past and present traffic and include some language value series and also an output variable of fuzzy known as Offset is considered to restitution amount of threshold that is include some language value series. To prevent the removed of past role in any range of time, traffic in each node is accounted as traffic average in past and current one. There is a series of fuzzy roles if—then that indicate traffic changes in threshold as language variables. For example, if amount of traffic be increase in various dimensions, then threshold or time out must increase and it prevent from replacing mistake identification of common messages with deadlock ones, so offset must be positive. Output fuzzy membership functions must identify as rectangular to input fuzzy parameter.

Fuzzy deadlock detection mechanism is performing in each Δt. we indicate it based on pulse. While number of input and output channels of busy and free in each pulse are computed and after each Δt, resulting traffic based on total channels of busy input and output channels in time range of Δt is reported. Network traffic is computed based on average of reporting traffic in Δt and past traffic. This action has not any effect on current decisions and is used to determine past traffic in next performance of mechanism of fuzzy detection in next Δt.

To defuzzification and accept appropriate Offset, it uses method of defuzzification of center of gravity. The offset is used to achieve amount of new threshold. (New threshold = offset + previous threshold).

In deadlock detection method in this article lower number of messages are introduce as deadlock involve messages, because deadlock cycle is recovered by extract one deadlock involve message and it is not necessary that in each node a message or several messages be introduced as deadlock involve message. Indeed, we try found the deadlock place and also we try introduce source of cycle creation by continue the deadlock cycle of dependency. This is feasible by choose an appropriate threshold to time of channel blocking and deadlock involving messages.

3 Network Pattern

In each node, shortest path is choosing to take a packet to destination. Wormhole switching is used to routing. Each packet is divided to some flits. Header flit is routing in each node and other flits are behind it. Routing unit in each pulse, routes a packet, indeed regarding to destination chose an appropriate output channel to packet. Each physical channel includes three virtual channels. Routing unit can choose each of virtual channels relate to physical channel. Politic of chose of virtual channel to use bandwidth of physical channel is Round Robin. In each node a queue is considered to packets that are waiting to routing, packets that are wait in this queue as Round Robin. In each pulse just one packet is routing and that is packet in begin of queue, but if there is not appropriate output channel routing to it, it was send to end of queue to be routing in next round.

It considered a buffer to store flits in output and input virtual physical channels and their capacity is 4 flits. Choosing politic to of packet to using output and input buffers is FIFO. In each pulse just one flit can pass from physical channel. In each node a crossbar switch is used to connection between output and input channels. In each pulse, crossbar switch allow several connects in output and input channels and connections are not effect on each other. During packet routing connection of input channel is performed to output channel, and this remained till packet cross to related switch and other packet can't to cross to related channel. To injection, there are four injection channels, in each pulse, each node produces some packet, if there were some free injection channels. If not, packet must wait in productive node to inject in other time.

4 Simulation Results

Simulation was performed on 2 dimensional mesh network with 3 virtual channels in each physical channel and in 20000 pulse. Traffic pattern of network is as same distribution, and in it each node can send a message to other nodes and can receive message from other nodes. Results of simulations are indicated packets as introduced deadlock packets mistakenly. Results of simulations to threshold considered 4 pulses—8 pulse—16 pulse and 32 pulse, dynamic threshold consider as 16–32 pulse and 16–48 and 24–48 to different traffics. Traffic is considered based on average of flits that receive in each node in each pulse. As it shows in figure, numbers of packets that mistakenly introduce as deadlock in Fig. 2 are lower than Fig. 1. And it is because of study both of input and output channels in traffic computations.

5 Conclusion

Deadlock is a phenomenon that can't be distinguished as certainly. It can be distinguishing deadlock by considering network deadlock. In this paper, deadlock is distinguished by fuzzy method and it be tried that messages that report as involve deadlock mistakenly were reduced. Regarding to simulation results, if low and high level of threshold be determined appropriately, incorrect determine of deadlock is reduced in fuzzy method threshold will change as fuzzy by network traffic changes. Therefore, threshold has any specific dependent to message length. But threshold choice is an important problem, if low level be considered to threshold, it may message that isn't involve deadlock, introduce involve deadlock message mistakenly, and if high level be considered to threshold, it may that involve deadlock message be detected lately, so channels been blocked in more time, and it leads to delaying in messages. Threshold is change based on network changes and its amount is determined as fuzzy by considering network traffic. Using fuzzy methods

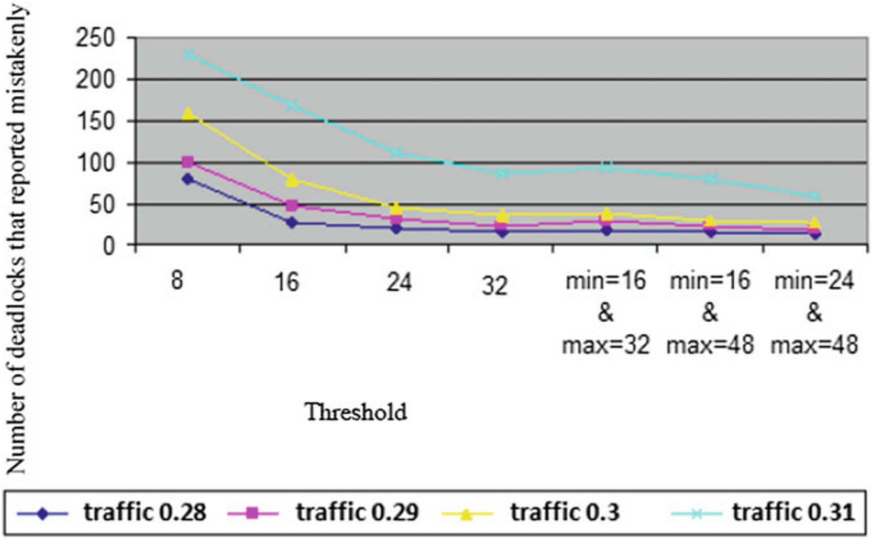


Fig. 1 Number of messages that introduce as deadlock, in a method that use extra blocked channels to deadlock detection (In this figure traffic is based on output channels)

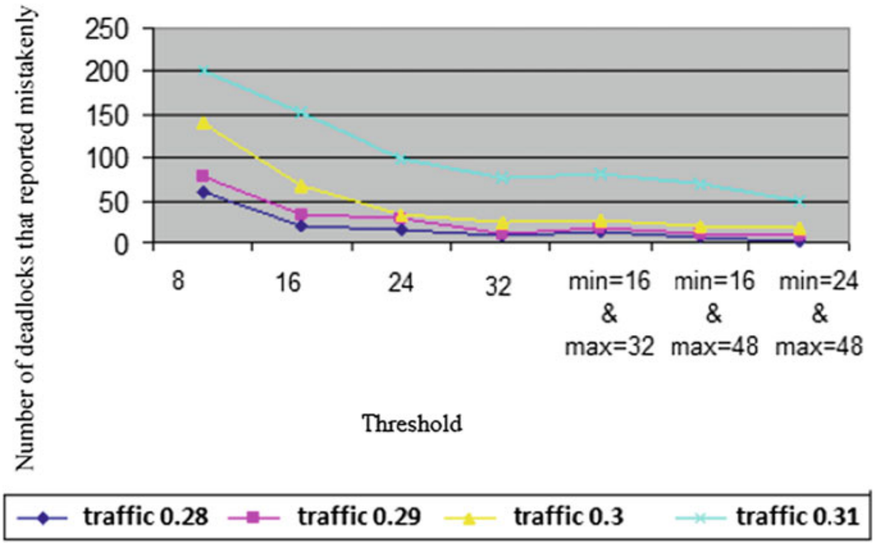


Fig. 2 Number of messages that introduce as deadlock, in a method that use extra blocked channels to deadlock detection (In this figure traffic is based on output and input channels)

in compare to classic methods is include better results in routing and deadlock detection, and it can be used fuzzy method to interconnection network. By using deadlock detection methods and by considering fuzzy laws, it can be decreased messages that introduced as involve deadlock message mistakenly.

References

1. Diatom J et al (2003) Interconnection networks: an engineering approach. Morgan Kaufmann Publishers. 600p. (In San Francisco, CA, USA). ISBN 1558608524
2. Zargham M (1996) Computer architecture: single and parallel system. Prentice-Hall International. 472p (In Upper Saddle River, NJ, USA). ISBN 0135294975
3. El-Rewini H et al (2005) Advanced computer architecture and parallel processing. Wiley, New York. 287p. ISBN 0-471-46740-5
4. Glass CJ, Ni LM (1991) The turn model for adaptive routing. In: Proceedings of the 19th Annual International Symposium on Computer Architecture (ISCA), Queensland, Australia, 19, 21 May 1992, pp 278–287
5. Dally WJ et al (2004) Principles and practices of interconnection networks. Morgan Kaufmann Publishers. 550p. (In San Francisco, CA, USA). ISBN 9780080497808
6. Martinez JM, Lopez P, Duato J (2003) FC3D: flow control-based distributed deadlock detection mechanism for true fully adaptive routing in wormhole networks. IEEE Trans Parallel Distrib Syst 14(8):765–779
7. Dally WJ, Seitz CL (1987) Deadlock-free message routing in multiprocessor interconnection networks. IEEE Trans Comput C-36:547–553
8. Kim J, Liu Z, Chien A (1997) Compressionless routing: a framework for adaptive and fault-tolerant routing. IEEE Trans Parallel Distrib Syst 8:229–244
9. Duato J (1993) A new theory of deadlock-free adaptive routing in wormhole networks. IEEE Trans Parallel Distrib Syst (TPDS) 4:1320–1331
10. Mohapatra P (1998) Wormhole routing techniques for directly connected multicomputer systems. ACM Comput Surv 30:374–410
11. Sharifian-Nia S, Vafaei A, Shahimohamadi H (2012) Deadlock recovery technique in bus enhanced NOC architecture. Int J VLSI Des Commun Syst (VLSICS), 3(4)
12. Dally WJ, Aoki H (1993) Deadlock-free adaptive routing in multicomputer networks using virtual channels. IEEE Trans Parallel Distrib Syst 4:466–475
13. Taktak S, Desbarbieux JI, Encrenaz E (2008) A tool for automatic detection of deadlock in wormhole networks on chip. ACM Trans Des Autom Electron Syst 13(6):1–22
14. Soojung L (2009) Deadlock detection and recovery for true fully adaptive routing in regular wormhole networks. J Inf Sci Eng 25:465–479
15. Soojung L (2007) A deadlock detection mechanism for true fully adaptive routing in regular wormhole networks. Comput Commun 30:1826–1840
16. Anjan KV, Pinkston T (1995) DISHA: a deadlock recovery scheme for fully adaptive routing. Proceeding of Ninth Int'l Parallel Processing Symposium, Santa Barbara, CA, USA, 25, 28 April 1995, pp 537–543
17. Khonsari A, Shahrabi A, Ould-Khaoua M, Sarbazi-Azad H (2003) Performance comparison of deadlock recovery and deadlock avoidance routing algorithms in wormhole-switched networks. IEEE Proc Comput Digital Tech 150:97–106
18. Soojung L (2006) Turn-based deadlock detection for wormhole routed networks. CIT'06 Proceedings of the Sixth IEEE International Conference on Computer and Information Technology, Seoul, Korea, 20, 22 September 2006

19. Mirza-Aghatabar M, Tavakol A, Sarbazi-Azad H, Nayebi A (2008) An adaptive software-based deadlock recovery technique. Proceedings of 22nd International Conference on Advanced Networking and Applications-Workshops, Okinawa, Japan, 25, 28 March 2008, pp 514–519
20. Al-Dujaily R, Mak T, Xia F, Yakovlev A, Palesi M (2012) Embedded transitive closure network for run-time deadlock detection in networks-on-chip. J IEEE Trans Parallel Distrib Syst 23(7):1205–1215
21. Anjan KV, Pinkston TM (1995) An efficient fully adaptive deadlock recovery scheme: DISHA. International Symposium On Computer Architecture, Santa Margherita Ligure, Italy, 22, 24 June 1995, pp 201–210
22. Anjan KV, Pinkston TM, Duato J (1996) Generalized theory for deadlock-free adaptive wormhole routing and its application to Disha concurrent. Proceedings of 10th Int'l Parallel Processing Symposium, Honolulu, HI, USA, 15, 19 April 1996, pp 815–821

Optimizing of Deadlock Detection Methods in Routing of Multicomputer Networks by Fuzzy Here Techniques

Maryam Poornajaf

Abstract One of the most important issues in parallel processing is routing message from a source node to destination that is done by routing nodes. Deadlock in routing is a damaging phenomenon that arises because of dependency of buffers and routing channels as insoluble dependency cycle. Therefor routing algorithms, which are inevitability different phenomenon, costs are incurred. Many routing algorithms to confronting deadlocks use prevention methods and deadlock avoidance. In these methods sources aren't used in optimal forms, because in these methods there are some limitations in routing and using sources. But in many algorithms of routing the method of deadlock detection and deadlock recovery is used because of inappropriate use of sources in prevention method. In this paper, a new method to detect deadlocks using fuzzy techniques is proposed that this method is combination of deadlock detection methods. And it tries reducing the messages that are introduced wrongly as deadlock involving messages.

Keywords Deadlock · Routing · Virtual channel · Threshold

1 Introduction

Two types of parallel computers are multicomputer and multiprocessors. Parallel process in multicomputer systems is using passing message, but in multiprocessors systems, processors are communicated by shared memory. In parallel architectures, interconnection networks facilitate the connection between network factors, and it lead to direct messages to destination. Interconnection network was classified based on various architectures. Their topologies are as two classes of dynamic and static [1–3]. In a static network, its nodes are connected as point to point are fixed in a static topology. Some types of topology in static network are: ring—binary tree—mesh—hyper cube—K—ray n cube and linear [4]. But a dynamic network allowed

M. Poornajaf (✉)
Technical and Vocational University, Ilam, Iran
e-mail: m_poornajaf1981@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_57

connections to change in network nodes as dynamic, so switch is used [3]. Some dynamic networks are: buss networks—crossbar and multilayer networks. One common approach of routing in multicomputer is wormhole routing [4, 5]. In wormhole routing a packet is transferred among nodes called Flit, Flit is the smallest unit of message. Header Flit of message is containing routing information and other Flits that are after it containing data. All flits move in pipeline [6, 7]. When header Flit reaches a router, all packet Flits are wait to free an appropriate output channel, if there isn't an appropriate free output channel. In this method, any router needs enough space of buffer in each channel to storing some Flits, and it doesn't need to buffer with high capacity [8, 9]. In this method of storing, routing and forward would do in same times, so time of transfer of packet is decreased, because each Flit doesn't wait to other one that is in its behind [10]. Wormhole routing decreases packet delay in compare to method of store and forward [6, 7].

Routing can be deterministic or adaptive. In deterministic routing, connection route between source and destination nodes is fixed, even there are several routes. But its flexibility is not sufficient to adapt with network conditions such as congestion and destruction. Adaption is an important feature that provides alternative routes to packets that confront to defect hardware, blocked channels and so on [6, 10, 11].

Deterministic routing prevents effective use of physical connections, because physical channels allocated inflexible. Routing by method of shortest path, direct packet in shortest route and so fault tolerance is low [4]. But in none short routing, chosen route may not be shortest one and it provide better fault tolerance. The best choosing route is dependent on network load. To reach good performance over all traffic loads, routing algorithm must be adaptive, and chooses routes based on network conditions. More flexibility in adaptive routers under same load and none same load can improve the performance.

In wormhole routing it may have occurred a deadlock event [12]. Indeed because of channels and buffers dependency, deadlock in interconnection networks is determine as dependency cycle, and in this dependency cycle each deadlock message, demands other sources that are in control of other messages of cycle, and this performance leads to an inevitable dependency cycle in routing [6, 12, 13]. Deadlock is different from message blocking. It may some packets been blocked without any deadlock, but if this deadlock occurred it must block all packets. Messages blocking is recovery by slowing traffic, but deadlock cycle is not recovered by slowing traffic, and only way to recovering it is break down the dependency deadlock cycle.

There are some methods to confronting to deadlock such as prevention and avoidance that sources in those are not used as optimal. In these methods, routing was done as none adaptive and may partially adaptive, and creates some limitations in networks to prevent from deadlock [11, 14–16].

But detection methods and recovering deadlock are not in this way, routing in detection methods and recovering deadlock are done as adaptive and it allowed to deadlock [14, 17, 18]. In these methods, involving messages with deadlock are identifying and they reported, and then related cycle to this deadlock is breakdown to deadlock recovering. Most of existence multicomputer networks, using deterministic routing, although there were several routes between source and destination, but to avoidance from deadlock, deterministic routing determines a route from source to destination [17, 19].

2 A Fuzzy Method to Deadlock Detection

There are two base methods to deadlock detection. One of these methods is that if a message be awaiting more than determine time out in a router to achieve appropriate output channel, that message is introduced as involve deadlock message [1, 6, 11, 14, 20]. Other method is that if a router can't route a message and all appropriate output channels be blocked more than determine time out so that message introduced as involve deadlock message [21, 22]. Comparing two methods show that first method introduces more messages as involve deadlock message mistakenly. In extra blocking of message, it may create a condition that a message introduce as involve deadlock message but appropriate output channels of that message don't block completely and been at used. This be occurred because in extra blocking of message method, attention is just toward time of message blocking in a router, that if be more than threshold, it introduced as involve deadlock message. In blocking channel method, it may be a message that inter to router recently, all appropriate output channels of that message been blocked more than threshold and so that message introduce as involve deadlock message but it is correct if that message interred to router recently and not be blocked more than threshold in input channel. If deadlock occurred so it is clear that if input channel be a part of dependency cycle, it must attention to input channel. In this paper, because of defects of both referred methods, it was considered a combination of two methods with fuzzy parameters of threshold, in that if a message be blocked in an input channel more than threshold and appropriate output channels to that message been blocked more than threshold, so message is introduced as involve deadlock message. This method tries to reduce messages that introduced as involve deadlock message but mistakenly. But choosing threshold is an important subject, in that if low level is considered to threshold it may a message be introduced as involve deadlock message but in fact not be involve deadlock message and if high level was considered to threshold it may involve deadlock message be detected lately. Time out of deadlock detection cannot be constant with every traffic or network load. This paper tries that threshold of deadlock detection be determined based on amount of past and present traffic as fuzzy method. To fuzzy the threshold of deadlock detection, it is necessary that determine high and low level of threshold and it performed to determine amount of threshold using fuzzy method and it leads to this that threshold amount that determine by fuzzy method is not be out of control. Primary value of threshold in each node must be determined, and also its value must be determined in high and low levels of threshold. In each pulse, the virtual output channels are investigated. By studying free and busy channels in output channels, we can calculate network traffic as local and approximately in each node. And if perform these application in one-time range, we can calculate traffic as an average of traffics in pulse of that time range. An input fuzzy variable is employee to show the difference between traffics of now and past that is include a language value series and a fuzzy output variable namely offset is considered to make correct threshold that also is include a language value series. To maintain role of past in time ranges, we calculate traffic as an average of now and previous traffics. A series the laws of if- that are defined that indicate traffic changes in threshold as language variables. For example, if traffic amount be increase in different dimensions, then threshold or time out must be more than now, to prevent introduce common massage instead to deadlock one, so offset most be positive. Output fuzzy membership functions must identify as rectangular to input fuzzy parameter.

To defuzzification and accept appropriate Offset, it uses method of defuzzification of center of gravity. The offset is used to achieve amount of new threshold. (New threshold = offset + previous threshold).

3 Network Pattern

In each node, shortest path is choosing to take a packet to destination. Wormhole switching is used to routing. Each packet is divided to some flits. Header flit is routing in each node and other flits are behind it. Routing unit in each pulse, routes a packet, indeed regarding to destination chose an appropriate output channel to packet. Each physical channel includes three virtual channels. Routing unit can choose each of virtual channels relate to physical channel. Politic of chose of virtual channel to use bandwidth of physical channel is Round Robin. In each node a queue is considered to packets that are waiting to routing, packets that are wait in this queue as Round Robin. In each pulse just one packet is routing and that is packet in begin of queue, but if there is not appropriate output channel routing to it, it was send to end of queue to be routing in next round.

It considered a buffer to store flits in output and input virtual physical channels and their capacity is 4 flits. Choosing politic to of packet to using output and input buffers is fifo. In each pulse just one flit can pass from physical channel. In each node a crossbar switch is used to connection between output and input channels. In each pulse, crossbar switch allow several connects in output and input channels and connections are not effect on each other. During packet routing connection of input channel is performed to output channel, and this remained till packet cross to related switch and other packet can’t to cross to related channel. To injection, there are four injection channels, in each pulse, each node produces some packet, if there were some free injection channels. If not, packet must wait in productive node to inject in other time.

4 Simulation Results

Simulation was performed on 2 dimensional mesh network with 3 virtual channels in each physical channel and in 20,000 pulse. Traffic pattern of network is as same distribution, and in it each node can send a message to other nodes and can receive message from other nodes. Traffic is considered based on average of flits that receive in each node in each pulse (Figs. 1, 2 and 3).

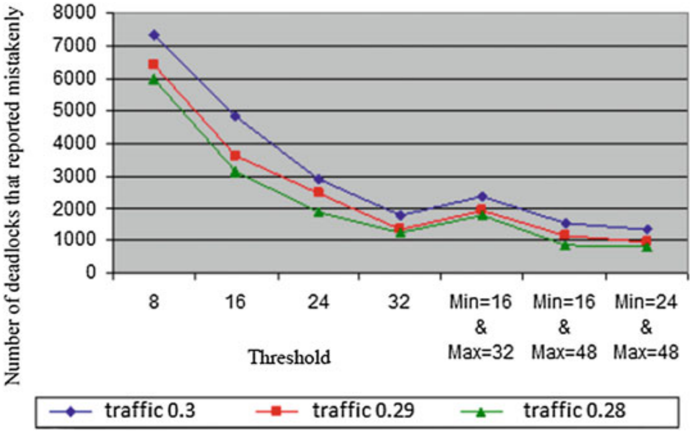


Fig. 1 This diagram is related to; number of error distinguishes of messages as deadlock in different amounts of time out, that use of message blocking method in deadlock detection. Indexes 16–32, 16–48 and 24–48 are related to fuzzy results. And indeed are low and high level of threshold

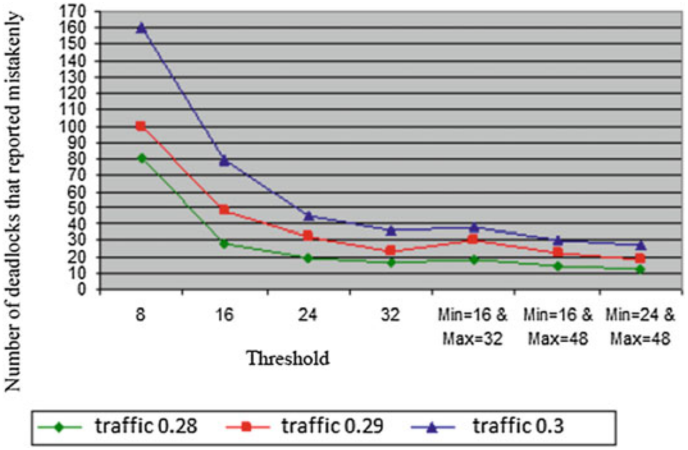


Fig. 2 This diagram is related to; number of error distinguishes of messages as deadlock in different amounts of time out, that use of channel blocking method in deadlock detection. Indexes 16–32, 16–48 and 24–48 are related to fuzzy results. And indeed are low and high level of threshold

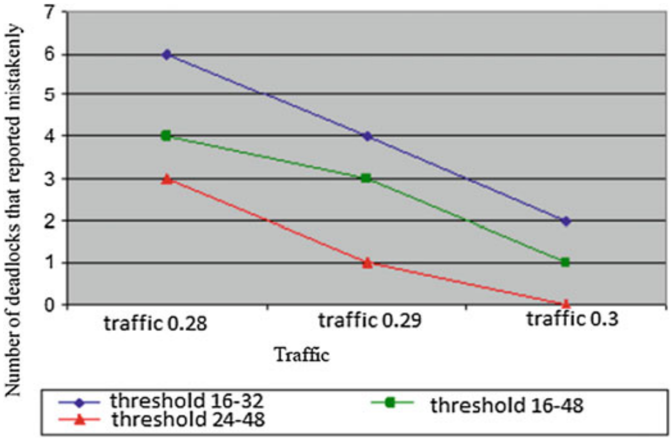


Fig. 3 This diagram is related to, number of error distinguishes of messages as deadlock in different amounts of time out, that use fuzzy method of message blocking and channel blocking (method of combine fuzzy) in deadlock detection. Indexes 16–32, 16–48 and 24–48 are related to fuzzy results. And indeed are low and high level of threshold

5 Conclusion

Deadlock is a phenomenon that can't be distinguished as certainly. It can be distinguishing deadlock by considering network deadlock. In this paper, deadlock is distinguished by fuzzy method and it be tried that messages that report as involve deadlock mistakenly were reduced. Regarding to simulation results, if low and high level of threshold be determined appropriately, incorrect determine of deadlock is reduced. in fuzzy method threshold will change as fuzzy by network traffic changes. Therefore, threshold has any specific dependent to message length. But threshold choice is an important problem, if low level be considered to threshold, it may message that isn't involve deadlock, introduce involve deadlock message mistakenly, and if high level be considered to threshold, it may that involve deadlock message be detected lately, so channels been blocked in more time, and it leads to delaying in messages. Threshold is change based on network changes and its amount is determined as fuzzy by considering network traffic. Using fuzzy methods in compare to classic methods is include better results in routing and deadlock detection, and it can be used fuzzy method to interconnection network. By combining deadlock detection methods and by considering fuzzy laws, it can be decreased messages that introduced as involve deadlock message mistakenly.

References

1. Diatom J et al (2003) Interconnection networks: an engineering approach. Morgan Kaufmann Publishers. 600p. (In San Francisco, CA, USA). ISBN 1558608524

2. Zargham M (1996) Computer architecture: single and parallel system. Prentice-Hall International. 472p. (In Upper Saddle River, NJ, USA). ISBN 0135294975
3. El-Rewini H et al (2005) Advanced computer architecture and parallel processing. Wiley, New York, 287p. ISBN 0-471-46740-5
4. Glass CJ, Ni LM (1991) The turn model for adaptive routing. In: Proceedings of the 19th annual International Symposium on Computer Architecture (ISCA), Queensland, Australia, 19, 21 May 1992, pp 278–287
5. Dally WJ et al (2004) Principles and practices of interconnection networks. Morgan Kaufmann Publishers. 550p. (In San Francisco, CA, USA). ISBN 9780080497808
6. Martinez JM, Lopez P, Duato J (2003) FC3D: Flow control-based distributed deadlock detection mechanism for true fully adaptive routing in wormhole networks. IEEE Trans Parallel Distrib Syst 14(8):765–779
7. Dally WJ, Seitz CL (1987) Deadlock-free message routing in multiprocessor interconnection networks. IEEE Trans Comput C-36:547–553
8. Kim J, Liu Z, Chien A (1997) Compressionless routing: a framework for adaptive and fault-tolerant routing. IEEE Trans Parallel Distrib Syst 8:229–244
9. Duato J (1993) A new theory of deadlock-free adaptive routing in wormhole networks. IEEE Trans Parallel Distrib Syst (TPDS) 4:1320–1331
10. Mohapatra P (1998) Wormhole routing techniques for directly connected multicomputer systems. ACM Comput Surv 30:374–410
11. Sharifian-Nia S, Vafaei A, Shahimohamadi H (2012) Deadlock recovery technique in bus enhanced NOC architecture. Int J VLSI Des Commun Syst (VLSICS) 3(4)
12. Dally WJ, Aoki H (1993) Deadlock-free adaptive routing in multicomputer networks using virtual channels. IEEE Trans Parallel Distrib Syst 4:466–475
13. Taktak S, Desbarbieux JI, Encrenaz E (2008) A tool for automatic detection of deadlock in wormhole networks on chip. ACM Trans Des Autom Electron Syst 13(6):1–22
14. Soojung L (2009) Deadlock detection and recovery for true fully adaptive routing in regular wormhole networks. J Inf Sci Eng 25:465–479
15. Soojung L (2007) A deadlock detection mechanism for true fully adaptive routing in regular wormhole networks. Comput Commun 30:1826–1840
16. Anjan KV, Pinkston T (1995) DISHA: a deadlock recovery scheme for fully adaptive routing. Proceedings of Ninth Int'l Parallel Processing Symposium, Santa Barbara, CA, USA, 25, 28 April 1995, pp 537–543
17. Khonsari A, Shahrabi A, Ould-Khaoua M, Sarbazi-Azad H (2003) Performance comparison of deadlock recovery and deadlock avoidance routing algorithms in wormhole-switched networks. IEEE Proc Comput Digit Tech 150:97–106
18. Soojung L (2006) Turn-based deadlock detection for wormhole routed networks. In: CIT '06 Proceedings of the Sixth IEEE International Conference on Computer and Information Technology, Seoul, Korea, 20, 22 September 2006
19. Mirza-Aghatabar M, Tavakol A, Sarbazi-Azad H, Nayebi A (2008) An adaptive software-based deadlock recovery technique. In: Proceedings of 22nd International Conference on Advanced Networking and Applications-Workshops, Okinawa, Japan, pp 514–519, 25, 28 March 2008
20. Al-Dujaily R, Mak T, Xia F, Yakovlev A, Palesi M (2012) Embedded transitive closure network for run-time deadlock detection in networks-on-chip. J IEEE Trans Parallel Distrib Syst 23(7):1205–1215
21. Anjan KV, Pinkston TM (1995) An efficient fully adaptive deadlock recovery scheme: DISHA. International Symposium on Computer Architecture, Santa Margherita Ligure, Italy, 22, 24 June 1995, pp 201–210
22. Anjan KV, Pinkston TM, Duato J (1996) Generalized theory for deadlock-free adaptive wormhole routing and its application to Disha concurrent. Proceedings of 10th Int'l Parallel Processing Symposium, Honolulu, HI, USA, 15, 19 April 1996, pp 815–821

Occupancy Overload Control by Q-learning

Mehdi Khazaei

Abstract Session Initiation Protocol (SIP) is considered as a signaling protocol for IP multimedia subsystem (IMS). IMS is introduced by 3rd generation partnership project (3GPP) as signaling foundation in next generation networks (NGN). Despite having the features such as: text based, IP based, independent of the data transmission, support for mobility and end-to-end, the SIP protocol has not suitable mechanism to deal with overload. Therefore, many mechanisms are proposed to control overload in SIP networks. One of their most famous is occupancy CPU (OCC) that is used in many researches. In traditional OCC, the value of parameters is indicated and they are used in subsequent documents. In this paper, optimal parameters value is obtained by Q-learning algorithm. Because modeling a large SIP network is impossible by mathematical relationships and it is a heuristic problem, Q-learning is the best method to compute the parameters. The simulation results demonstrate that the Q-learning output is comparable with traditional OCC.

Keywords Occupancy · Overload control · SIP · Q-learning

1 Introduction

SIP protocol is widely used in modern communications as the signaling protocol. In the near future, current networks are integrated based on IP and named next generation networks (NGN). SIP is considered as a signaling protocol for NGN. SIP has very good properties but it has not suitable mechanism to deal with overload. This challenge will cause a wide range of NGN users facing with a sharp drop in quality of service.

The overload occurs when the rate of incoming requests to the SIP server is more than the server capacity like a voting in the TV show. In the overloaded SIP servers,

M. Khazaei (✉)
Information Technology Department,
Kermanshah University of Technology, Kermanshah, Iran
e-mail: m.khazaei@kut.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_58

retransmission mechanism begins, which increases the load in the networks. Retransmission mechanism is begun on UDP protocol. Since it is not economical to consider the worst case to be designed for the offered load, SIP protocol uses 503 responses (Service Unavailable) to control the overload. The overloaded servers reject the requests with 503 responses. Rejecting the requests has a negative effect on overload because the server must spend all resources to reject the requests.

In RFC 6357, three overload mechanisms are introduced: local, hop-by-hop and end-to-end. In local control the server begins to reject the requests by 503 messages. In the hop-by-hop mechanism, the upstream servers are notified by itself or downstream servers. In end-to-end mechanism, the edge servers adjust the amount of traffic sent to the overloaded server and middle servers forward the information about the overloaded server to the edge servers. End-to-end overload control causes fewer resources involve in the requests which finally are rejected [1].

The innovation of this paper is that, it uses Q-learning to calculate the parameters of occupancy SIP overload control. These parameters are computed by authors but calculating method is not noticed in [2].

The paper is organized as follow: Sect. 2 describes background of the problem. Section 3 discusses related works. Proposed method is outlined in Sect. 4. Section 5 presents the simulation and conclusion are made in Sect. 6.

2 Background

Since the proposed method is the combination of SIP protocol and Q-learning method therefore, it is necessary to briefly discuss these two concepts.

2.1 SIP Protocol

SIP is the most important application layer signaling protocol standardized by IETF in RFC 3261. It is used to manage the session in many internet applications such as: VOIP, instant messages and video conferencing. SIP architecture is composed of user and server agent entities. The user agents contain two categories, client user agent (UAC) and server user agent (UAS). UAC is applicant and UAS is respondent. The servers or proxy servers perform routing to transfer message to a user location. SIP transaction is a request and all the relevant responses are exchanged between two adjacent components. The SIP transaction in terms of confirmation and retransmission is classified to invite and non-invite. Proxy servers are configured depending on the circumstances and the need of the network in stateful or stateless forms. The stateful, server maintains information for each transaction while in stateless server there is no need to store any record of a transaction [3].

Figure 1 shows the call establishment between two UAs in the stateful proxy servers. The call establishment begins by sending invite message on behalf of the

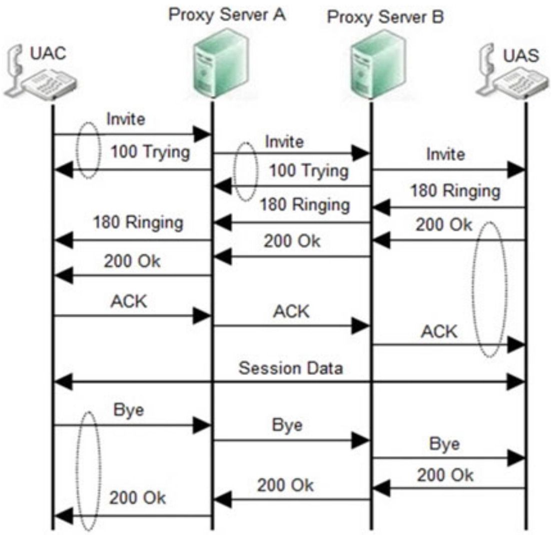


Fig. 1 The session creation between two UAs

UAC. When the proxy server receives the invite message, replies with 100-trying message as confirmation and passes invite message to the UAS. When UAS receives the invite message, the response 180-Ringing is sent then 200-OK is sent by answering the call. Finally, after sending ACK by UAC, a call is established between the parties then, data is exchanged without passing through a proxy server. Finally, ACK is sent by the UAC and call is established between UAs. The call will be terminated by sending a BYE by one of the parties and sent 200-OK by another [1].

2.2 Q-learning

In artificial intelligence, we can define the agents with the learning capability. Like a living organism, an intelligent agent can learn based on trial and error. The goal is achievable according to agent abilities, environment states and the value of actions. For this purpose, agent does an action on environment then it is receives a response.

In an undefined environment, reinforcement learning is agent ability to reach a goal without any outside help. Different algorithms are offered to implement reinforcement learning. The basis of these algorithms is relation between environment states, action performed in this states and the reward obtained to do this action. One of the famous reinforcement learning is Q-learning.

Standard reinforcement learning or Q-learning is an online learning. With reward function, Q-learning can learn optimal policy on Markov process [4]. A table include of couples < action, status > is the output of Q-learning algorithm. Table 1 shows the Q-learning algorithm.

Table 1 Q-learning algorithm

1	Initialized to the table	
2	Obtain the current status of the environment (S)	
3	Repeat the following loop until the end condition	
4	4-1	9-Select the action (a) as Random (Exploration)
	4-2	11-Get the reward of environment (r)
	4-3	13-Get the new status of environment (S′)
	4-4	15-Change the value of Q-table based on (1)
		16-$Q(s,a) = \alpha\left(r + Max'_a Q(S', a')\right) + (1-\alpha)Q(s,a)$ (1)

3 Related Works

In telecommunication networks, overload is the challenge about which an extensive research has been done. In this regard, the studies about overload control in SIP networks have been continued. Since the proposed method improves an end-to-end mechanism, we glance to suggested end-to-end overload control methods as related works [5].

In [2], OCC algorithm (occupancy algorithm) is introduced. In OCC, each proxy server calculates restriction on CPU load based on (1) and maintains list of targets restrictions. F is a probability to accept the session requests that is calculated dynamically.

$$f_{t+1} = \begin{cases} f_{min} & \text{if} \quad \emptyset_t f_t < f_{min} \\ 1 & \text{if} \quad \emptyset_t f_t > 1 \\ \emptyset & \text{otherwise} \end{cases} \tag{1}$$

Φ_t is multiplicative increase/decrease factor is given by (2). F_{min} shows the threshold for the minimum fraction of traffic accepted. Φ_{max} defines the maximum possible multiplicative increase in f from one epoch to the next and φ_{targ} is a threshold to occupancy the CPU. Φ_t is the current CPU occupation [2].

$$\emptyset_t = \min\left\{\frac{\rho_{targ}}{\rho_t}, \quad \emptyset_{max}\right\} \tag{2}$$

Targets servers send theirs restrictions to direct upstream neighbors. Server j maintains a list of (f^{j1}(d), f^{j2}(d),…, f^{jL}(d)) values if its corresponding downstream neighbors for target d are (j^1(d), j^2(d),…, j^L(d)), where f^{j1}(d) is the f value the server receives from its downstream neighbor j^1 for target d. When a server receives a restriction of the target server, corresponding item of that target server is updated in restriction list. Furthermore, server calculates the new restrictions based on f^i(d) = min{f^{j1}(d), f^{j2}(d),.., f^{jL}(d), f^i} then forwards it to the direct upstream neighbors. Therefore, the restrictions are propagated to all edge servers. The edge servers based on these restrictions decide to forward or reject the invite requests [2].

In [6], when server is overloaded, scheduler transferred overload to direct upstream neighbors until they finally reach the edge servers. Then the edge servers apply traffic restrictions. In [7], for each target server one DEOC is created, controlling entry the load into the destination. The core servers only implement local overload control. If necessary, requests are rejected by 503 responses and 503 responses are forwarded to the edge servers. According to 503 responses, DEOC implements restrictions to forward or reject the request to the target server. In [8], as [7], for each target server, one PEOC is created, following SIP requests and 503 responses to apply limitations to the destination. In [3], a window-based end-to-end mechanism is offered to control and manage overload in SIP networks. Holonic multi-agent system is used by the due method and named WHOC. Based on past observations, the normalized least mean square algorithm is used to estimate the each agent window size. In WHOC, the size of the windows is adjusted in the way that no overload will occur in network paths, which could be fulfilled through using holonification properties and communications. WHOC offers an appropriate window size for edge servers to control the load from the beginning of the network and prevent network overload. Reference [1] reuse multi-agent with holonic organization to implement a fairness end-to-end overload control, called HOC. In this mechanism, fitting is done based on received requests and used as predictor. When overload is occurred, causing is noticed to reduce sending load based on announced percent. In repression of the servers, fairness is observed.

4 OCC with Learning Capability

In the many overload researches like OCC, the algorithm performance is depends on the parameters value. In order to confirm this issue, φ_{targ} plays important role in OCC efficiency. If the value of φ_{targ} is selected very high the throughput is strongly decreased because server exits from the safe mode. As result, the transactions delay is increased and retransmission mechanism is activated. On the other hand, low value of φ_{targ} causes power of the processor is wasted. Therefore the suitable value of φ_{targ} has very important efficacy in OCC performance. In addition to φ_{targ}, $\Phi_{\max}$ should given proper values. The question is how these parameters should be determined to be optimal?

The mathematical relations are the best and most accurate method to calculate the above parameters. In SIP networks, overload control is a heuristic and complex problem. Therefore; only one or two serial servers could be approximately modeled by mathematical while it's impossible to modeling large network in this manner [9]. Another method is trial and error. In this method, the values are generated by simulation and the best ones are selected. These values are appropriate for current state of the network and if the conditions change, the values must be re-calculated. When the number of parameters is increased, this method is more time consuming. In such circumstances, the learning method could be used hence the Q-learning is selected in this paper.

In some circumstances, the value of Φ_{max} and φ_{targ} should be different for each server but in [2], these two parameters are respectively considered 5 and 0.9 on all servers. In this paper, Φ_{max} and φ_{targ} are separately calculated by Q-learning algorithm.

In SIP network, each server could be considered as an agent that is learning the environment conditions. These agents estimate the optimal value of Φ without the expert knowledge and previous environment information. The necessary entities to design the Q-learning algorithm are defining states scope, actions scope and reward function. In this paper, it is tried to reduce the number of actions and the states to accelerate the algorithm convergence.

4.1 States Scope

In order to define the states scope, one main state is considered for each of the Φ_{max} and φ_{targ}. The numbers of sub-states are allocated to each Φ. Figure 2 shows the states scope. Φ_0 and φ_0 are the initial state while Φ_{i} and φ_{i} are sub-states. N is the number of sub-states for Φ_{max} and m is the number of sub-states for φ_{targ}. Number of all states is calculated by (3).

$$|\mathrm{States}| = \sum_{\mathrm{i}=0}^{\mathrm{n}} \emptyset_{\mathrm{i}} + \sum_{\mathrm{j}=0}^{\mathrm{m}} \varphi_{\mathrm{j}} \tag{3}$$

In learning process, each of the states specifies the value of Φ. For example suppose, the value of Φ_0 is three and each of its sub-states is one, if current state is first sub-state the value of Φ_{max} is four.

4.2 Actions Scope

In different states, six actions are considered to select. These actions are as follow: increase Φ_{max}, decrease Φ_{max}, unchanged Φ_{max}, increase φ_{targ}, decrease φ_{targ} and unchanged φ_{targ}. Increase means that the current state is changed to the next state and decrease is changed to the previous state.

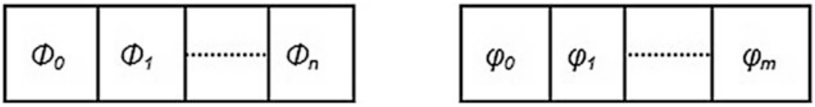


Fig. 2 States scope of OCC parameters

4.3 Q-table

In learning process, Q table is used to save update value based on (1). This table is considered as matrix which states are shown by rows and actions are displayed by columns. Table size is 6*|states| and its initial value is zero.

4.4 States Change Function

Another requirement to design Q-learning algorithm is state's change function. This function determines how the states are changed by doing an action. In this paper, the states are changed based on selected action as follow:

If the selected action is increase and there is not in the last state, the state changes to next state. But if current state is the last state, next state is current state according to (4).

$$\begin{aligned} &\mathrm{Transfer}(\Phi_{\mathrm{i}}, \mathrm{increase}) = \Phi_{\mathrm{i+1}};\quad \Phi_{\mathrm{i}} \neq \Phi_{\mathrm{n}} \\ &\& \mathit{Transfer}(\varphi_{\mathrm{i}}, \mathrm{increase}) = \varphi_{\mathrm{i+1}}\quad ; \varphi_{\mathrm{i}} \neq \varphi_{\mathrm{m}} \end{aligned} \tag{4}$$

If the selected action is decrease and there is not in the first state, the state changes to pervious state. But if current state is the first state, next state is current state according to (5).

$$\begin{aligned} &\mathrm{Transfer}(\Phi_{\mathrm{i}}, \mathrm{decrease}) = \Phi_{\mathrm{i-1}}\quad ; \Phi_{\mathrm{i}} \neq \Phi_{0} \\ &\& \mathit{Transfer}(\varphi_{\mathrm{i}},\ \mathrm{decrease}) = \varphi_{\mathrm{i-1}}\quad ; \varphi_{\mathrm{i}} \neq \varphi_{0} \end{aligned} \tag{5}$$

If the selected action is unchanged, the next state is current according to (6).

$$\begin{aligned} &\mathrm{Transfer}(\Phi_{\mathrm{i}},\quad \mathrm{unchanged}) = \Phi_{\mathrm{i}} \\ &\& \mathit{Transfer}(\varphi_{\mathrm{i}},\ \mathrm{unchanged}) = \varphi_{\mathrm{i}} \end{aligned} \tag{6}$$

4.5 Reward Function

In states scope, guidance of agent to achieve the optimal policy is done by reward function. Indeed, the reward function distinguishes the value of the agent action in current status and how much the agent comes close to the goal. In this section, server throughput is considered as reward function. The obtain reward by a server is proportional with increase throughput of itself than previous states. Due this action, if throughput is increased in upstream servers the server achieves more reward and vice versa. The reward function is defined based on (7). G^{oldi} and G^{newi} are throughput of server ith before and after the selected action on current status. K is the number of upstream severs of server ith.

$$\mathrm{Rewrd_i} = \left(\mathrm{G_i^{new}} - \mathrm{G_i^{old}}\right) - \frac{1}{\mathrm{k}}\sum_{\mathrm{j=1}}^{\mathrm{k}}\left(\mathrm{G_j^{old}} - \mathrm{G_j^{new}}\right) \tag{7}$$

5 Simulations

The network simulator (NS-2) and the standard topology as shown in Fig. 3 is selected for simulation. The capacity of all proxy servers is equally considered 300 calls per second (cps). Hence, network capacity is equal to 900 cps.

UDP is used as transport layer protocol. Therefore, proxy server work in stateful mode and the reliability of message transmission is achieved by SIP retransmissions. Except to φ_{targ} and Φ_{max}, other parameters and methods is set as OCC algorithm [2]. The proposed method is called L_OCC and the parameters of Q-learning are as Table 2. Goodput and call setup delay are considered as performance metric. The Goodput is defined as a number of successful calls in a second. A call is successful when UA receives 200-OK responses in less than 10 s. Call setup delay is considered as the time between sending the invite request and receiving 200-OK responses.

After simulation for different offered load and Q-learning convergence, Φ values are extracted from Q table and they are set in OCC algorithm. In Fig. 4, a part of Q table is shown for server 6. The obtained Φ_{Max} is 5 equal to the amount of OCC algorithm.

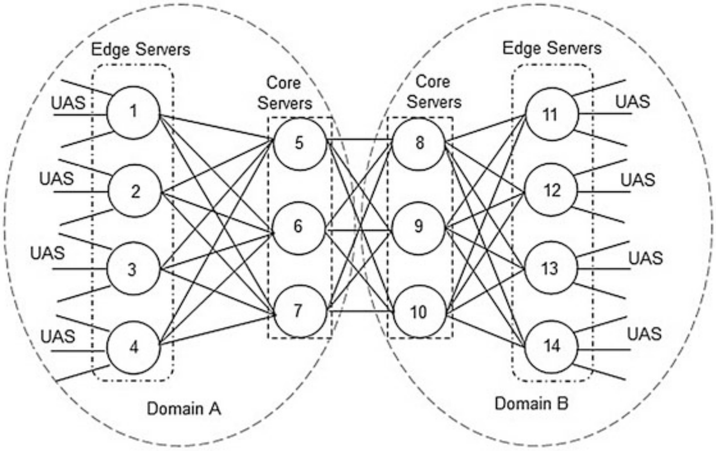
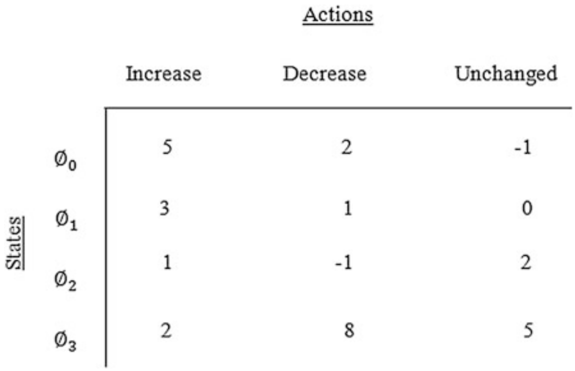


Fig. 3 Simulation Model [1]

Table 2 Q-learning parameters

Parameter	Value
Value of Φ_0	3
Value of Φ_i	1
Number of Φ_i	4
Value of φ_0	0.5
Value of φ_0	0.1
Number of φ_i	5
Learning rate (α)	0.9 and decrease linearly
Discount factor (γ)	0.4



		Actions		
States		Increase	Decrease	Unchanged
	$\emptyset_0$	5	2	-1
	$\emptyset_1$	3	1	0
	$\emptyset_2$	1	-1	2
	$\emptyset_3$	2	8	5

Fig. 4 A part of Q table for server 6

5.1 Goodput and Call Setup Delay

In Fig. 5, goodput are evaluated under deferent offered load. As depicted in Fig. 5, when offered load is below the network capacity, goodput is equal to the offered load for both methods. When the offered load goes beyond the network capacity, L_OCC and OCC present approximately constant and equal goodput. Nevertheless, goodput is below the network capacity because based on [6], a threshold is considered for CPU occupancy, so under heavy load percent of CPU capacity is wasted and is not used.

Figure 6 shows the results of call setup delay. When the offered load is under the capacity of the network, call setup delay are negligible. When offered load is above the network capacity, OCC and L_OCC have approximately equal call setup delay with fluctuated. The reason for the fluctuation is that if CPU occupancy exceeds the threshold, servers begin to reject the requests. Therefore process time of requests is increased.

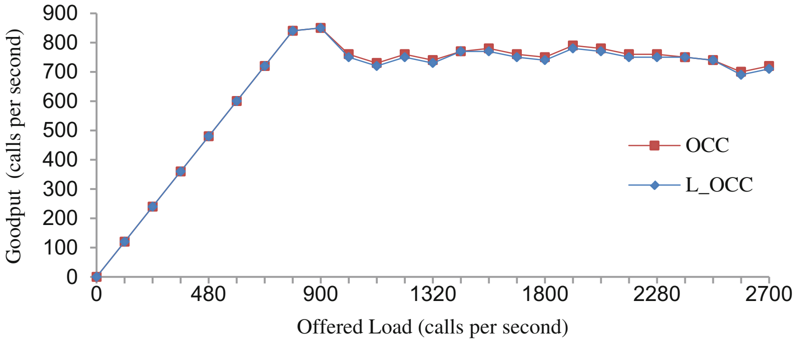


Fig. 5 Goodput for OCC and L_OCC

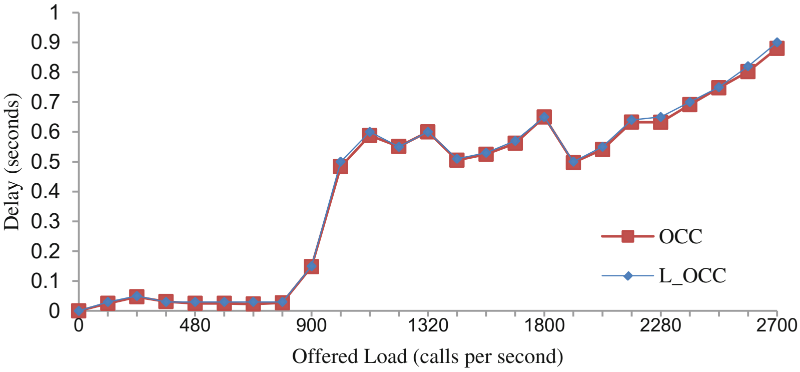


Fig. 6 Call setup delay for OCC and L_OCC

5.2 *Stability and Reactiveness*

Stability means that the overload control methods should be designed in a fashion that it does not impose fluctuation on SIP servers. Reactiveness requires an overload control mechanism to react quickly to the overload [3]. To test stability and reactiveness, in the beginning, the offered load is 600 cps, below the network capacity. The offered load is suddenly increased to 2700 cps at 100 s, three times the network capacity. The offered load is gone down to 600 cps at 200 s. The Goodput and the call setup delay are measured as shown in Figs. 7 and 8.

In Fig. 7, both mechanisms respond quickly to the offered load sharp changes. When the offered load is above the network capacity, the goodput is not increased immediately because the control parameters update in period T seconds. As a result, when an invite request is accepted, all subsequent related messages should be accepted but after overload parameters update, L_OCC the better goodput.

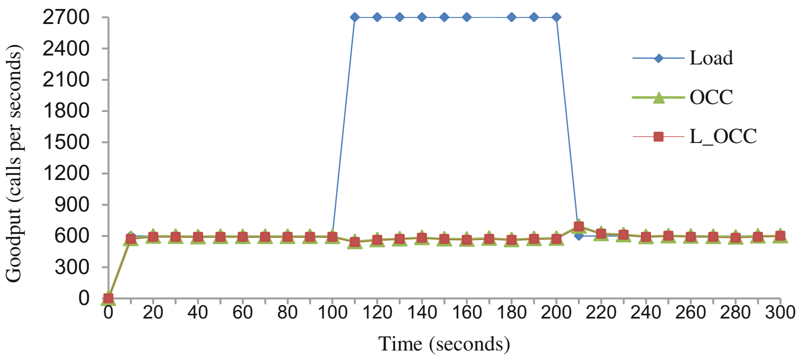


Fig. 7 Goodput when the offered load is suddenly changed

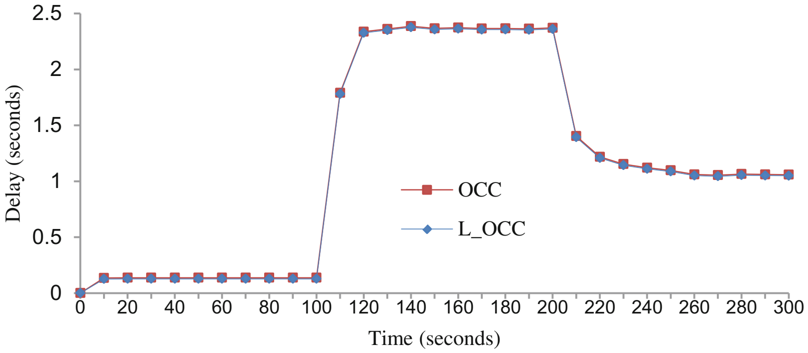


Fig. 8 Call setup delay when the offered load is suddenly changed

In Fig. 8, when the offered load is abruptly increased, many requests enter the network and cause delay in OCC and L_OCC. But after updating the parameters, entering requests is controlled by edge servers. Therefore, call setup delay is slightly decreased.

6 Conclusion

Due to the importance of multimedia networks and their increasing development based on SIP protocol, the issue of overload control in SIP networks becomes more and more essential. Therefore, moving from traditional methods to modern and intelligent methods is inevitable. In SIP network, the overload control is introduced as a complex system. In such case, mathematical modeling is impossible to obtain proper value of parameters. But learning method can be implemented for any SIP network with any scale and it is fully scalable. Simulation results demonstrated that learning method can keep high throughput even when the offered load exceeds the capacity of the network. Also it is stable, when the offered load is suddenly changed. In the SIP networks, the overload control based on learning is a new method and has not been fully touched. Nevertheless, with the progress in multimedia networks and its intelligence, much research can be carried out in the case of this issue.

References

1. Khazaei M, Mozayani N (2017) Overload management with regard to fairness in SIP networks by holonic multi-agent systems. Int J Netw Manage 27(3)
2. Hilt V, Widjaja I (2008) Controlling overload in networks of SIP servers. IEEE International Conference on Network Protocols 2008, pp 83–93, Orlando, FL
3. Khazaei M, Mozayani N (2015) A dynamic distributed overload control mechanism in SIP networks with holonic multi-agent systems. Telecommun Syst 63(3):437–455
4. Abdoos M, Mozayani N, Bazzan ALC (2011) Traffic light control in non-stationary environments based on multi agent Q-learning. In: 14th International IEEE Conference on Intelligent Transportation Systems, 2011, pp 1580–1585
5. Khazaei M, Mozayani N (2017) An end-to-end overload control method in SIP networks. 1st International Conference New Perspective in Electrical and Computer Engineering-May 12th, 2017, Tehran, Iran
6. Wang Y (2010) SIP overload control: a backpressure-based approach. In: Proceedings of ACM SIGCOMM (poster), August 2010, pp 399–400
7. Liao J, Wang J, Li T, Wang J, Wang J, Zhu X (2012) A distributed end-to-end overload control mechanism for networks of SIP servers. Comput Netw 56(12):2847–2868
8. Wang LJ, Li T, Wang J, Wang J, Qi Q (2014) Probe-based end-to-end overload control for networks of SIP servers. J Netw Comput Appl 41:114–125
9. Hong Y, Huang C, Yan J (2013) Modelling chaotic behaviour of SIP retransmission mechanism. Int J Parallel Emerg Distrib Syst 28(2):95–122

Mobile Smart Systems to Detect Balance Motion in Rehabilitation

Saedeh Abbaspour and Faranak Fotouhi Ghazvini

Abstract In the present paper, a mobile smart system is introduced to assess the individual's balance in remote rehabilitation. Gyroscope sensor is used in order to analyze the individual's balance during the remote rehabilitation for five movement activities. The data acquired from the sensor are transmitted via edge layer and SDN controllers to the server database leading to reduction of costs, control of network traffic, and mitigation of delay. The transmitted raw data are analyzed using unsupervised K Means algorithm. The respective algorithm performed the best separation with K values equal to 5 and 8. In fact, accuracy of this method for the 5 movement activities is equal to 0.7 and 0.8 for k = 5 and k = 8, respectively.

Keywords Smart system · Sensor · Movement balance · Rehabilitation
Unsupervised algorithm

1 Introduction

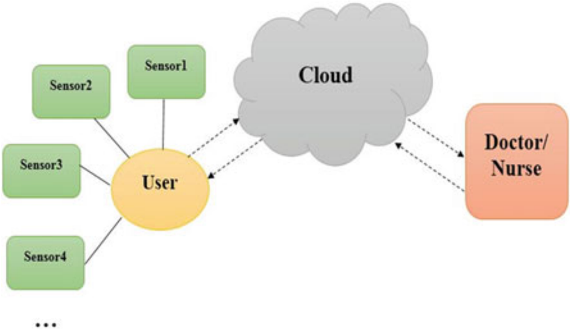
The recent progress in medicine and the growth of communication technologies have increased the life expectancy among societies. Some people lose their ability to move as a result of incidents such as driving accidents, heart attacks or cerebral strokes. They return to normal life through short-term or long-term rehabilitation. Nowadays, for the purpose of patient's comfort as well as saving the expenses, it has become possible to remotely rehabilitate patients and improve their health conditions. In remote rehabilitation a system is equipped with sensors which sends the information acquired from the patients' conditions to the physician for analysis. It then applies the appropriate feedback commensurate with the patient's conditions. Figure 1 schematically illustrates the remote rehabilitation system.

S. Abbaspour (✉) · F. F. Ghazvini
Department of Engineering, Qom University, Qom, Iran
e-mail: s.abbaspour@qom.ac.ir

F. F. Ghazvini
e-mail: f-fotouhi@qom.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_59

Fig. 1 Remote supervision rehabilitation system



The sensors available in this system can be used as wearable and ambient devices. Kinect-Sensor is an instance of ambient sensor which is used for performing different movement activities during rehabilitation [1]. Wearing sensors could be also utilized for carrying out rehabilitation of injured persons.

For instance, different sensors such as accelerometer, gyroscope, and PPG[1] are applied in Fig. 1 where vital and movement data extracted from the patient are instantaneously provided to the physician and nurse via cloud space. Subsequently, the data received by the physician or nurse are analyzed in order to apply the right remedy. Then the necessary measures will be transmitted for improvement of patient's condition via cloud space. The respective data are kept in the database and are stored as a history of individuals' health data. Actually, the diagnosis and treatment process of patients are carried out more rapidly and easily with the implemented system. Therefore, when the person is doing his/her routine daily activities, alarm will be sent for performing a specific act to the patient based on the data transmitted from sensors to the medical center [2]. In the present paper, gyroscope movement sensor along with Sim900 module was used as a smart system so as to detect movement balance of the person versus time and also to instruct the conventional exercises for improvement of movement activity. Furthermore, edge technology with SDN[2] was used for better transmission of data acquired from the sensors. The advantage of using this technology over cloud servers is the fact that edge technology is closer to accessible sources and computational and storage sources are placed in the vicinity of mobile systems and sensors [3, 4]. It also overshadows issues like network traffic control, costs, and bandwidth. In fact, implementation of edge technology will lead to delay reduction of the transmitted data [4], reduction of costs and interval of data transmission, improvement of service quality [5], decrease in energy consumption, reduction in number of servers, and support of viability [3]. Accordingly, if there is any problem in sending the data, they will be resent to edge layer. In fact, the respective layer will act like a local network. This layer is sometimes referred to as "fog". Just as "fog" is close to

[1] PhotoPlethysmography (PPG)

[2] Software Defined Network (SDN)

earth surface, the respective layer is also close to accessible sources. Edge layer is between the cloud structure and sensors available in the health system. This system encompasses health sensors existing in wearable and ambient varieties. Edge technology is therefore used to quickly analyze the data received from the sensors. Yet, SDN[1] network can be utilized to control the transmitted data within a single package, where SDN actually serves as the intermediate layer between edge and central computer. SDN architecture is viable, cost-effective, and flexible.

On contrary to the traditional network, data panel is separate from the control panel in SDN network. The controller is like an operating system of the network that undertakes control of the hardware and facilitates automatic administration of the network. This operating system provides the whole network with a concentrated and integrated programmable interface. Like the operating system on a computer enabling reading and writing for applied programs, the network's operating system also enable observation and control of the network. Hence, the controller does not administer the network solely but also merely acts as a programmable interface that

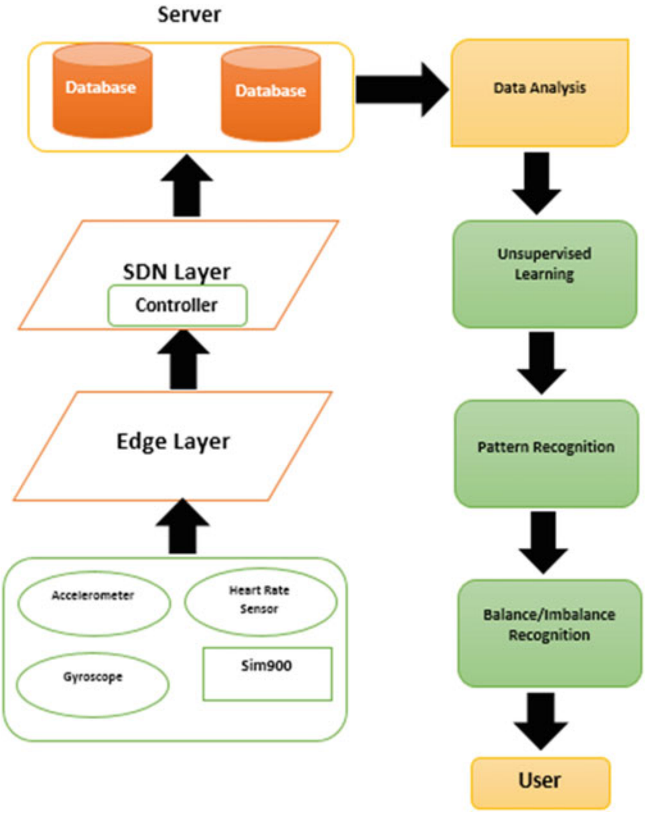


Fig. 2 Relational structure of edge and SDN layers

enables network administration of the user's software programs. As a result, more advantages will be achieved by using SDN network instead of traditional networks. Figure 2 represents the relational structure between edge layer and SDN layer in the field of rehabilitation.

2 Sensors

Different types of sensors used in remote rehabilitation will be analyzed in the present section. Table 1 includes wearable and ambient sensors. Each of the sensors is mounted in different parts of body or in the space around the patient. MPU6050 gyroscope module was used in the present study to receive the movement data for implementation of a smart system. The data received from the sensor are monitored based on three axes i.e. X, Y, and Z. They actually show the movement angle precisely in addition to acceleration. The respective data are transmitted via Sim900 module within the internet environment and stored in the database for analysis of movement balance of the patient and can be restored for knowing about the person's progress level during consecutive exercises. This sensor has been designed as a small kit and installed around the ankle.

Table 1 The sensors used in analyzing movement activities and remote rehabilitation

Sensors	Type	References
Accelerometer	Wearable/ Smartphone	[2, 6–12]
PPG	Wearable	[2, 8]
Accelerometer and Gyroscope	Wearable/ Smartphone	[2, 6, 13, 14]
ECG[a]	Wearable	[2, 6, 8]
Magnetometer	Wearable	[13]
FSR[b]	Wearable	[15]
Gyroscope	Wearable/ Smartphone	[8, 16, 17]
EMG[c]	Wearable	[8, 18]
PIR	Wearable	[16]
Accelerometer, gyroscope and magnetometer	Wearable/smartphone	[19–21]
Kinect sensor	Wearable	[1]
EEG[d]	Wearable	[2]

[a] Electrocardiography (ECG)
[b] Force Sensitive Resistor (FSR)
[c] Electromyography (EMG)
[d] Electroencephalogram (EEG)

3 Research Method

In the current research, 10 men and women between the age intervals of 25–62 years were chosen to investigate and analyze the data received from the gyroscope sensor. For the purpose of analyzing the balance status, the chosen individuals performed the movement activities namely walking, running, standing, sitting, and lying for 5 h. The sensors specified in Table 1 were used in the previous plans for controlling the movements in the field of rehabilitation. The intent is to implement Sim900 module along with gyroscope sensor for sending the data via internet. Aimed at improving information transfer via internet, edge technology was used to process the data and send them to the central system. Also, SDN was utilized after edge layer for better online transmission of data. These two technologies were applied along with each other aimed at reducing the costs, controlling the network traffic, and mitigating the delay in transmitted data acquired from the sensors. Figure 3 illustrates the Edge technology along with the SDN for transmission of gyroscope data to the internet-based server. The transmitted data represent precise angle in addition to movement acceleration. As shown in Fig. 2, the respective data are stored in the server and are assessed for further analysis of the person's movement balance.

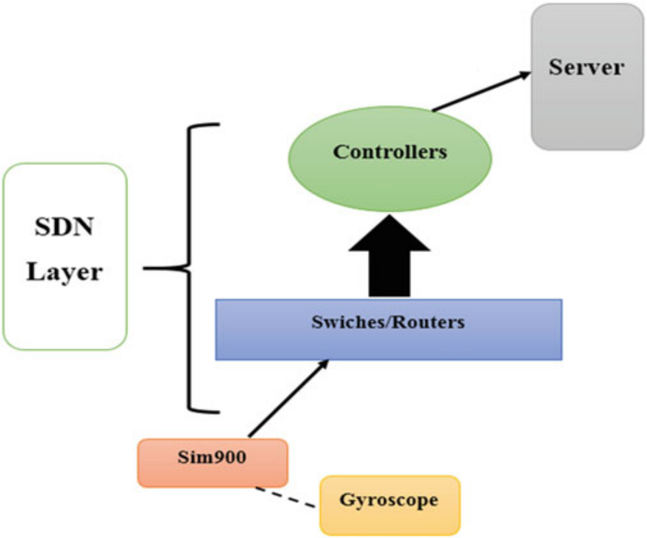


Fig. 3 Transmission of sensor data via edge and SDN

4 Data Analysis

The data received from the gyroscope sensor are saved as a text file and analyzed for detecting movement balance of the person during the movement activities. The respective data are stored in the database for subsequent usage. Unsupervised algorithms are used in order to analyze which movement activity the new data belong to and how much progress and improvement exist in the person's performance. K-Means clustering algorithm was applied for this purpose. Then, the algorithm output is analyzed. The analysis criteria are accuracy and Calinski_Harabaz. Higher value of Calinski_Harabaz index output signifies more suitable number of clusters (k) selected for the respective value. Equation (1) represents the calculation procedure of the mentioned index. B(k) in the numerator of the fraction denotes sum of distances of cluster centers from the center of the whole data. Also, W (k) in the denominator of the fraction denotes sum of distances of the samples inside each cluster from its center.

$$\mathrm{CH(K)} = \frac{B(K)/(N-1)}{W(K)/(N-K)} \tag{1}$$

$$\mathrm{B(K)} = \sum_{i=1}^{k} n_i |c_i - c|^2$$

$$W(\mathrm{K}) = \sum_{i=1}^{k} \sum_{j=1}^{n} |x_j - c_i|$$

Figure 4 illustrates the diagram related to algorithm indexes such that accuracy and Calinski_Harabaz (CH) assume the greatest values in K equal to 5 and 8. In the respective algorithm, K value was set equal to 5 and 8 taking into account the five movement activities. K values are number of separation clusters corresponding to the aforementioned movement activity classes. According to Fig. 4, CH value is maximum at k = 5 and k = 8 and this value is almost constant according to the figure. The grey line and the orange line represent CH value and accuracy respectively. Therefore, as mentioned in the present section, the greater value of this parameter signifies more suitable k value or number of selected clusters.

According to the acquired output in Fig. 5, K values equal to 5 and 8 have done the best separation. In fact, five movement activities have been correctly separated. The present design has been compared with a design in which the smartphone sensors were used. K Means algorithm is used as one of the algorithms under analysis for determining the accuracy value. It was observed that the separation of 5 movement activities had been performed correctly in reference [22] only with k value equal to 5. However, in the present design, it is possible to separate 5 movement activities with different k values (5 and 8) using wearable gyroscope sensor. In fact, different movement activities were separated accurately using

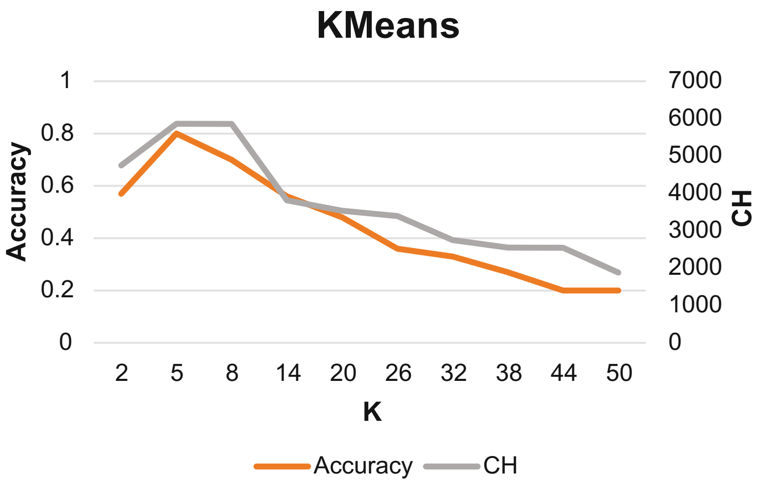


Fig. 4 Evaluation of K Means clustering corresponding to gyroscope data for different values of k

wearable gyroscope sensor in the present design. Also, accuracy is equal to 0.8 for K = 5 and 0.7 for K = 8 whereas accuracy equaled 0.7 only for k = 5 in reference [22]. In Fig. 5, k values equal to 5 and 8 yield the best separation. Actually, five movement activities are separated correctly. The respective 5 movement activities can be observed in 5 different colors in the same figure.

Therefore, the information is sent as real time data via the wearable kit for analysis of the person's status and the results are compared with the previous time in each analysis. If any difference is detected, a variation in the person's behavior can be recognized versus time. As an instance, the information obtained during one week may indicate the person's movement trend. Then, all conducted analyses will be sent to the user or family members via email or SMS.

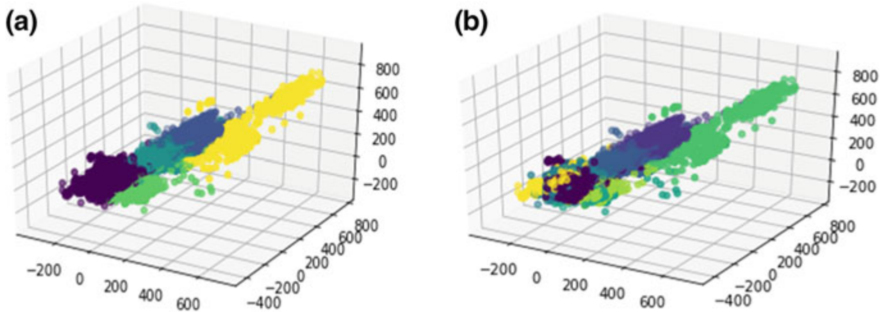


Fig. 5 Different values of k for clustering the movement activities using K Means

5 Conclusion

The objective of the present paper is to propose a mobile smart system in remote rehabilitation aimed at analyzing the person's movement balance. Also, gyroscope sensor was used to assess the person's balance during remote rehabilitation for the respective scenario. The data obtained from the sensors are transmitted through edge layer and SDN controllers to the database in the server leading to reduction of costs, control of network traffic, and mitigation of delays. The transmitted raw data are analyzed using K Means clustering algorithm. The respective algorithm performs the best separation with K values equal to 5 and 8. In fact, accuracy of this method for the 5 movement activities is equal to 0.7 and 0.8 for k = 5 and k = 8, respectively.

References

1. Fakhar M, Behzadipour S, Mobini A (2013) Motion performance measurement using the Microsoft Kinect Sensor, vol. 2, no. December, pp 28–37
2. Patel S, Park H, Bonato P, Chan L, Rodgers M (2012) A review of wearable sensors and systems with application in rehabilitation. pp 1–17
3. Directions F, Baktir AC, Ozgovde A, Ersoy C (2017) How can edge computing benefit from software-defined networking: a survey, use cases. no. c, pp 1–34
4. Satyanarayanan M (2015) of Edge Computing. no. June, 2015
5. Aggarwal C (2016) Securing IOT devices using SDN and edge computing. no. October, pp 877–882
6. Yudi MB et al (2016) SMART phone-based, early cardiac REHABilitation in patients with acute coronary syndromes [SMART-REHAB Trial]: a randomized controlled trial protocol. BMC Cardiovasc. Disord., pp 1–8
7. Postolache G, Carvalho H, Catarino A, Postolache OA (2016) Smart clothes for rehabilitation context: technical and technological issues
8. Vogiatzaki E, Krukowski A (2015) Modern stroke rehabilitation through entertainment
9. Smondrk M, Krohova J, Cerny M, Cernohorsky J (2015) Rehabilitation system for the motion analysis of the wobble board. pp 529–532
10. Begovac J, Šeketa G, Celi L, Lackovi I, Magjarevi R (2015) Quality of human activities measurement from accelerometer data of a smartphone. pp 268–272
11. Miyoshi H, Kimura Y, Tamura T, Sekine M (2015) Smart living—home rehabilitation training system using an interactive television
12. Willmann RD, Lanfermann G, Saini P, Timmermans A, Vrugt J (2007) Home stroke rehabilitation for the upper limbs, pp 4015–4018
13. Postolache O, Pereira JMD, Ribeiro M, Gir P (2015) Assistive smart sensing devices for gait rehabilitation monitoring, vol. 1, pp 234–247
14. Lai C, Lai RHY (2017) An intelligent body posture analysis model using multi-sensors for long-term physical rehabilitation
15. Chen W, Xu Y, Wang J, Zhang J (2016) Kinematic analysis of human gait based on wearable sensor system for gait rehabilitation. J Med Biol Eng
16. Wang Q, Markopoulos P, Yu B, Chen W (2017) Interactive wearable systems for upper body rehabilitation: a systematic review, pp 1–21
17. Editor M (2016) Handbook robotics

18. Kusayev E, Naftali S, Ratnovsky A, Setup AE (2015) Analysis of skeletal muscle performance using piezoelectric film sensor
19. Balbinot A, Crauss J, De Freitas R, Côrrea DS (2015) Use of inertial sensors as devices for upper limb motor monitoring exercises for motor rehabilitation
20. Kulbacki M, Koteras R, Szcz A, Daniec K, Bieda R (2015) Scalable, wearable, unobtrusive sensor network for multimodal human monitoring with distributed control, vol 2 pp 914–917
21. Zulj S, Seketa G (2015) Virtual reality system for assisted exercising using WBAN, pp 719–722
22. Kwon Y, Kang K, Bae C (2014) Unsupervised learning for human activity recognition using smartphone sensors. Expert systems with applications

A Novel Algorithm Developed with Integrated Metrics for Dynamic and Smart Credit Rating of Bank Customers

Navid Hashemi Taba, Seyed Kamel Mahfoozi Mousavi and Ahdieh Sadat Khatavakhotan

Abstract There are a wide variety of algorithms for bank customer credit rating. Over-allocation or under-allocation of credit arises from weakness in algorithms and lack of software programs involving efficient metrics. This in turn gives rise to legal and criminal issues between banks and customers, poor utilization of customer capabilities, and inappropriate provision of banking services. This study intended to propose qualitative metrics to identify the best customer credit rating model with a focus on financial transitions. Instead of focusing on customer credit, this study employed a concept known as discredit derived from the concepts concerning system quality assurance. The new model was validated through efficiently developed software including metric information and customer data. Over the past four years, the account information about 56,000 customers of an international bank branch was studied to determine the criteria and metrics of their credits using different modeling techniques. The developed software was used to define, analyze, and statistically test multiple financial metrics for the financial information of an international bank branch, while fitting the best metrics in a dynamic model for discredit detection. The best coefficients for combination of financial metric were calculated by weighting based on time, while extracting and validating appropriate equations for the newly proposed model. More specifically, the current year account balance was correlated with discredit, whereas the previous year account balances were not correlated. In addition, the discredit data involved a somewhat greater regression than the numerical discredit data.

Keywords Dynamic credit rating · Discredited · Integrated metrics

N. Hashemi Taba (✉) · A. S. Khatavakhotan
Department of Computer Engineering, Islamic Azad University,
Tehran Central Branch, Tehran, Iran
e-mail: Nav.hashemitaba@iauctb.ac.ir

S. K. Mahfoozi Mousavi
Department of Computer Engineering, Islamic Azad University,
United Arab Emirates Branch, Dubai, United Arab Emirates
e-mail: kamel.mousavi@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_60

1 Introduction

From the bank's perspective, the credit of a customer depends on their fulfillment of liabilities [1]. Particularly, clearing of cheques and repayment of loans can be two criteria to allocate credit to a customer [2]. In terms of quality assurance, however, credit violations are more reliable. More specifically, discredit can be demonstrated based on the relationship between the amount of cleared cheques and the amount of bounced cheques.

When a cheque-account customer draws a cheque, money is simply generated. In fact, a customer obtains goods/services by providing a cheque with a credit equivalent [3]. The cheque is supposed to be credible to its minimum amount from the date of drawing until the date of cashing at the bank. When the cheque is bounced on due date because of insufficient balance, it shows that false credit has been assigned to the customer, at least as much as the difference between the current balance and the cheque amount. Nonetheless, there might have been other cheques drawn.

2 Modeling for Dynamic Discredit Calculation

As noted earlier, the banking codes of conduct prescribe that cheque-account balance cannot be an indicator of credit or discredit [4]. That is because the real-time or average balance over a given banking period cannot indicate the issuance of cheques, and the account holder is at full elbow-room. Nonetheless, the account balance is supposed to be sufficient when the cheque is submitted to the bank for collection. However, balance and, in particular, the average balance of an account in a specific interval, e.g. three months, is the easiest way to allocate credit to customers for a new chequebook or loans. It can also leave a loophole for abuse or fraud [5].

2.1 Discredit

Failure to fulfill liabilities leads to discredit [6]. Nevertheless, the decision on credit usually depends on recorded and registered data as well as official actions rather than exchanged information [7]. Specifically, we defined discredit as official action of cheque owners when they are informed that the account balance is insufficient for clearance and request a certificate of absence for judicial authorities. In this perspective, the following metrics can be defined:

N-CoA-x-d: The number of discredit events per day d for customer X (the number of issued certificates of absence)
A-CoA-x-d: The amount of discredit events per day d for customer X (the number of issued certificates of absence)

2.2 Discredit in a Fiscal Year

The number of discredit events for a customer in one fiscal year can be calculated through the following equation:

$$\mathrm{N-CoA-x_{-y}} = \sum_{d=1}^{d=\text{number of working days}} N - CoA - x - d \tag{1}$$

(Number of certificates of absence issued during one fiscal year)

The amount of discredit events for a customer in one fiscal year can be calculated through the following equation:

$$\mathrm{A-CoA-x_{-y}} = \sum_{d=1}^{d=\text{number of working days}} A - CoA - x - d \tag{2}$$

(Amount of certificates of absence issued during one fiscal year)

3 Hypotheses

This study involved three pairs of hypotheses as follows:

3.1 Hypothesis One

H1—There is a significant relationship between average balance and discredit for the number of cheques.
H0—There is no significant relationship between average balance and discredit for the number of cheques.

3.2 Hypothesis Two

H1—There is a significant relationship between average balance and discredit for the monetary amount of cheques.

H0—There is no significant relationship between average balance and discredit for the monetary amount of cheques.

3.3 Model Description

Figure 1 illustrates the conceptual model proposed in this paper. The following steps were sequentially included in the model:

- Data of cleared and bounced cheques by time
- Monetary amount of cleared cheques
- Monetary amount of bounced cheques
- Access to cheque-account database and archives
- Calculating balance for current year accounts and previous years
- Calculating metrics designed for discredit
- Calculating data regression and specifying the significant relationship between metrics

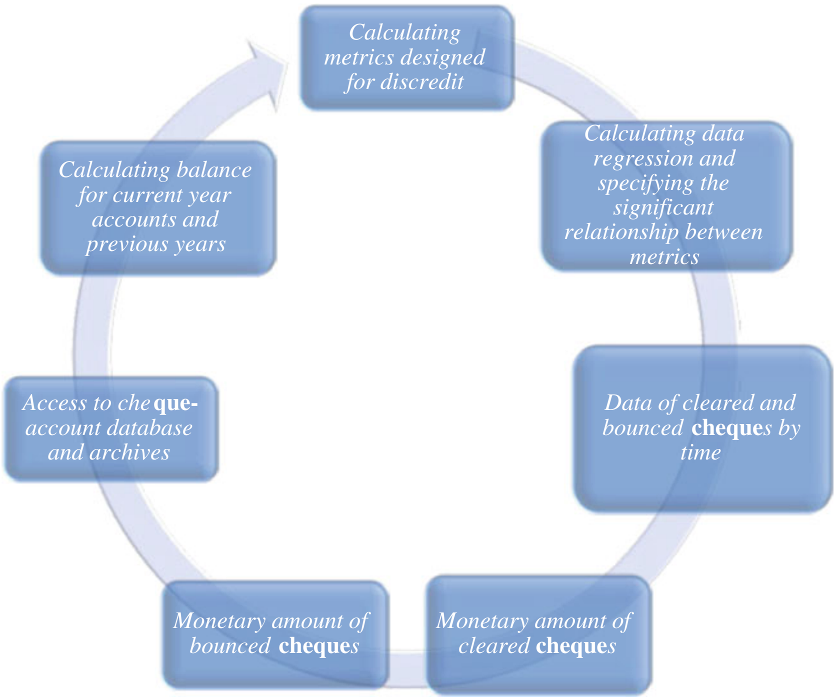

Calculating metrics designed for discredit
Calculating data regression and specifying the significant relationship between metrics
Data of cleared and bounced cheques by time
Monetary amount of cleared cheques
Monetary amount of bounced cheques
Access to che que-account database and archives
Calculating balance for current year accounts and previous years

Fig. 1 The conceptual model for account credit

3.4 Main Input Variables of the Model

We selected a total of 56,000 customers holding cheque-accounts with available records over the past three years. For each customer, the account balance information and the number and amount of cleared and bounced cheques were extracted and calculated. The variables in Table 1 were adopted for calculations.

Table 1 Main input variables

Variable	Description
Bal – Ave – x − 3y	Average account balance of customer X over three years
Bal – Ave – x – py	Average account balance of customer X over the past year
Bal – Ave – x – ty	Average account balance of customer X from the beginning of the year to the study day
N – C – x – d	The total number of cheques submitted to the bank on day d for customer X
A – C – x – d	The total amount of cheques submitted to the bank on day d for customer X
–B – C – x – d	The number of bounced cheques on day d for customer X
A – B – C – x – d	The amount of bounced cheques on day d for customer X
N – P – C – x – d	The number of cleared cheques on day d for customer X
N – P – C – x – d	The amount of cleared cheques on day d for customer X
N – P – C – pctg = N – P – C – x – d / N – C – x – d * 100	Percentage ratio (number) of cleared cheques to total cheques drawn on day d for customer X
N – B – C – B – pctg = N – P – C – x – d / N – C – x – d * 100	Percentage ratio (number) of bounced cheques to total cheques drawn on day d for customer X
A – P – C – pctg = A – P – C – x – d / A – C – x – d * 100	Percentage ratio (amount) of cleared cheques to total cheques drawn on day d for customer X
A – B – C – B – pctg = A – P – C – x – d / A – C – x – d * 100	Percentage ratio (amount) of bounced cheques to total cheques drawn on day d for customer X

4 Dynamic, Weighted Discredit Calculation Model for Cheque-Account Customers

Discredit is calculated through two methods: (1) the number of bounced cheques to the total cheques submitted to the bank, and (2) the amount of bounced cheques to the total amount of cheques submitted to the bank. The calculations covered discredit for chequeing-account customers of the branch holding a chequebook for three years and available banking history of at least three years. Discredit was also calculated separately for the previous year. Finally, discredit was calculated from the beginning of the year to the day of empirical study. There were three separate criteria inserted into the model:

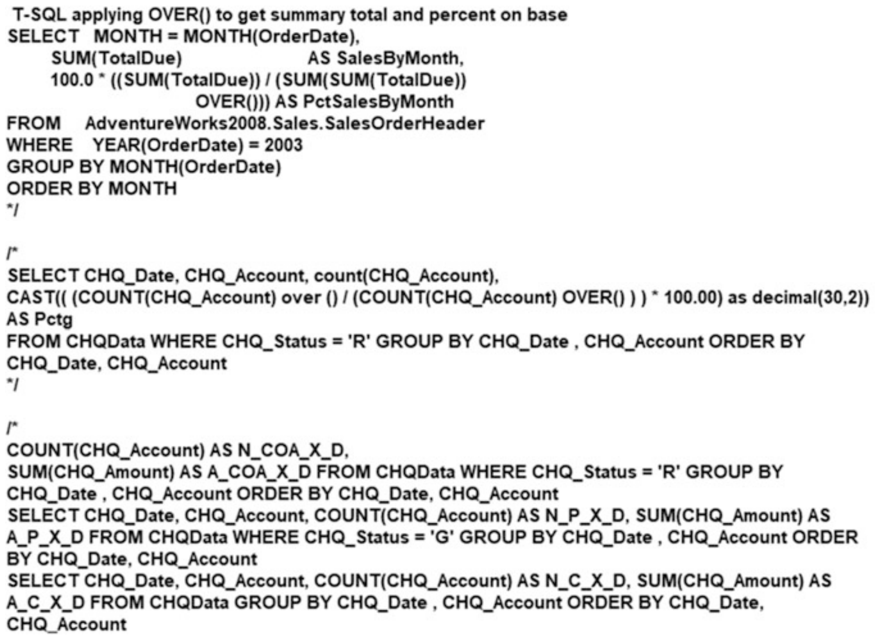
- Three-year discredits
- Recent-year discredits
- Discredit from the beginning of the year to the study day.

The final metric was selected because the fiscal year for individuals and companies are mainly specified by closing the previous year account, opening new ones, and transferring over items from accounts of the previous year. Usually, a separate budgeting is assigned to each year; and resources, costs, and capital are supplied using the new plan.

4.1 Model Validation Through Developing Discredited Software

The presented model was completed and checked through designing new software in which the mentioned metrics were embedded and the cheque-account information was presented without personal information. The information about cheques issued by cheque-account customers, including cleared and bounced cheques were imported into the software, while calculating the values of metrics separately.

The information was compiled into four databases, covering three current and archived time periods. Several fetches were run through SQL Query in Discredited software developed in this phase (see Fig. 2) until essential data were extracted from numerous tables involved in each database. There were no system tabs in databases. Fetches were run on log files which were extremely time-consuming. Most core banking systems do not store but calculate balances. Therefore, the data sources were adopted to make the essential calculations for balance inquiry. A more precise average balance was obtained by accumulating the individual day balances instead of dividing the difference between the first day and last day of the period by working days.



```
 T-SQL applying OVER() to get summary total and percent on base
SELECT   MONTH = MONTH(OrderDate),
       SUM(TotalDue)                    AS SalesByMonth,
       100.0 * ((SUM(TotalDue)) / (SUM(SUM(TotalDue))
                               OVER())) AS PctSalesByMonth
FROM    AdventureWorks2008.Sales.SalesOrderHeader
WHERE   YEAR(OrderDate) = 2003
GROUP BY MONTH(OrderDate)
ORDER BY MONTH
*/

/*
SELECT CHQ_Date, CHQ_Account, count(CHQ_Account),
CAST(( (COUNT(CHQ_Account) over () / (COUNT(CHQ_Account) OVER() ) ) * 100.00) as decimal(30,2))
AS Pctg
FROM CHQData WHERE CHQ_Status = 'R' GROUP BY CHQ_Date , CHQ_Account ORDER BY
CHQ_Date, CHQ_Account
*/

/*
COUNT(CHQ_Account) AS N_COA_X_D,
SUM(CHQ_Amount) AS A_COA_X_D FROM CHQData WHERE CHQ_Status = 'R' GROUP BY
CHQ_Date , CHQ_Account ORDER BY CHQ_Date, CHQ_Account
SELECT CHQ_Date, CHQ_Account, COUNT(CHQ_Account) AS N_P_X_D, SUM(CHQ_Amount) AS
A_P_X_D FROM CHQData WHERE CHQ_Status = 'G' GROUP BY CHQ_Date , CHQ_Account ORDER
BY CHQ_Date, CHQ_Account
SELECT CHQ_Date, CHQ_Account, COUNT(CHQ_Account) AS N_C_X_D, SUM(CHQ_Amount) AS
A_C_X_D FROM CHQData GROUP BY CHQ_Date , CHQ_Account ORDER BY CHQ_Date,
CHQ_Account
```

Fig. 2 A sample of SQL script using in this study

4.2 *Discussion*

We selected the relationship between the average balance and the percentage of discredit was one of the most important issues explored in this paper [8]. Therefore, the model included and sorted the relationship between average balance in three years and percentage of discredit over three years, the relationship between average balance over one year and percentage of discredit over one year, and the credit of high-ranking and low-ranking customers. The ratio of these rankings to the values of metrics was inserted into the following mathematical models below and the corresponding charts were discussed.

Table 2 was derived from the data on average balances for the current year, previous year and past three years by bounced cheques. The results indicated that the average balance of people who had at least three bounced cheques from the beginning of the current year until the research period was 50% less than the average balance of those with one bounced cheque and 35% less than those who had two bounced cheques, which is significant. The ratio was much stronger for previous year in that the average balance of people who had at least three bounced cheques was 95% less than the average balance of those with one bounced cheque and 55% less than those who had two bounced cheques. The ratio does not suggest a significant relationship for the past three years. It implies that the recent data has more weight and is the current criterion of the account holders.

Table 2 The relationship between the number of bounced cheques and the average balance

Current year		Previous year		Past three years	
Title	Amount ($)	Title	Amount ($)	Title	Amount ($)
The average balance of those who had a bounced cheque in the past three years	136,127.38	The average balance of those who had a bounced cheque previous year	241,816.18	The average balance of those who have a bounced cheque bcurrent year	101,682.82
The average balance of those who had two bounced cheques over the past three years	119,602.02	The average balance of those who had two bounced cheques previous year	77,117.47	The average balance of those who had two bounced cheques bcurrent year	695,067.34
The average balance of those who had at least three bounced cheques over the past three years	73,699.08	The average balance of those who had at least three bounced cheques previous year	12,196.39	The average balance of those who have at least three bounced cheques bcurrent year	692,653.06

As shown in the charts (Figs. 3 and 4), there is significant relationship between the number of bounced cheques, discredit, and account balances. This relationship was insignificant for all three periods of the current year, the past year, and the last three years. Nonetheless, there was a significant relationship between balance and amount of bounced cheques in all three periods. Furthermore, there was a significant relationship between discredit and account balances for the last period of current year.

Focusing on the "bounced cheque to the average balance" metric led to significant results and relationships. To this end, three levels of balances were considered. The balances were divided into deciles and customers were identified at each level. The first three deciles were regarded as low level, the three upper deciles as high level and the three middle deciles as medium level. By pressing the necessary scripts at each level, the amount of each customer's bounced cheques in the current year, previous year and past three years was determined. Table 3 contains the results of the metric data.

The hypothesis of inverse relationship between the average balance and the bounced cheque rate was confirmed. Figures 5, 6 and 7 are pictorial view of

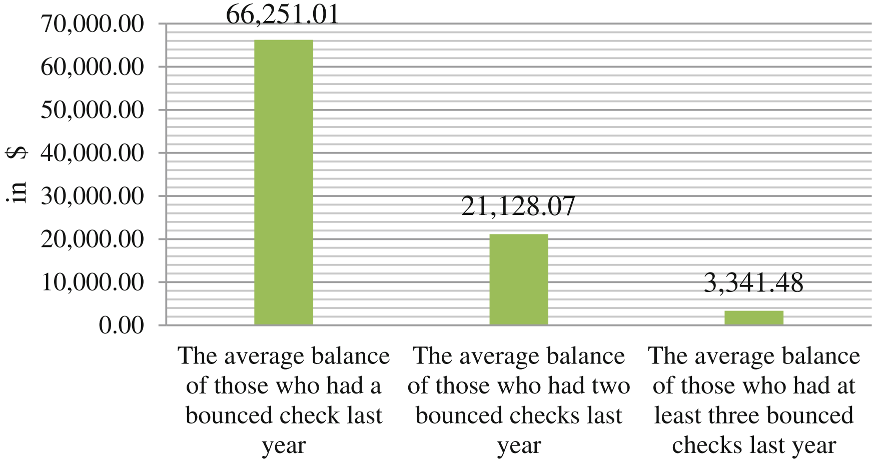


Fig. 3 The relationship between the number of bounced cheques and the average balance for previous year

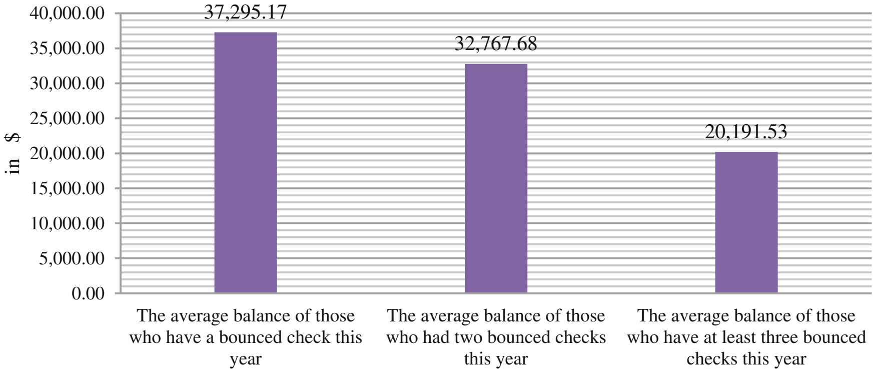


Fig. 4 The relationship between the number of bounced cheques and the average balance for current year

relations in form of diagrams. The amount of the bounced cheques for low balances bcurrent year was seven times higher than that for medium balances in the same year and 80 times higher than the amount of the bounced cheques for the upper-level customer balances. Similarly, the amount of the bounced cheques for lower balances in the previous year was 3.5 times higher than that for the medium balances in the previous year and 25 times higher than the amount of the bounced cheques for the upper-level customer balances. Finally, the amount of the bounced cheques for lower balances over the past three years was 6 times higher than the amount of the bounced cheques for the medium balances in the past three years and 35 times higher than the amount of the bounced cheques for the upper-level

Table 3 The relationship between the amount of bounced cheques and the average balance

Current year		Previous year		Past three years	
Title	Amount ($)	Title	Amount ($)	Title	Amount ($)
The amount of bounced checks for high balances in the current year	8842.24	The amount of bounced cheques for medium balances in the current year	91,187.03	The amount of bounced cheques for low balances in the current year	649,052.13
The amount of bounced checks for high balances in the previous year	31,592.96	The amount of bounced cheques for medium balances in the previous year	218,491.87	The amount of bounced cheques for low balances in the previous year	743,305.28
The amount of bounced checks for high balances over the past three years	67,336.26	The amount of bounced cheques for medium balances over the past three years	455,892.33	The amount of bounced cheques for low balances over the past three years	2,554,846.39

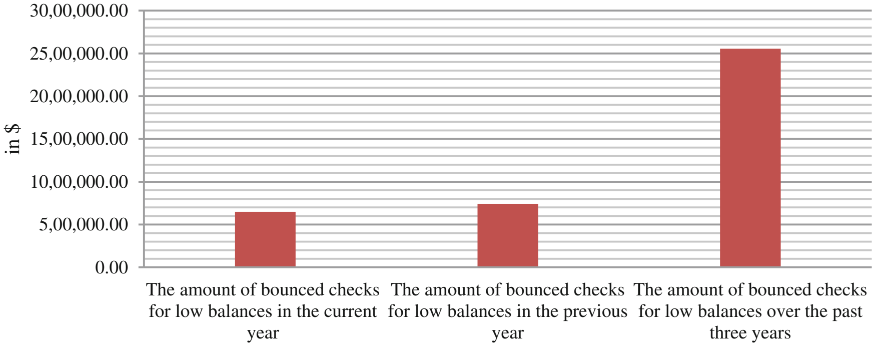


Fig. 5 The relationship between the amount of bounced cheques and the average balance for Low balances

customer balances over the past three years. Therefore, the average customer balance is negatively correlated with the bounced cheque rate with a statistical strength of tens of times. Unlike the relationship between the number of cheques (and the bounced cheque rate) that was measured with the previous metric, the relationship between balances and the amount of the bounced cheques over the past three years

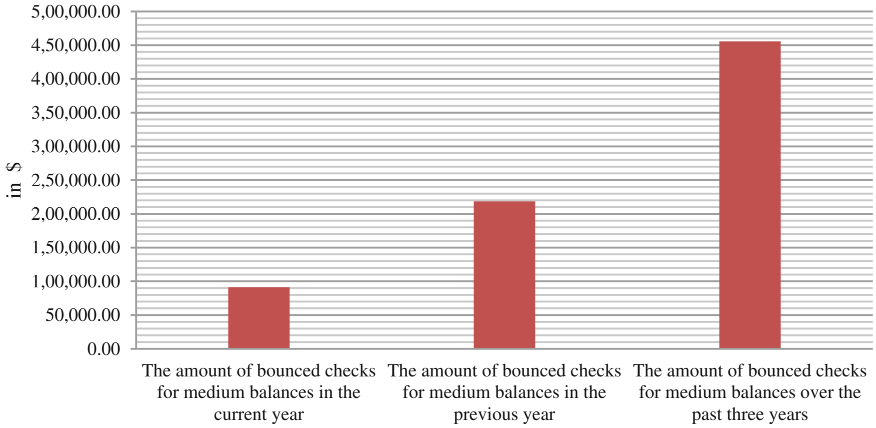


Fig. 6 The relationship between the amount of bounced cheques and the average balance for Medium balances

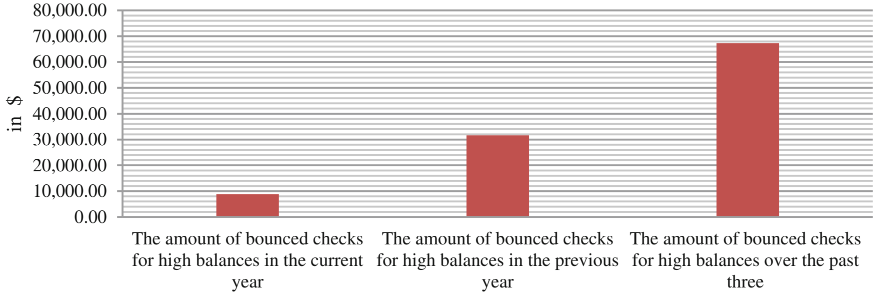


Fig. 7 The relationship between the amount of bounced cheques and the average balance for High balances

is stronger and more significant than the previous year and current year. Therefore, the average balance over the past three years can be considered a criterion for customer credit.

5 Conclusions

The information obtained in this paper suggested that the average discredit over past years was not significantly correlated with current credit of customers. In fact, the average amount of discredit of past year and discredit from the beginning of the year could provide a better criterion. Moreover, the number of cheques was not significantly correlated with customer credit, unlike what is usually applicable. In fact, the amount of discredit provided a more important criterion than the percentage number of discredit. Finally, the new model can be adopted as a desirable criterion for granting financial facilities to bank customers.

6 Future Studies

The information Considering the important points and recommendation in this section will pave the way to continue the research from different supplement aspects.

6.1 Important Recommendation for All Banking Systems

We selected The field studies in this paper indicated that cheques exchanged through the banking system between banks and branches are fully recorded and traceable. When individuals visit the banks to cash the cheques and there is insufficient account balance, however, information is not recorded in most cases for defective or incompatible cheques. In this scenario, a certificate of absence is issued and recorded if the customer requests one. However, nothing would be recorded and the bank would remain unaware if the customer decides to personally contact the drawer to increase account balance or resolve other issues. It is strongly recommended to formulate a code of conduct where the information of each bank-submitted cheque is recorded [9].

6.2 Limitations and Delimitations

The modeling could be improved by some information unavailable to the bank. For example, the interval between cheque issuance and bank delivery (due date on the cheque) is unspecified even though it can be an important criterion to determine customer credit. This can be achieved through self-declaration, which tends to be unreliable [10].

6.3 Model Improvement by Inserting Other Effective Variables

The future studies can be more accurate by covering the loans granted, and the repayments with/without delay. Also by presenting a customized method similar to that of the current paper, the amount of discredit in repayment of loans can be obtained and examined. Ultimately, a combination of discredit for cheques and discredit for repayment of loans can double the validity and reliability of customer classification in terms of credit [11]. For this purpose, it is recommended to adopt integrated modeling together with statistical tests involving real information as presented in this paper.

References

1. Hahm J-H, Shin HS, Shin K (2013) Noncore bank liabilities and financial vulnerability. J Money Credit Banking 45(s1):3–36. https://doi.org/10.1111/jmcb.12035
2. McHugh S, Ranyard R (2016) Consumers' credit card repayment decisions: the role of higher anchors and future repayment concern. J Econ Psychol 52:102–114. https://doi.org/10.1016/j.joep.2015.12.003
3. De Muynck M (2011) Credit cards, overdraft facilities and european consumer protection: a blank cheque for unfairness? SSRN Electron J. https://doi.org/10.2139/ssrn.1970212
4. Obstfeld M (1984) Balance-of-payments crises and devaluation. J Money Credit Banking 16 (2):208. https://doi.org/10.2307/1992546
5. Bad Credit? No Credit? (2003) Comput Fraud Secur 2003(3):10–12. https://doi.org/10.1016/s1361-3723(03)03009-4
6. Discredit to whom discredit is due (1902) The Lancet 160(4120):453. https://doi.org/10.1016/s0140-6736(01)41625-4
7. Gong GM, Huang JV (2014) Research on the bank's credit decision making based on the exploration data analysis. Appl Mech Mater 644–650:5721–5724
8. Varol İyidoğan APDP (2014) The analysis of the causal relationship between budget balance and current account balance by MGARCH methodology: Turkey experience. Anadolu Üniversitesi Sosyal Bilimler Dergisi 14(2). https://doi.org/10.18037/ausbd.62277
9. Mehta M, Sanchati R, Marchya A (2010) Automatic cheque processing system. Int J Comput Electr Eng 761–765. https://doi.org/10.7763/ijcee.2010.v2.224
10. Li J, Zhu Y (2016) Combined forecasting model based on cuckoo search algorithm for personal credit assessment. J Softw Eng 10(3):297–301. https://doi.org/10.3923/jse.2016.297.301
11. George A (2016) The past of the pay cheque. New Sci 230(3079):31. https://doi.org/10.1016/s0262-4079(16)31144-7

Data Mining Based on Standard Analysis

Ali Saberi

Abstract Data investigation in any type of database or any smart systems intends for achieving sound and practical information, so that data analysis can be used for extracting defined models and patterns, which can be generalized and utilized, and such patterns and structures are discovered and perceived that can be used as instruction or application in similar structures. In fact, data mining is important because working process of a system can be perceived by analysis of data content, and useful and constructive information can be obtained by standard and expert analysis. Data mining can be defined at any smart system or interactive and communicative structure that has outcome and different results, and data analysis can help to develop performance of any system with any method and working purpose in interactive and competitive space. Utilizing data mining science is a factor for producing better content, performance improvement, and achievement of optimal outcome and result. Thus, standard analyses are able to provide sound information about activities because they can properly identify and investigate nature of the activity and data. This information are practical and security, quality, and quantity of data as well as performance can be improved using these information, new working patterns and models can be formed. Data mining actually based on standard analyses causes development of data-dependent processes.

Keywords Data mining · Smartization · Data analysis · Organizational intelligence

1 Introduction

Data are important because the information is shaped by their analysis, and represent activity status of a system based on specific and defined indexes. Data can be a set of commands and codes, which are given to a software or smart data system,

A. Saberi (✉)
Iranian Researchers Network, Saqqez, Iran
e-mail: sharemjazi@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_61

and information, are obtained from these data according to a programming command, which is practical. That is, both the software and smart system performs a specific action based on these data and the users and developers achieve their goal, i.e. an interaction is established based on the data. A smart interaction actually forms a databased structure. Data can also be defined in other way: at a specific level, data can be analyzable information. That is, if the set of activities of a store system is investigated, some patterns and results can be obtained considering trend of sale and welcome to the products, and accordingly solutions can be offered for progress of the activities. Perhaps the data alone do not represent a fixed indicator based on which conclusion can be made, but if the practice based on the data were standard and systematic, data mining would be an effective action in the system, and would bring about effective outcomes. If data mining is utilized in progress of activities and in smart and multimedia communications, where the data content is important and communication is shaped based on which, effective and reliable results would be achieved that develops working model. Therefore, effective role of data is an analyzable subject that requires knowledge and expertise. It is important both in virtual smart environments and in other activity structures, which follow technology or patterns other than virtual space, and data and information in this types of organization are obtained from data investigation. Anyway, the purpose of data and information analysis is acquiring the knowledge, which develops the activities.

2 Statement of Problem

In the past, the process according which the trend of activity of a collection was informed was the statistics science. That is, based on the statistics from different parts of a collection an outcome was concluded. Using analysis of statistical data and information, both the way of activity could be expressed and solutions could be provided for improvement of the conditions. Thus, making statistics from the way of activities was common and effective. However, the way of activities and required tools changed by mechanization and smartization of interaction and communication space and using electronic and computer sciences, and statistics for activities of a collection were provided in a smart and electronic manner. To this end, information are developed in various ways, and information storage spaces is cloud space and databases that store collection of different information and data, and they are used in different ways. Thus, it can be stated that data mining is the sciences that extracts necessary information and knowledge from large database. Therefore, data mining somehow leads to optimization of interactions of websites and software contents. In line with data mining, expert analysis is required considering all effective elements in data science, so that the analyses lead to constructive generalizable and utilizable patterns and they can be used in similar structures.

3 Description of Problem

Data includes knowledge, awareness, possessions, statistics, thoughts, signs, symbols, imaginations, …that if are understandable to humans, then the data contain understandable content and a kind of information, as in the past, the symbols and signs were used to express a kind of reality or to communicate. They were the signs and patterns that expressed a certain meaning, and certainly the type of use and application of these data was limited to those times and conditions. However, new data has become meaningful with the advancement of technology and transformation of thoughts and life styles. It can be said that data is a product of different processes. Science growth, impact of knowledge in life, formation of innovations with specific capabilities, each contributes to the production of different data. For example, before emergence of intelligent technologies, statistics from individuals and collecting data were used for making conclusions and providing hypothesis. Thus, they concluded about the situation. Alternatively, they could achieve knowledge and awareness that could understand the information, use it in different structures, and develop human insight. Therefore, this was the process of collecting data and using it in a specific activity. In smartization and the computer science, which is the subject of the article, the data have the same meaning and nature. Only their content and method of use have been changed. Data are generated in other ways and are used in more diverse ways, and certainly have different effects and results. Data are used in programming and computer knowledge, that is, a program can written using codes of a particular language, which has content and application. Another type of data usage is a variety of software for data analysis. For example, a software such as SPSS, which analyzes the data entered by the user and then quantifies the result. That is, using the structure of a software, information can be obtained. Therefore, the method of acquiring information has turned to an intelligent method from physical method, and one can reach the results using software. In fact, it is a kind of data usage. For example, when a software is made based on data in programming language and has specific application, here there is a kind of data mining, or the outcome of data mining has been transformed into a comprehensible and communicative software or content with the programming commands. In this case, these data for a specific programming software are meaningful, and these commands and codes, which made the software, are not meaningful in the other software, and do not generate information. In other words, they do not generate the same information, content, or application software, thus data have extensive meanings, and data mining is the science with various results and outcomes. The other type of data usage is in software structures. That is, data and information are defined in the store or production process in a goods sale software. It means that the sale processes represent data about status of activities, and if these data are investigated and analyzed in standard manner, the experts will achieve information that has special meaning. Thus, it is a kind of data mining, and yields a result and model, which includes indexes, information, and values. Based on these information working development can be obtained. For example, sale can be increased purposeful and effective promotion can be developed.

3.1 Purpose and Nature of Data Mining

Data mining is not merely data collection; it also covers data management, information analysis, and prediction or conclusion. Thus, data mining is a useful phenomenon for advancement of information and communication technology as well as software engineering and multimedia files based on the content and user interaction. Data are meaningful signs, which if they are systematically selected and utilized, as the goal of data mining, then they lead to novel technologies in computer knowledge-based technology era. The top ideas and thoughts are the major factor for changes and advancements in science in different areas. In fact, the ideas and assumptions are also regarded as a kind of data. When the individuals grow and prosper by creativity and knowledge, and make a change, they turn into useful information. This useful information can be an intellectual model with specific structure or a software with practical capability, each of which are obtained from data analysis, and create a result. Hence, the goal of data mining is discovering knowledge through data analysis. If the analysis were done in an expert manner, a top and better knowledge would be achieved. For example, if information of the community's doctors were used for generation of a medical software, systematic and proper data would be selected. That is, the basis for data selection would be correct. If one who makes the uses these information properly in the program, then the respective result would be obtained. In fact, data analysis or data mining is valuable because develops sound and practical information. For example, assume that various people are researched about a social problem, and obtained data represent the reality of the problem, then if the social experts can acquire reliable interpretation and hypothesis from the obtained data and perform standard data mining, such results can be achieved that would be effective in improvement of the community's condition. Data mining is used in society security, education, banking, marketing, e-commerce, smart webs, software, etc. in different ways. Especially with presence of intelligent systems in different interactions, data storage and management is systematic and scientific ways, so that it leads to applied knowledge generation, is crucially important, and the society security and social development is somehow dependent on it, and it would be realized as a result of proper data analysis.

3.2 Importance of Content in Information Generation

Standard analysis of data generates sound information. These content includes image, voice, film, text, etc. Content is generated and disseminated in a software or website and social network. In fact, the content is also a kind of information or data. For example, the software, file, or website that provides medical information, expresses texts, image, film, and other information types, is regarded as a kind of content, and it is clear that if the contents are sound and standard, the efficiency and efficacy of the program will be higher. That is, supervision of content and information generation leads to production of healthy and sound contents in virtual space. Because of virtual influence in different interactions and higher access of individuals to this technology, the content can be influential in various aspects, certainly sound and healthy contents have positive impact, and unhealthy contents have adverse impact in different parts. Having proper content and information promotes quality of software programs. For example, if an advertisement website publishes proper information through various ways, it causes that many people refer to this website and use it, and it means progress and success of the website, because its content and information are accurate and expertized, and since it has systematic and standard structure, the people tend to visit the website. Thus, since the website data and information are selected based on proper analysis and data mining, it means that the information contains knowledge. It causes progress and excellence of the website, and data mining process in website is satisfactory. Security is one of the indexes of information storage. That is, structure of software, websites, and intelligent systems should be so that information and data are transferred and stored in a secure space. For example, if a software is properly designed and planned and security tips are considered in it, it is an advantage for the software. Obtaining security and reliability of the program depends on various aspects. In fact, selection of a secure environment, which results from data analysis, is a success in the program working process, because such security is achieved when the process is specifically investigated so that data and information are delivered to the users through a secure path. Such secure path is obtained because of analysis of different situations based on acquired information. That is, if data and information were not analyzed in a standard manner, there would be no safe program or website. In addition, the pattern which is used in the structures should be based on the healthy and specialized content. Thus, all factors influence in the process so that its security and efficiency is acceptable. Therefore, supervision of information and content quality and quantity is an index for valuing the data. In fact, data analysis should cause generation of useful content and information, and when data mining is done based on standard analyses, proper information and contents are selected.

3.3 Organizational Intelligence for Data Analysis

Data, in addition to virtual space, can also be defined in a collection. For example, when working trend of an organization is to be investigated, there should be awareness about the way and structure of its action. Obtaining accurate information and statistics about the organization's function helps to gain awareness about its situation, and results and predictions can be provided by information analysis, and at advanced level, a pattern can be proposed according to the acquired theories. Since change in structures of systems and institutions of the society and formation of new knowledge such as organizational management, knowledge management, psychological management, human resource management, living quality, and other cases represent role of different intelligence factors in the working process of a collection, all activity aspects and effective elements are considered in organizational intelligence, so that a sound and dynamic system can be achieved. Thus, different activity aspects of the organization are investigated by data mining, and necessary information are obtained, and these information with standard analysis leads to the solutions that put the system at optimal status, human resources act directed, the goals are defined for the collection, excessive costs and capital loss is prevented, and other cases, which are result of reporting of activity process in the organization. Hence, providing statistics according to data and its analysis and obtaining information that are expressed in the form of model and pattern are parts of organizational intelligence structure. For smart activity, the collection should investigate status of its activities and acquire such knowledge from it that helps organizational development.

3.4 Database Optimization

Data gain different nature given the type of activity structures. For example, in an organization or store, data express the information acquired about the way of activity, and certainly effective storage, management, and usage of these data requires the knowledge and proper analyses so that a practical model and principles can be achieved. However, in information technology space and given the electronic nature of services in this space, the data have a nature matched to this structure, and thus information management and security should be done in line with the respective knowledge. Data not only need to be managed and analyzed, but also they should be stored safely. Information in software and computer systems is stored in a space known as database. Database is made by the software. That is, information are entered into the tables and forms of the software-made database and stored there, and the prepared file is used in software or websites and smart systems. Thus, since database is the place for information storage, database loading in the website or software would be difficult in case of high load or illegibility of commands. For example, the website may face data traffic, work slowly, or do not act properly. The software and websites with such problems certainly have fewer users, and it is considered as a defect. Therefore, the importance of optimizing the database is essential. Deleting excessive information, using the necessary data and information, using low-volume information, using a high RAM for servers, and similar practices are used in order to optimize the database in computer science-based systems, and optimizing the database is a step towards the correct and effective use of the data.

4 Discussion and Conclusion

Achieving logical results, providing organizational models based on results and reports, using these patterns and theories in generalization and development of different structures are main goals of data mining. Thus in order to acquire practical knowledge from the analyses, the criteria, and indexes of data mining should be defined according to expert and standard foundations. Since the knowledge obtained from statistics and performances should be practical and influential knowledge, and in order to obtain constructive and effective knowledge, the analyses should be expert. Therefore, data mining analysis is significant to achieve respective result and goal. It is important both in terms of data type in virtual space, information, and findings in other organizational structures. With development of technology, need for data analysis is crucially important so that with offering effective patterns, activities can be progressed in competitive space at global level, and obtain leadership and systemic power. Thus, data development and data analysis is an action for development of the society in different aspects.

References

1. Saberi A (2017) Studying main indicators in an optimal standardization. J Sci Eng 2(2)
2. Saberi A (2016) Knowledge management, second international management and economics conference in the 21st century. Allameh Tabatabai University, Tehran
3. Saberi A (2017) Design, production and development of electronic content, second international conference on knowledge based research in computer engineering and information technology, Allameh Tabatabaei University
4. Saberi A (2017) Path of software development in intelligent systems. In: Second national conference on modern research in electrical and computer engineering, Basir Institution of Higher Education
5. Saberi A (2017) Organizational Management, Islamic Republic of Iran Humanistic Magazine, ISSN: 2588-2635
6. Saberi A (2013) Management in information technology. Comput Sci Res J. ISSN: 2584-2174
7. Saberi A (2017) Multimedia in cyberspace. Sci J Comput Sci Res. ISSN: 2584-2174
8. Saberi A (2017) Technology management in globalization. Sci Res J Comput Sci. ISSN: 2584-2174
9. Saberi A (2017) Cognitive science in artificial intelligence. Sci J Comput Sci Res. ISSN: 2584-2174
10. Saberi A (2013) Human resource management. Sci J Manage Econ Account. ISSN: 2783-2588
11. Saberi A (2017) Management in smart technology. In: Third annual management conference and business economics, Tehran
12. Saberi A (2017) Health and electronic health. In: National conference on new accounting and management researches in the 3rd millennium
13. Saberi A (2017) E-learning and social knowledge development. In: Annual conference on research in humanities and social studies

14. Saberi A (2017) Importance of advertising programs in the age of development. In: Annual conference on research in the humanities and social studies
15. Saberi A (2017) Using intellectual capital with standard strategy. In: Third annual management conference and business economics, Tehran

Development of Software with Appropriate Applications in Smart Tools

Ali Saberi

Abstract Technology development has caused that smart tools to become part of interactions and communications, that is, the smart systems have developed along with the human activities. Thus, smart tools have become service providers by utilization of developed computer and electronic structure. Smart tools are composed of a hardware system. This hardware system, using an operating system or under several operating systems like Microsoft Windows, Android, Linux, IOS, Windows Phone, etc. develops a smart tool like cell phone, computer, and other smart tools. In order to promote the interaction between smart tool and user or to provide more facilities for the smart tool, certainly software appropriate to the operating system and smart tool is used. This process of change and advancement is crucially important in the current technological era, and the better the tool is able to communicate, it is considered as an advantage and superiority. Various software with different applications plays the major role in the superiority process. If the smart tool enjoys side facilities or suitable software, tendency to use it would be increased because of its usefulness and efficacy. Hence, developing practical software with healthy and effective contents is a major issue in software design and production, which would lead to more capabilities and facilities for the smart tools, and people, would be more tended to use them.

Keywords Software · Smart tools · Markets · Society development

1 Introduction

With change in interaction and communication structures at different levels as well as practical role of science and skill in the modern technologies, full-scale development in different scientific areas is observed. Novel structures of scientific technologies have become emerged and life styles have been influenced consid-

A. Saberi (✉)
Iranian researchers network, Saqqez, Iran
e-mail: sharemjazi@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_62

erably. Scientific development index is the process that defines meaningful and effective progress for technologies. In fact, in order to use the knowledge or expertise in optimal way and consistent with other scientific changes, it should be developed, and optimal results should be obtained from the knowledge. Development, with reliance on fundamental concepts of the science, attempts to pursue a purposeful growth, and makes the sciences as applied. That is, the science should be able to serve individual and social needs, and it can be turned into applied from theoretical science. For example, transformed economy or transformed education utilizes technological structures of virtual space. Thus, using management, information technology, computer, and economy gives birth to electronic commerce science, and it can be considered as outcome of development of different sciences. If this meaning of development is to be defined in smart tools, it means interaction of different knowledge for creating smart structure in the tools. In technology development discussion, computer science plays basic role, so that different sciences are required to utilize computer science, information and communication technology, and electronic for smartness and achievement of optimal efficiency. Virtual space, interactive environments, and smart tools use this science in different ways. Thus, it can be stated that software development includes comprehensive and general development of different scientific sectors. For example, a software that is able to perform banking services, causes change in the economy and business science; a multimedia educational software transforms learning and education process. Many examples of this case can be enumerated in the current era, which have been formed and are currently growing and developing. It can be said that action for software development is acting for a full-scale and comprehensive development, and it has different effects and outcomes. This structure can be perceived by studying and acquiring knowledge in this regard, and it is an essential reason for significance of software development. Specialized and purposeful planning is required for realization of this process.

2 Statement of Problem

Lack of attention to factors affecting the society development causes that the society is damaged from different aspects, and absence of optimal social conditions is due to negligence to knowledge development. Developed communities and deprived communities are examples of paying attention and negligence to technology, knowledge, awareness, and social welfare. Successful and advanced communities certainly have standard programs, and attempt to achieve the respective goals, and its outcome is such a social environment with facilities, security, and welfare. In contrast, the communities that do not care for the growth and development process have turned into the communities that are deprived of the minimal facilities and have problems and difficulties in different aspects. Hence, we should step in the global development path with such attitude, and we should be aware that negligence and not using abilities and talents causes failure of the individuals and community, and consequences of such negligence to progress and development obviously requires investigation. The reason for expressing significance of development and progress is paving path for development and growth in the society. To this end, a standard structure should be followed, which has plan for talents, resources, and abilities of the society. In the context of expert planning, society is flourishing and the ground for scientific advancement is achieved in different fields. When targeted society is on the path to technology development, definitely all sectors of society will be affected, and in the discussion of computer technologies, which software production is part of it, these actions should be more specialized and based on knowledge and skills. Smart tools such as computers, mobile devices, and other operating system-based tools require software to make tools and operating systems more efficient. Designing, manufacturing, and distributing software is a process for generating a software to provide it to the user, and If we want to properly underground this path of development, so that content and application programs are produced, then we have to take steps in this direction. The use of national software is an act that transforms both the community and its talents. It also increases the software power or the ability to create useful software. Therefore, development of software, considering changing technology space and tendencies to use this science in communications and interactions, requires that such a process be developed in societies, in order to provide the path of transformation and community development, and the individuals and systems of the community will benefit from this technology in different sectors.

3 Statement of Plan

The cases that should be taken into account in production of software can be investigated from different dimensions. Part of which is related to the aesthetics and design of software, which should be in such a way that is suitable with the software application. It should possess great graphics, and in terms of the size, it should need low space. The other point is quality and content of the program. Healthy, specialized, and effective content should be considered in production of the software. The content should be designed appropriate to the type of tool, such as cell phone, or laptop, or other smart tools. On the other hand, given different applications of tools, their software should also meet this need. The other important case in software production is that their providers, i.e. Software markets, should offer software variety, and choose accurate options so that the users use their programs ensuring proper application of them. Given changes in technology area, tendencies are mostly toward such tools as cell phone and tablets. Thus, paying attention to the designs appropriate to these highly used smart tools is constructive for software development in the smart tools. National knowledge and production of national programs should be taken into account. It is also needed that software localization is implemented and planned properly; so that both the community talents are used and constructive steps are taken in national culture.

3.1 Internet of Things in Smart Tools

Internet of things is one of the changes or developments in the software area, which is applied in appropriate tools. Considering the fact that most tools use smartization process, thus Internet of things can have many applications, and it can be a change in line with software development or smartization of tools. When a tool, such as a home machine or device, communicates part of its functional structure with the user in a smart way, Internet of things is shaped. If Internet network, i.e. Software interface, is used in the structure of tools for communication, reception, and transmission of data, it is smartization under Internet network. The tools are able to communicate with the network and user using an administrative software system. Therefore, if smart communication without Internet ground is to be considered, object-oriented software engineering gets important. For example, such a tool as cell phone performs the user demand by a command. Thus, the software can be designed with and without internet network, in such a way that they can appropriately communicate. Therefore, objectivism means that a tool interact with the user under software structure, and properly performs human demands, for example, search in smart televisions using audio is a kind of smart structure. That is, such a tool as television communicates with the voice of the user, checks it out, and executes the request. Smartization of a tool is done for interacting and executing a request, and there are certainly plenty of these kinds of smart communication in tools. With the growth of technology and use of software engineering, more capabilities can be observed in the smart tools.

3.2 Software Progress

One of the factors causing more use of smart tools especially cell phones and tablets is presence of software applications, which are defined and constructed for these tools and operation systems. Because of tendency of people toward these tools and their suitable physical structure, developers of mobile applications and the companies acting in this field produce numerous numbers of software, because they have variety in content and application. Therefore, mostly demands and tendencies of individuals to different subjects and various applications are designed in the form of software. It should be noted that software engineering technology, in addition to smartization using Internet of things, can act as a database in various subjects, for example, a dictionary, novel, educational material, or any information content can be provided as the software. Hence, in addition to progress in this regards, software structure requires fewer hardware facilities and it is more optimal in terms of cost and virtual space saving. Therefore, software development should be analyzed with such attitude so that appropriate actions are taken for its development. Since the goal of science and technology is welfare, security, and other credible indexes, then if software engineering is developed in a community, it would certainly benefit also in cultural, economic, and political dimensions. The power of income generation and the influence of cyberspace in the present era imply this issue. Therefore, it is necessary for governmental institutions to take the necessary measures in identifying specialists and researchers, as well as providing the necessary facilities for the development of software technology in various aspects. It is imperative that the development of cyberspace and software engineering is a step towards the modernization and advancement of society in different fields, so that any action in this field is appropriate and in line with the progress of society.

3.3 The Role of Markets in Software Development, Identifying Experts and Attracting Diverse People with Various Tendencies

In the development of software, in addition to paying attention to the necessary steps in production of standard software, it is important to be aware of the way the programs are released. With active and validated markets, advertising and introducing the software will be done correctly, and the people will take their interested and needed programs with awareness. Since mostly software development is in smart tools such as cell phone and other similar smart phones, thus the markets active in this area are more important. Fortunately, there are active software markets in Iran, which most people use them given their constructive and optimal performance. On the other hand, these markets have succeeded to absorb part of experts in programming, which is a factor for employment and income creation. In addition, since these markets are Iranian, and due to importance to the national history and nature, they have been effective in production of national programs and games in various areas. It is a factor for development of the society culture and monitoring the individuals' tendencies. It means that considering heathy and standard activity of these markets, people can receive their needed software in the optimal quality from them. Additionally, since a large number of people use markets for access to the programs, the markets can serve as a source and effective index for promoting the culture of proper use of software. Of course, governmental institutions should support the knowledge-based companies that are developing so that the creative and talented individuals can perform their activities and these national resources and assets can be effective in the community progress.

3.4 Promoting Software Power; a Step for Technology Localization

Since localization and nationalization is occurring in different sectors of the society in the current era, and most developed communities pave ground for their community's growth and flourishing through reliance on national and local capacities, and utilize human and non-human resources in this regard, hence paying attention to localization of technology and knowledge is essential for progress and development. Given nature of localization, which means aware strategic use of resources and assets for community development at national and international level, and since localization for technology progress influences different aspects of the community, therefore promoting software power using national experts and giving facilities to them is an action in line with technology localization. With paying attention to software technology development, both the community talents and capacity is directed toward the right path, and the community in terms of security cultural, and economic aspects is improved. This is certainly a constructive action in line with purposeful localization of society.

4 Discussion and Conclusion

Today modern world is such that all its aspects and elements should be taken into account for the success, and considering relationship and influence of sciences on each other, various knowledge should be utilized for achievement of standard progress. In this regard, software development in smart tools was investigated from different aspects. On the other hand, it should be accepted that using knowledge in different areas would result in more optimal impact of the produced technology and science. Given variety in software and global activities in this area, it is necessary that activity in software engineering to be based on novel knowledge, so that it is possible to act based on international criteria and indexes in the global competitive space. In order to succeed in this regard, both individual and social attempts are needed, and national measures should be taken for improvement of conditions for activity in this area, so that sustainable development of community based on successful software engineering is implemented. Its constructive outcomes leads to progress in the community and the individuals take benefit from modern and healthy life due to such social progress.

References

1. Saberi A (2016) Knowledge management. In: Second international management and economics conference in the 21st century. Allameh Tabatabaei University, Tehran
2. Saberi A (2017) The path of software development in intelligent systems. In: Second national conference on modern research in electrical and computer engineering. Basir Institution of Higher Education
3. Saberi A (2016) Human resource management and organizational behavior. In: Second international conference on new ideas in management, economics and accounting. Singapore
4. Saberi A (2017) Nationalization of software. Second international conference on knowledge based research in computer engineering and information technology. Allameh Tabatabaei University

5. Saberi A (2017) Design, production and development of electronic content. In: 2nd International conference on knowledge based research in computer engineering and information technology. Allameh Tabatabaei University
6. Saberi A (2017) Management and planning in informatics business. In: International conference on modern research in management, economics, empowerment of the tourism industry in development. Shandiz Institution of Higher Education in Mashhad
7. Saberi A (2017) Technology management in the age of globalization. In: Third national conference on a new look at the transformation and innovation in education, Shiraz
8. Saberi A (2017) Management in smart technology. In: Third annual management conference and business economics, Tehran
9. Saberi A (2017) Iranian software development, supported by software markets. In: National conference on new accounting and management researches in the 3rd millennium
10. Saberi A (2016) Human resource management. In: National conference on new accounting and management researches in the 3rd millennium
11. Seryasat OR, Haddadnia J, Ghayoumi-Zadeh H (2015) A new method to classify breast cancer tumors and their fractionation. Ciência e Natura 37:51–57

Investigating IPv6 Addressing Model with Security Approach and Compare It with IPv4 Model

Asieh Dehvan, Amir Reza Estakhrian and Ahmad Changai

Abstract Internet Protocol Version 6, or IPv6, is the latest version of the Internet Protocol which Internet connections formed based on it. This version is expected to replace version 4 of this protocol that is currently in use, therefore, in this study, the IPv6 addressing model has been utilized with a security method and compared with the IPv4 model, which is the reason of IPv6 features should be held as the benefits of IPv4, the disadvantages of it have been abandoned or diminished or added new features.

Keywords IPv6 · Security · Threats · Vulnerabilities

1 Introduction

Since its origin, in most countries of the world the Internet is the source of increasingly development in human life and its use rate enhanced. It is obvious, designing in some cases, faced serious challenges after decades, and there is no expectation of this. For example, the IP protocol which is one of the main protocols on the Internet, it is not designed to support a large number of devices and users connected to the Internet, but IP addresses or Internet Protocol Address, or the so-called public IP, identifier is assigned to any device connected to the Internet or to a network using the Internet Protocol. This identifier is completely separate and the protocol understands which data or requests to which computer is sent or from which one they will receive. The use of the IP address is similar to the email

A. Dehvan (✉)
Department of Information Technology Engineering,
Barkhat Institute of Higher Education, Ahvaz, Iran
e-mail: asiehdehvan@yahoo.com

A. R. Estakhrian
Department of Engineering, Islamic Azad University, Sepidan, Iran

A. Changai
Islamic Azad University, Ramhormoz, Iran

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_63

address. If we do not have an individual email address, we cannot send him an email; this is also true about the Internet protocol.

The IPv4 protocol was invented in 1970, and at the time no one thought that it would take time to use the above protocol to become a necessity for many tasks. But even though the IPv4 protocol has a great performance but has its own limitations, for example, IPv4 has not been secured. For this reason, other protocols such as IP Sec are implemented with security approach. On the other hand, the most important challenge is IPv4, its address space is limited, and so after many years of the public Internet, the lack of IP addresses has become one of the main concerns on the Internet.

NAT was invited to overcome the restrictions of the number of IP addresses. The above-mentioned technology has made it possible for computers on a dedicated network of private addresses to communicate with one another. But they use a public IP address shared by all Internet connections. In the early 1990s, the IETF is responsible for standardizing the Internet, informed that the IPv4 protocol has to address limitation and from that time on, the new version of the protocol was emphasized and finally, in early 1995, the original version of IPv6.0 was prepared [1].

This version versus version 4 has a very wide range. Also, in version 6, IP addresses are 128-bit (including 8 16-bit sections) and each section segment is separated by two characters (:). The IPv6 structure is more complex than IPv4, and an IP address 6 versions is as follows:

2601: F0A0: 9002: E051: 0000: 0000: 0000: C91D

The segments separated by characters (:) include numbers and letters of the hexadecimal standard (letters A, B, C, D, E, F) that can be varied from 0000 to FFFF.

For the convenience of reading this phrase, sections with four digits can be deleted. Note that this compression can only be performed at a single IP address. The simplified IP address above is as follow: 2601F0A0: 9002: E051:: C91D

As stated, these terms do not have any meaning for the computer and should be converted into binary expressions. In this conversion, the values of each section are converted to a 16-bit term.

That is, converted binary is the following code:

1100100100011101

The reason that each section becomes a 16-bit term is that each character in each segments according to the hexadecimal conversion table, it converts to a four-digit binary term. That is, the first section (2601) becomes 0010011000000001. So, putting these 4 digits together, we will have 16 bits in each section.

The range and magnitude of the IP address of version 6

Given the 128-bit version of IP-6, with a calculation (2128), you can count all the IPs that are usable. So the IP version 6 can be 1038 × 3, 4 different modes. That's about 340 Andsilion or exactly 340282366920938463463374607431768211456 different IP. This means there are 79228162514264337593543 times address more than the IP address of version 4! [2].

How to address in IPv6

How to addressing, which is also called multicasting, is the method and technology used to send a packet of data to multiple destinations during a sending process within a network. Using the multicasting feature, the bandwidth consumed within the network is reduced and the sending process time to multiple destinations and the processing pressure inside the network will be heavily optimized. According to RFC3307, the routing method in Internet Protocol Version 6 is divided into three following modes:

Unicast: it is a transition in which data is sent from one source to a single specified destination within the network;

Any cast: it is transition method that directs the envelope to a specific group of nodes that may be in different places. But in that router, select the closest and best node that the envelope can reach and it sends the envelope only, but it may also be sent to nodes with the same destination address;

Multicast: it is a transition method in which envelope is sent from one source to groups in a group and each node receives this envelope only once; [3].

2 Research Background

Certainly, IPv6 is one of the biggest and most widespread changes to the basic Internet infrastructure. The change has been unprecedented over the past 20 years. In fact, versions 1–3 of the IP protocol have never been used, which is the same for the fifth version.

IPv4 was useful until all of the protocols correctly performed their tasks, however, this series had a series of problems; each computer on the network needed its own web address and IPv4 only had 32 bits.

30 years ago, due to technical issues, the total number of networking devices was less than 2^{32} or 429467296—something close to 2 billion—that did not seem like a big problem.

But today, with hundreds of millions of Internet users, while Internet addresses on mobile phones that use small Internet searches and gaming and home appliances such as fridges and television … and which are soon to be connected to the Internet for ease of use. It is well known why these 32-bit addresses will end soon. So the need to use IPv6 was introduced. In this regard, the most significant change in the address space was quadrupled from 32 to 128 bits in the new IP version, IPv6.

In practice, the new address space will not be used well; however, there will never be a problem with address shortages. In short, we will provide 60,000,000 IP addresses for each of our servers. The expansion of IPv6 will require a rethink of network systems and a change in the configuration of hundreds of millions of computers around the world.

The security issue is also one of the most important issues in this release, which needs to be addressed extensively [4].

Security: security is one of the internal features of the IPv6 protocol, which has both attributes of authentication and encryption in the new IP protocol layer. The public key management system adopted by IPv6 designers is the ISAKMP mechanism [5].

ISAKMP messages are exchanged using the UDP protocol and use the port number 500. IP Sec is a VPN standard that prevents data from being accessed or manipulated by using authentication and encryption mechanisms. It uses both network and network access (computer to the network). The IP Sec protocol uses the IKE as the key management. Designed specifically for IP protocols, it does not provide security for other protocols, unlike PPTP. It also includes two modes of encryption called transport and tunnel. This protocol operates in Layer 3, which includes AH and ESP protocols. The IP Sec uses AH for the authentication and originality of the source without the use of encryption, while ESP provides authentication and originality to cryptography [6].

AH protocol

In summary, the AH protocol will actually provide the following security services:

1. Data integrity sent
2. Authentication of data uploaded origin
3. Ignore re-uploaded packages

This protocol uses HMAC for the integrity of the data sent and to do this, it's based on the secret key, which will pack the payload and the unchanging parts of the IP header like an IP address.

HMAC review

HMAC protects the integrity of the data sent. Because only peer-to-peer points know the secret key created by HMAC and checked by the same.

ESP protocol

The ESP protocol provides the following security services:

- Data integrity sent
- Authentication of data uploaded origin
- Ignore re-uploaded packages

IKE protocol

IKE is a protocol that repairs several important security challenges: peer-to-peer authentication and symmetric exchange keys. This protocol makes a security assembly (SA) and places it on the Security Association data base (SAD). The IKE protocol uses the port number UDP/500.

The IKE consists of two stages

Step 1: The first step is the formation of the Key Management Security Assembly or ISAKMP SA.

Step 2: In the second step, ISAKMPSA is used to negotiate and configure IP Sec, SA.

While IPv4 has the ability to use IPsec derived from Internet Protocol security, but the above feature was added as a new functionality to the above protocol to be used in cases such as tunneling, network encryption in order to access VPNs remotely from Virtual Private Networks and to communicate with sites. A large number of organizations use the IPsec protocol in certain cases, but the existence of barriers such as NAT can make it difficult to use.

In IPv6, the IPsec protocol has been introduced as a necessary part of the implementation of an appropriate security infrastructure to provide authentic services such as authentication, integrity, and reliability. IPsec's operational capacity is such that organizations can help improve their security status and develop their security policies [7].

Investigating IPv6 security issues

Perhaps the biggest controversy was over the security issue. Everyone agrees on the principle that security is needed. The fight was about how to get to the security and how to address it. The first place addresses the security of the network layer. Advocates argued that implementing network-level security would provide a standardized service that all applications could take advantage of without any prior planning. Opponents also argued that secure applications generally do not require any more mechanism than end-to-end encryption. Then the source process encrypts its data and decrypts the destination process. Anything less than that can pose a risk to the user, which is due to security bugs in the network layer, and no fault is detected by him. The answer to that argument was that the user can ignore the security of the IP layer and do its job! The opponents' final response was that those who did not trust the proper functioning of the network (in terms of security).

Why should you pay for the slow implementation of IP? One aspect of security was the fact that many countries have imposed strict laws on the export of cryptographic products. Examples include France and Iraq, which even restrict the use of internal cryptography, and the public cannot keep anything from the police. As a result, any IP implementation that uses strong encryption methods will not be licensed by the United States (and most other countries). The implementation of two software's is one for internal use and one for export is a topic that opposes the suppliers of the computer industry [8, 9].

3 Research Method

This study in term of method is a theoretical kind of research that is based on the deductive and creativity method based on textual and non-textual documentation.

4 Conclusion

In the countries throughout worldwide, IPv4 is used as a platform for Internet addresses. IPv4 protocols are essentially composed of 32-bit addresses. This means that this space allows the addressing of about 4.3 billion addresses, as well as the security in the IPv4 protocol, is optional. In the sense that there are a number of communication security protocols such as IP Sec, which are optional for IP Version 4 but security is the most important issue in the today's world, and security should not be optional but must be compulsory since version 4 of the Internet Protocol is no longer capable of answering needs, and the lack of this address and the low security posed problems for public and private companies. A new protocol called IPv6 was defined, which requires the use of mechanisms and methods known as migration. Accordingly, using these mechanisms, the current IPv4 protocol can be migrated to the new IPv6 protocol. Increased address space, auto configuration, higher security, character routing, ease of multiple- communication and strong support for multidisciplinary facilities include the benefits of the 6th version of the addressing system that is referred to the strategic document for the transition to this system.

It is important to note that IPv6 is an IPsec protocol as an essential part of the implementation, in order to provide an appropriate security infrastructure to provide proven services such as authentication, integrity, and dependability. The IPsec operational capability is such that organizations can help improve their security model status and develop their security policies. Therefore, security is one of the internal features of the IPv6 protocol, which has both IP authentication and encryption features in the new IP protocol layer. As a result, IPv6 features have kept the benefits of IPv4 out of the way, eliminating or reducing its defects or adding new features. Generally speaking, IPv6 is not compatible with IPv4, but it is compatible with all the Internet's protocols such as DNS, BGP, OSPF, IGMP, ICMP, UDP, TCP.

References

1. Learning to Understand IP Version 6, Joseph Davies
2. Atkinson R, Security Architecture for the Internet Protocol, RFC2401
3. Atkinson R, IP Authentication Header(AH), RFC2402
4. Atkinson R, IP Encapsulating Security Payload(ESP), RFC2406
5. Douglas E. Comer. Computer Networks and Internets‘ Prentice Hall 2004
6. Huitema C, IPv6 The New Internet Protocol, Prentice Hall, New Jersey 1996
7. http:/en.wikipedia.org/wiki/pv6_IPv6, Wikipedia.org, …
8. http://www.internetworldstats.com/me/ir.htm, Iran Internet …
9. Doyle Jeff, CCIE Professional Development: Routing TCP/IP_Edition, Chapter …

Design of Dual-Band Band-Pass Filters with Compact Resonators and Modern Feeding Structure for Wireless Communication Applications

Mohammadreza Zobeyri and Ahmadreza Eskandari

Abstract In this paper two new planar dual-band band-pass filters (DBBPF), that each of them is designed using an exclusive type of feed structure and by planar technologies of microstrip have been presented. First DBBPF is implemented, using compact open-loop meandering resonators, and open stub-loaded feed-lines. The cause of introducing the transmission zeros near the pass-bands of the first filter is source/load coupling at stub-loaded feed structures. The longer resonators, which are fed by open-stub loaded feed-lines, independently control the lower pass-band and the shorter resonators which are magnetically coupled, independently control the higher pass-band. The second DBBPF are designed by using 0° feed structure and by combining of two types of resonators. The main resonators with diagonal feeding structure and similar to stepped-impedance resonators (SIR) construct the lower pass-band. The sub-resonators with double-spiral and inter-coupling, can be embedded in the structure of main-resonators and construct the higher pass-band. There is a cross-coupling between the four sub-resonators, that suppresses any harmonics and introduces the extra transmission zeros at each side of second pass-band. Both sets of DBBPFs are designed to miniaturize, improve the gain, decrease losses and are fabricated to apply to wireless communication applications. The obtained experimental results validate the simulation results.

Keywords Quasi-elliptic response · Compact resonator · Diagonal symmetry feeding · Open stub-loaded · Dual feeding structure · Dual-band band-pass filter (DBBPF)

M. Zobeyri · A. Eskandari (✉)
Department of Electrical Engineering, East Tehran Branch,
Islamic Azad University, Tehran, Iran
e-mail: ar_eskandary@yahoo.com

M. Zobeyri
e-mail: mohammadrezazobeyri@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_64

1 Introduction

In recent year, dual-band and multi-band band-pass filters with high performance are one of the essential components in the design of transceivers used in the wireless communication applications. Already, to meet the requirements of the wireless systems, different design approaches have been presented [1]. Most of the designed filters are based on Stepped-Impedance Resonators (SIR) [2–6]. The recent results are aimed at obtaining high-efficiency filters that are used at multiple standard frequencies, especially, in Wireless Local Area Networks (WLAN) also in Global Systems for Mobile, (GSM) communication. For instance, a new generation of WLAN standards such as IEEE 802.11 b/g, which are working at 2.4 GHz. As reported in [3] and [7], GSM and their generation wideband code division multiple access (WCDMA) work at both 0.9 and 1.8 GHz. Recently, many efforts have been made to miniaturize the dual-band band-pass filters (DBBPFs), which had led to the design of the most compacted DBBPFs, of course, in many cases, Stepped-Impedance Resonators (SIR) has been used. In many filters basing on new WLAN standards, it's necessary to avoid the lots of losses, large circuit area, and low selectivity. So, to reduce the size and compact the dimensions of filters and resonators are especially important in new design methods of planar filters. Some of the recent designs have higher selectivity because of using 0° feed structures [2–7]. At the first time, it was shown in [2] how can be improved the skirts of pass-bands. By applying a diagonal symmetric at tap-coupled I/O position, due to adding 0° feed structure, two transmission zeros (TZs) will appear near the passband edges. In [3] a diagonal symmetric feed structure has been introduced for hair-pin type main-resonators, but the embedded sub-resonators are fed in a mirror symmetric form, which is called dual-feeding structure. However, the presented frequency response of this filter has not enough between-band isolation. In [4] proposed that the 0° feed structure with diagonal symmetry form are applied for both sets of main- and sub-resonators. In [5] by applying a dual-feeding structure, using the presented methods in [3], two DBBPF for different resonant frequencies are implemented, without using SIR feature. Although, the stop-band of this filter needs more suppression. Also, in [6, 7] a 0° feed structure are applied only for two coupled resonators without using any feed-line. Always the SIRs at fundamental resonant frequency introduce a pass-band. Moreover, by adjusting the impedance ratio of the SIRs, another pass-band can be tuned in the desired frequency, which is a harmonic of first pass-band. The major challenge of this type of resonators in some cases is that the harmonic resonant frequency has high insertion loss and low selectivity. Besides using SIR, in most of the new works, open or short stub-loaded or short-circuited resonators or a composition of mentioned methods have been utilized [6, 7]. Extra open/short stubs can be efficiently used to control the center frequency of pass-bands. Although, some of above filters have an efficient performance, it's impossible in most of them, to create another pass-band by adding extra resonators in their structures to promote the filters to a higher order response.

In this paper, two sets of DBBPFs are presented, which in by combining the different types of coupling methods and by embedding the sub-resonators in the structure of main-resonators two final DBBPFs, have been designed and miniaturized. Besides, in these methods with the use of different kinds of feed structures, such as 0° feed structure and stub-loaded feed structure, the DBBPFs have been simulated and optimized. In the first set of DBBPFs, for creating two passbands, two pairs of compact meandered open-loop resonators are utilized and open-stub are loaded in order to improve the coupling structure of resonators. Due to the presented coupling scheme of the filters and cause of parasitic coupling between the source/load, extra TZs near the pass-bands are created that improved the stop-band performance. In the second set of DBBPFs instead of using the traditional feed structure, a 0° feed structure are applied for the two open-loop SIRs. This state of coupled main-resonators, results in a single-band frequency response at 0.9 GHz, with two TZs at both sides of pass-band. By embedding the dual-spiral resonators (DSRs) in the structure of the main-resonators, an upper pass-band at 1.8 GHz has been realized. The proposed DSRs are symmetric and have a compact size and uniform-impedance [8]. Because of the cross-coupling between the embedded sub-resonators, two TZs have been introduced at both sides of upper pass-band. Sharp skirts of the double pass-bands at the GSM standard frequencies, make a good between-band isolation and high-frequency selectivity due to the introduced TZs and near-zero insertion losses, all of these good features of the filters are provided.

At the end of the paper, two sets of filters have been compared in terms of structural and functional parameters. Consequently, two prototypes of the designed filters are presented after fabricating and measuring. The simulated and measured results of fabricated prototypes are in a good agreement which validates the design methodology.

2 Design of the First Set of DBBPFs

2.1 The Coupling Scheme and the Primary Structure

Figure 1 shows the primary structure of the first set of DBBPFs, that consists of feed-lines and open-loop resonators. According to the simple structure in Fig. 1, by changing the length of the two resonators at above and below the feed-lines, the dual pass-band can be tuned easily. Besides, the primary structure of the proposed filter in Fig. 1, has compatibility of adding more resonators and making a cascaded structure [9].

Figure 2 shows the proposed coupling topology for the first set of DBBPFs, which illustrates the arrangement of the coupling between the resonators and feed-lines. So, using a coupling scheme to realize the first set of DBBPFs is necessary. The black nodes 1 and 2 represent the longer open-loop resonators, whereas the black nodes 3 and 4 represent the shorter open-loop resonators. The hollow

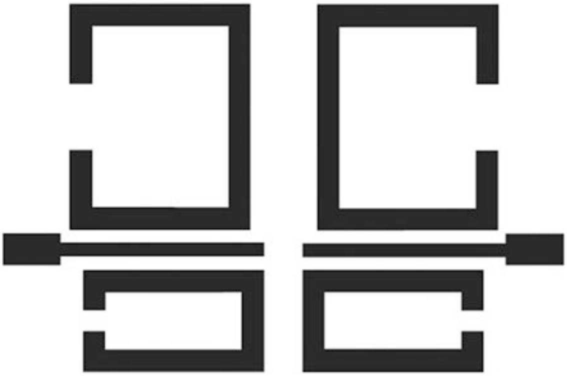


Fig. 1 Primary structure of the first set of DBBPFs

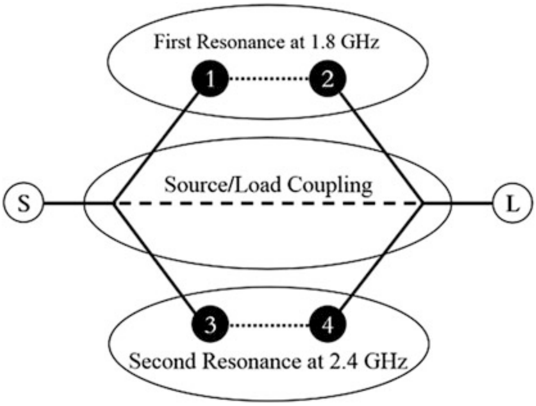


Fig. 2 Coupling scheme of the first set of DBBPFs

nodes in Fig. 2, represents the source and load. According to the signal paths between In/Output ports in Fig. 2, it is found that the coupling between the two resonators 1 and 2 at upper signal path causes to resonating filter, and construct a pass-band at 1.8 GHz as a first independent band. The coupling between the two resonators 3 and 4 causes the second resonant frequency at 2.4 GHz, as second pass-band of the filter. The middle path causes source/load coupling and creates TZs in the stop-band and improves the performance of the DBBPF.

The proposed primary structure in Fig. 1, is capable of cascading by adding more serial resonators. The features of resonating in the above proposed filter can be achieved by employing the common methods of designing the microstrip SBBPFs and transmission line theory. As investigated in [9], there are two types of traditional coupling structures for microstrip filters with coupled resonators as:

1. Tapped-line feeding structure
2. Coupled-line feeding structure

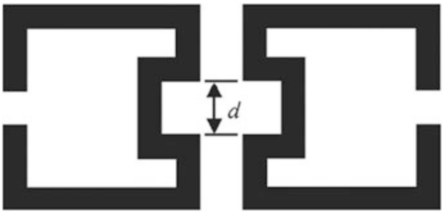


Fig. 3 The improved structure of the shorter resonators

At the first set of DBBPFs of this paper the coupled-line feeding structure has been employed. So that the two proposed sets of compact resonators as shown in Fig. 1, are placed on the two sides of the coupled feed-lines. But at the second set of DBBPFs, the tapped-line feeding structure has been applied. Due to the gap between the In/Output feed-lines in Fig. 1, the coupling coefficients for both passbands are limited. To improve the coupling coefficient between the two resonators 1 and 2, the open side-coupled stubs and internal stubs are used. Due to the short length of the two resonators 3 and 4, these resonators need to smaller coupling spacing to improve their coupling coefficient. So the geometry in Fig. 3 is proposed to control the coupling coefficient in the second pass-band by the indentation length between the two shorter resonators. In this case, the two pairs of resonators on upper and lower sides of the coupled feed-lines can have more parallelism, which is studied in [9]. Due to the wide occupied area of the proposed structure in Fig. 1, also to strengthen the magnetic coupling, beside the modified indentations (are shown in Fig. 3) the structure of the resonators must be more meandered to save on occupied area [10].

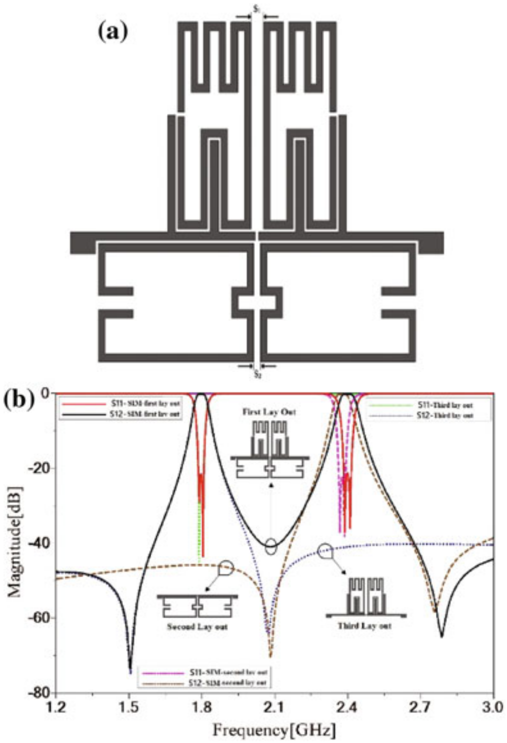
2.2 The Configuration of the DBBPF After Meandering the Resonators

In this paper the applied resonators in the structure of the first DBBPFs with λ/2 electrical length, have been compacted and meandered, so, they follow the theory of microstrip transmission lines. Thus, the frequency response of each passband is related to the total length of the corresponded resonators, and the desired pass-band frequency can be obtained by tuning the total length. After meandering and miniaturizing all the resonators, and modifying the shorter resonators exclusively, based on the primary structure and the coupling scheme (as shown in Figs. 1, 2 and 3) the results are obtained as Fig. 4.

As shown in Fig. 4, the changes in the frequency response if any pair of coupled resonators removes, are depicted and compared. What is inferred from the frequency responses in Fig. 4b is that each pair of the resonators and their corresponded pass-band are independent of another pair.

In Fig. 4a, the distance between the two resonators 1 and 2 is determined by "S_1", and the distance between the two resonators 3 and 4 at lower path is

Fig. 4 **a** The configuration of the DBBPF including the shorter and longer meandered resonators, **b** the frequency response of the filter with and without each of the resonators



determined by “S_2”. The coupling spacing between each pair of the resonators directly affects the magnetic coupling. As shown in Fig. 5b, by reducing the S_1, the bandwidth and subsequently the ripple of the first pass-band will increase abundantly. Also, in Fig. 5b is shown that any change in the S_2, will affect the bandwidth and ripple of the second pass-band. Of course, increasing over the distance between the two resonators, makes the bandwidth strongly reduced that is not desired. So, over increase, the coupling spacing beyond a value causes cause the destruction or sloppy integration of the transmission poles in pass-bands. So, the coupling spacing must be optimized to obtain the desired frequency with maximum bandwidth and minimum ripple.

The feed-lines of the filter has the ability to displacement and inscribing the shorter resonators as shown in Fig. 6a. inserting shorter resonators inside the feed lines, on the one hand, helps miniaturizing the shorter resonators and on the other hand, saving the occupied area of the filter that consequently decreases the bandwidth of the second pass-band. According to the Fig. 6b, by increasing the length of the coupled feed-lines and aggregating both In/Output port on one side of the filter, the TZ at the upper edge of the second pass-band will be shifted to a higher frequency beyond 3 GHz. According to the frequency response of above DBBPF, can be inferred that the fundamental coupling occurs between the shorter and

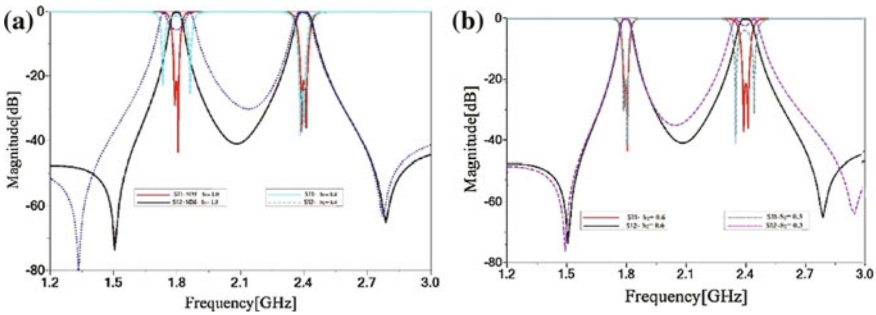


Fig. 5 **a** The frequency responses of the proposed filter in the Fig. 4a, with two different values of "S_1", **b** the frequency responses of the proposed filter in the Fig. 4a, with two different values of "S_2"

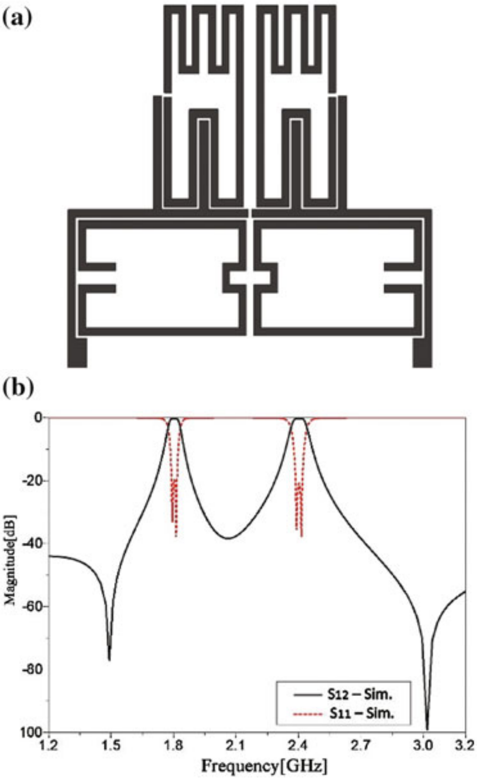


Fig. 6 **a** The proposed filter with a change in the coupled feed-lines, **b** the frequency responses of the filter after increasing the length of coupled feed-lines

longer resonators but the coupled feed lines have not a direct effect on the resonance phenomenon. As shown in Figs. 4, 5, 6, 7 and 8, in comparison of frequency responses, this can be determined that the increasing the feed lines in the process of miniaturizing the filter, moves the higher TZ, which effects on the bandwidth of second pass-band. Thus, the bandwidth and coupling coefficient depends on many factors, not only to coupling spacing between the coupled resonators (S_1 and S_2).

The coupling coefficient can be calculated as [11]

$$\frac{F_{h2} - F_{l2}}{F_{h2} + F_{l2}} = K_i, K_{i+1} \tag{1}$$

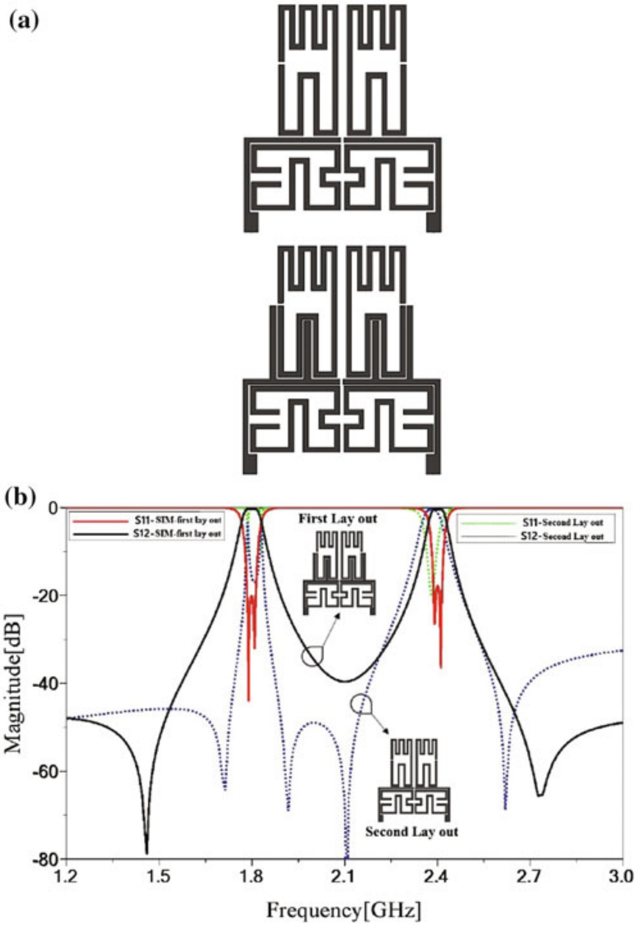


Fig. 7 **a** The proposed structure in Fig. 6a with and without the open stubs, **b** the comparison of frequency responses of the above filter with and without the open stubs

Were the K_i, K_{i+1} is the coupling coefficient at each pass-band, and the F_{h2} is the higher and the F_{l2} is the lower resonant frequency.

In Fig. 7, the internal open stubs at the middle of the longer resonators, and side-coupled stubs are removed respectively to compare the results in the different states of loading or removing two pairs of stubs. Consequently, as shown in Fig. 7b, the filter response at the first pass-band without the open stubs is poor, and functionally the performance of the filter has been disturbed. So, the open stub-loaded in the proposed configuration, improved the coupling of the filter, effectively. One of the advantages, presented for the first set of DBBPFs is that the different bandwidths can be obtained by changing the length of coupled feed-lines, but achievable bandwidths are limited.

Figure 8 shows the maximum rate of compactness can be applied to the first set of DBBPFs without disturbing the resonant modes in second pass-band. By over compacting the two inscribed resonators between the feed lines (resonator 3 and 4), the bandwidth of the second pass-band will be decreased. Often decreasing the

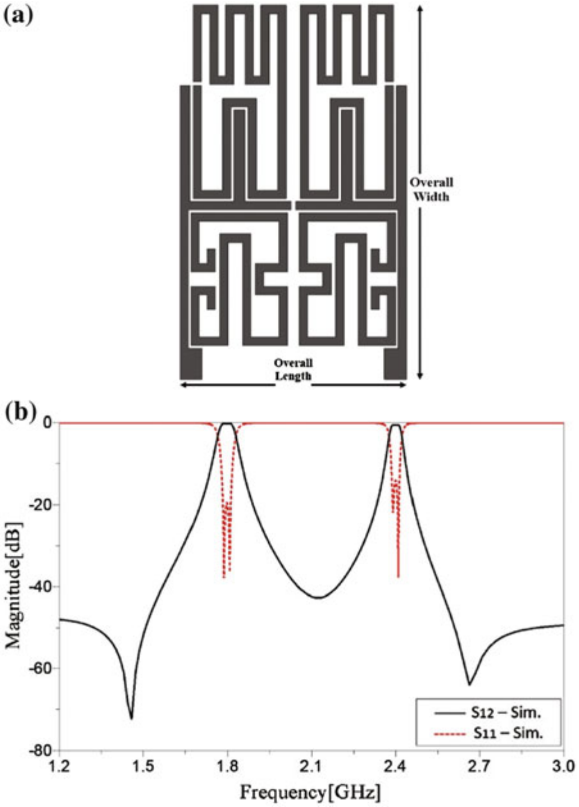


Fig. 8 **a** The final structure from the first set of DBBPFs, **b** the frequency responses of the final filter at 1.8 and 2.4 GHz

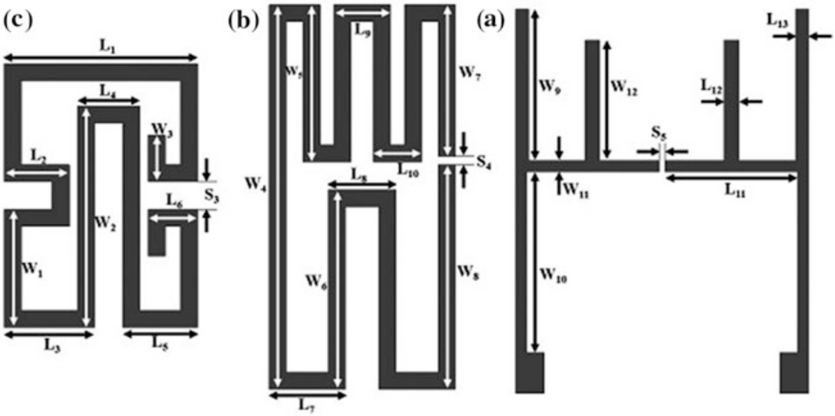


Fig. 9 **a** The dimensions of the feed lines, **b** the dimensions of the longer resonators, **c** the dimensions of the shorter resonators

bandwidth of the pass-bands is not desired, and sometimes is considered as a disorder, also according to the low distance between the transmission poles (especially in the second passband), it might the transmission poles be tangled or suppressed if the bandwidth further decreased.

2.3 *Numerical Parameters and Information*

Based on the final design of the first DBBPF (Figs. 8 and 9) the overall size of the most compacted filter is 24.3 * 16 mm^2. The coupling spacing between the resonators and feed-lines is 0.2 mm, and the size of the internal open stubs and side-coupled stubs are 6.5 * 0.8 mm^2 and 8.2 * 0.7 mm^2 respectively. The first set of DBBPFs are simulated on the substrate of Duriod 5880 with a dielectric constant of 2.2 and thickness of 0.508 mm.

The rest of the dimensions are determined as follow:

L_1 = 6.85 mm, L_2 = 2.3 mm, L_3 = 3.2 mm, L_4 = 2.2 mm, L_5 = 2.65 mm, L_6 = 1.75 mm, L_7 = 2.7 mm, L_8 = 2.4 mm, L_9 = 2.0 mm, L_{10} = 1.7 mm, L_{11} = 7.25 mm, L_{12} = 0.8 mm, L_{13} = 0.7 mm, W_1 = 4.2 mm, W_2 = 7.9 mm, W_3 = 1.65 mm, W_4 = 13.7 mm, W_5 = 5.6 mm, W_6 = 7.1 mm, W_7 = 5.4 mm, W_8 = 8 mm, W_9 = 8.2 mm, W_{10} = 9.8 mm, W_{11} = 0.6 mm, W_{12} = 6.5 mm, S_1 = 1.0 mm, S_2 = 0.8 mm, S_3 = 1.0 mm, S_4 = 0.3 mm, S_5 = 0.3 mm.

The insertion loss in the first and second pass-bands are 0.3 and 0.4 dB and return loss are 18 and 20 dB respectively. The between-band isolation and out-of-band rejections are 40 and 48 dB respectively. The maximum fractional bandwidth obtained in 3 dB, at two 1.8 and 2.4 GHz pass-band, according to the Fig. 4b are equal to 3.33% (1.77–1.83 GHz) and 3.75% (2.36–2.45 GHz) and the minimum fractional bandwidth that obtained in 3 dB, according to the Fig. 4b is equal to 3.33% (1.77–1.83 GHz) and 2.1% (2.38–2.43 GHz) respectively.

3 Design of the Second Set of DBBPFs

3.1 The Basic Theory of 0° and Non-0° Feed Structure

In the second section of this paper a new feed structure, are introduced for the Fundamental structure of second proposed DBBPFs. As investigated in [11], there are different types of coupling to construct microstrip filters which include the electrical, magnetic and mixed coupling. To create the electrical coupling two electrical-coupled open-loop resonators with a common feed structure in Fig. 10a, are presented. As can be seen in Fig. 10a, the electric delays between the upper and lower paths, between the two In/Output ports, are different for the open-loop resonators. The symmetric structure shown in Fig. 10a is considered as non-0° feed structure.

There is another type of new feed structure is shown in Fig. 10b that has a zero-degree difference between the electric delays of the upper and lower paths.

For the first time in [2] after examining a transmission matrix at upper and lower signal paths for the two U-shaped resonators, a 0° feed structure was proven and analyzed. The results that achieved was, two TZs at both sides of pass-band edges can be obtained, which improves the selectivity of the pass-band. In [11] it is proven that the TZs would be introduced when the delays at upper and lower electric paths (with electrical length θ_1 and θ_2) approach to $\pi/2$.

3.2 Basic Structure of Proposed DBBPF

Figure 11a shows a Fundamental structure for realizing the second DBBPF, that consists of two electrical-coupled main-resonators. Each of the main-resonators consist of two low-impedance sections connected by a folded high-impedance

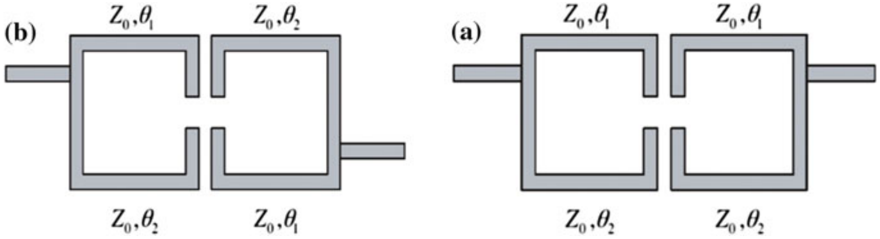


Fig. 10 **a** The open-loop resonator with the non-0° feed structure, **b** the open-loop resonator with the 0° feed structure

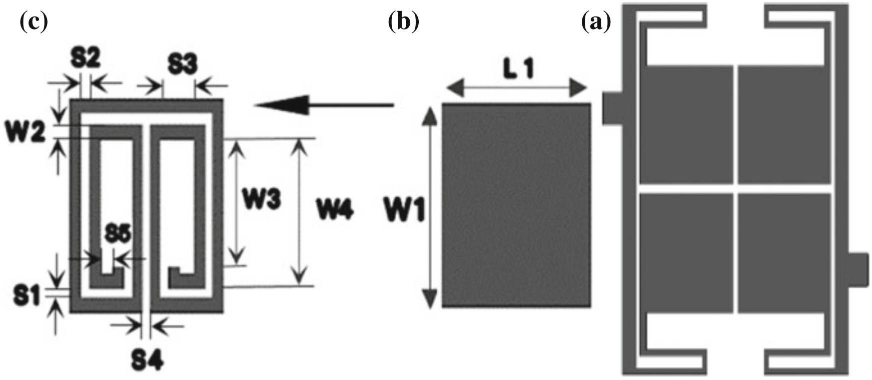


Fig. 11 **a** The proposed fundamental structure of the second DBBPF with 0° feed structure, **b** the low-impedance section of the main-resonators, **c** the embeddable sub-resonator instead of a low-impedance section of the main-resonators

section. In Fig. 11a, at the middle of each main-resonators a tap-coupled transmission lines is connected, as In/Output ports.

To create a pass-band at 1.8 GHz, it is necessary to replace the low-impedance sections of main-resonators (are shown in Fig. 11b) with open-loop dual-spiral resonators (DSRs) which are introduced as sub-resonators (as shown in Fig. 11c). Consequently, the resulted configuration is a DBBPF, consists of four cross-coupled sub-resonators with a high performance.

Figure 12a shows the fundamental structure with a 0° feed structure that is capable to excite a pass-band at 0.9 GHz as a single-band band-pass filter (SBBPF). The proposed DSRs with λ/2 electrical length and uniform-impedance as sub-resonators (depicted in Fig. 10c) can be embedded in the structure of the main-resonators. The resulted configuration excites another pass-band at 1.8 GHz next to the first pass-band at 0.9 GHz. There is an inter-coupling inside the sub-resonators which allows the 1.8 GHz pass-band to be adjusted conveniently. In Fig. 12a by moving the position of two tap-coupled ports in contradictory directions, a diagonal symmetry has been achieved which has the 0° electric delays between the two upper and lower paths. The frequency response of this process with a different length ratio of W_6 to W_5 are compared in Fig. 12b. The simulation results of the fundamental structure with diagonal symmetry has a 0.9 GHz pass-bands in frequency response with two realized TZs, a high selectivity and, low insertion loss. The distributions of electric and magnetic fields are presented in [3].

3.3 The Proposed DBBPF 0.9/1.8 GHz

In an analysis of fundamental structure as shown in Fig. 12a where the length ratio of W_6 to W_5 approaches to 0.5 a 0° feed structure will be created which is proved in

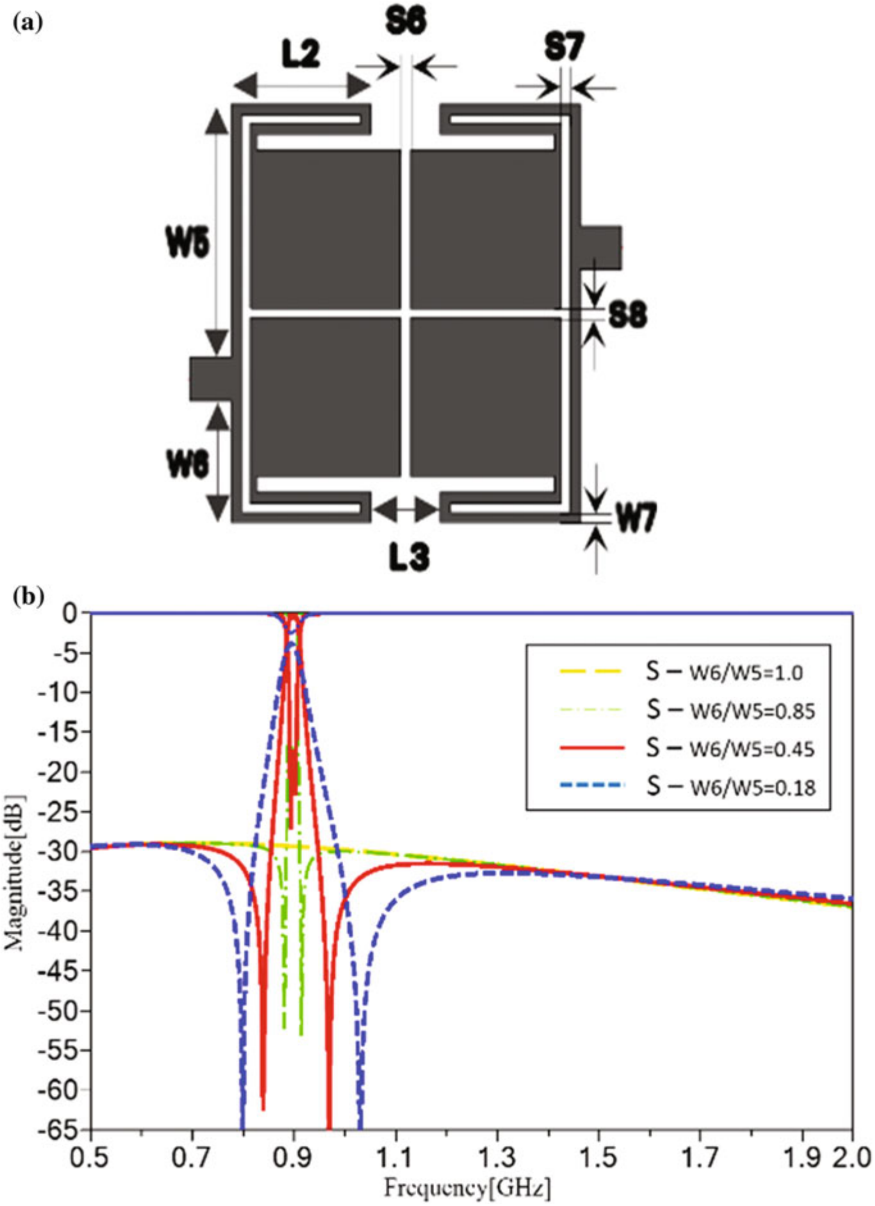


Fig. 12 **a** The dimensions of the fundamental structure with 0° feed structure, **b** the frequency response of the fundamental structure with 0° feed structure

[2]. By applying this basic theory for fundamental structure and by embedding the half-wavelength sub-structures a dual-band response with two independent pass-band can be obtained.

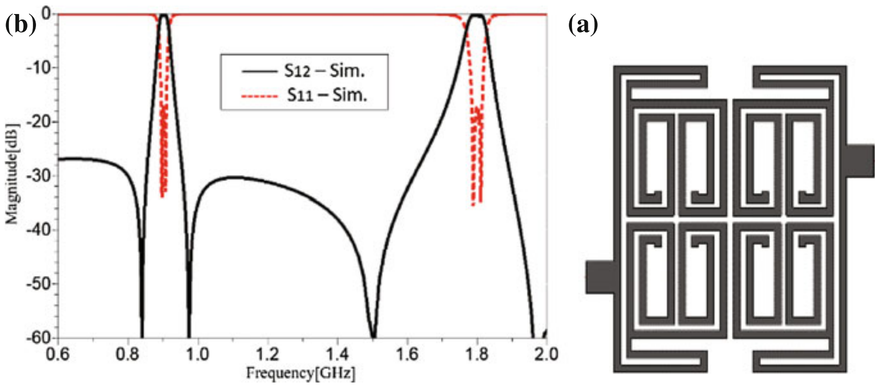


Fig. 13 **a** The final structure second DBBPF, **b** the frequency response of the second DBBPF

Figure 13a shows the combination of two features, consist of coupled main-resonators with 0° feed structure and embedded sub-resonators to construct the corresponding pass-band. Using the compact sub-resonators and folding the main-resonators are very effective in miniaturizing the overall structure. The simulated results shown in Fig. 13b is an optimization of the three dimensions (W_3, W_5 and W_6). The proposed sub-resonators were used before in [12].

3.4 *The Proposed DBBPF at 0.9/1.8 GHz*

After tuning the dimensions of the proposed structures based on the above method, at last, two prototypes of SBBPF and DBBPF are fabricated on the substrate of Duriod 5880 with a dielectric constant of 2.2 and thickness of 0.787 mm. the photograph of the fabricated SBBPF is shown in Fig. 14, and the simulated and measured results are compared in Fig. 15 that are in a good agreement.

Fig. 14 The photograph of the fabricated SBBPF

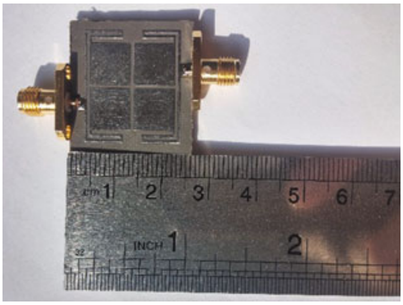


Fig. 15 The comparison of simulated and measured results of fabricated SBBPF

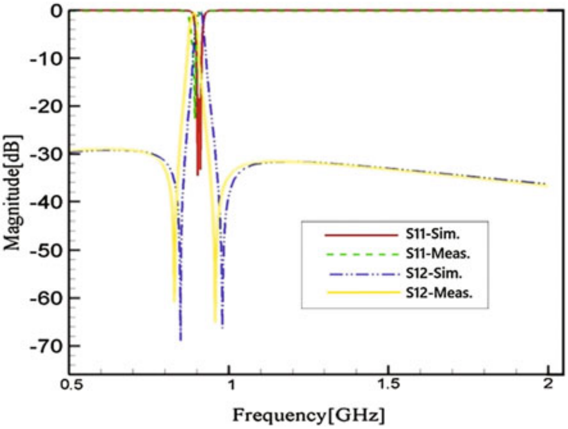
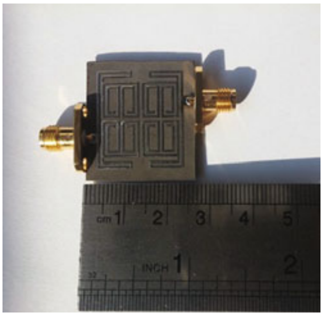


Fig. 16 The photograph of the fabricated DBBPF



By examining the frequency responses in Fig. 15, the experimental results compared to the simulation results have a slight shift down, but in general, it can be said that the results have a good agreement (Fig. 15). The photograph of the fabricated DBBPF is shown in Fig. 16. The simulated and measured results of the DBBPF are illustrated and compared in Fig. 17. There is a good accordance between the depicted results and the measurement is similar to the prediction as Fig. 17 shows.

In this paper, all the presented structures, are simulated by EM-Simulator (ADS). The "S" parameters of the fabricated prototypes are measured by Agilent network analyzer N5230A. The fractional bandwidth in 3 dB at two 0.9 and 1.8 GHz pass-band according to the Fig. 17, is equal to 3.65% (884–916 MHz) and 3.9% (1765–1835 MHz). The overall size of the both SBBPF and DBBPF is the same and is equal to 23.3 * 20.4 mm^2. The rest of the dimensions are as follow:

W1 = 8.75 mm, W2 = 0.6 mm, W3 = 5.3 mm, W4 = 6.75 mm, W5 = 14.85 mm, W6 = 6.05 mm, W7 = 0.6 mm, L1 = 8.8 mm, L2 = 8.2 mm, L3 = 4 mm, S1 = 0.35 mm, S2 = 0.5 mm, S3 = 1.9 mm, S4 = 0.4 mm, S5 = 0.8 mm, S6 = 0.6 mm, S7 = 0.5 mm, S8 = 0.6 mm. The insertion loss of the SBBPF and DBBPF are below the 0.9 dB and 0.7 dB respectively. The return loss of both SBBPF and DBBPF are above 18 dB. Two designed filters have a high freedom degree in design, high selectivity and are capable to suppress any harmonics.

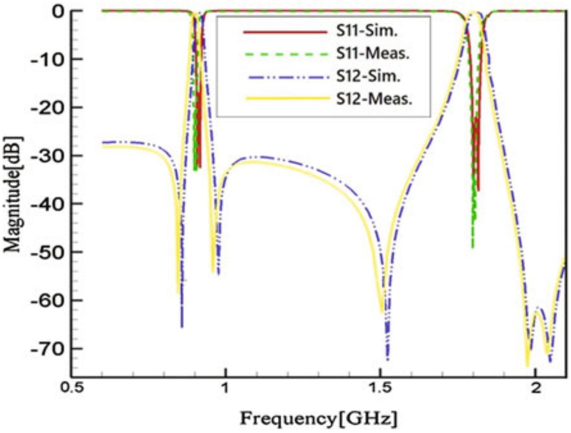


Fig. 17 The comparison of simulated and measured results of fabricated DBBPF

4 Comparison of the First and Second DBBPFs

Although the first pass-band in the first DBBPF is located at 1.8 GHz but in the second DBBPF located at a frequency of 0.9 GHz, this makes it impossible for the first and second DBBPFs to be compared in different ways, such as miniaturizing and compacting the dimensions. This is because their first passband is located at different frequencies, so, the two DBBPFs have different guided wavelengths (λ_g). However, in the case of the coupling structure, the existence of the gap between the two input and output feed lines in the first DBBPF makes it possible to create a multi-path propagation mode that results in the production of TZs in the edges of the transmission bands. The TZs in the second DBBPF are primarily due to the existence of a 0° feeding structure and, secondly, because of the cross-coupling between the four sub-resonators, which are quad-coupled, and create two TZs near the second pass-band. The presence of two TZs between the two pass-bands (between the 0.9 and 1.8 GHz pass-band) in the second DBBPF causes a good isolation, which improves the selectivity of the filter, but this feature is weaker in the first filter due to the absence of TZs. The response of first DBBPF is Chebyshev, but the response of the second DBBPF is quasi-elliptic type [9]. These two DBBPFs have a high degree of freedom in a variety of ways, including the fact that both filters have the ability to change from dual-band to single-band mode. In the first DBBPF, by removing one of the pairs of resonators at upper side or the lower side of the feed-lines, the remained pair of resonators can be used to resonate a single-band response, but in the second filter structure, by electrically-coupling the fundamental resonators, one single-band response can be made at a frequency of 0.9 GHz, but with the alone use of compact sub-resonators, without the initial structures, the 1.8 GHz pass-band can't be realized.

5 Conclusion

In this paper, by using two different design methods, two sets of DBBPFs are presented. After the fabricating and measuring the second method, the simulation results and the fabrication of its filters were compared. The purpose of the design of two DBBPFs is to provide compact filters with high performance and low loss for standard GSM and WLAN frequencies. In the first set of dual-band filters, after the presentation of a coupling scheme, the miniaturization process is performed. In the first method of DBBPF design, several layouts with different bandwidths are recommended, which can be used as needed. Finally, the layout that has the most compactness, the lowest bandwidth, and the least possible ripple (in comparison to others), has been investigated in terms of dimensional parameters. In the following, the second DBBPF is presented, which is designed, constructed and measured by developing a fundamental structure with a stepped-impedance shape. By utilizing a 0° feed structure that is applied in the electrically-coupled main-resonators, the low pass-band GSM-band is obtained at 0.9 GHz. Also, by embedding the four dual-spiral sub-resonators in the fundamental structure, a second pass-band in the high-frequency GSM-band is obtained at 1.8 GHz, which can be adjusted by changing the electrical length of the sub-resonators. Two pairs of TZs near the second DBBPF are introduced, due to the cross-coupling of four compact sub-resonators (DSRs). The use of the 0° feed structure, created a wide stop-band (between 970 and 1550 MHz) between the dual band. Both methods lead to Low insertion loss and high return loss in all generated transmission-bands. In the above filters, the ideal TZs near the pass-bands with a high attenuation rate were realized, due to the multi-path propagation modes, which resulted in a sharp skirt pass-band and decreasing the stop-band level. All the desirable features, including the high rate of suppression in stop-band and the high rejection for all proposed filters and the isolation between the two bands for the second DBBPF are realized. There is a very good match between the simulation and measurement results. The results make the designed DBBPFs, use suitable for GSM and WLAN applications.

References

1. Hong J-S (2011) Microstrip Filters for RF/microwave Applications. 2th edn, Wiley, New York. ISBN 978-0-470-40877-3 (Hardback)
2. Tsai C-M, Lee S-Y, Tsai C-C (2002) Performance of a planar filter using a 0° feed structure. IEEE Trans Microw Theory Tech 50(10):2362–2367
3. Chen C-Y, Hsu C-Y, Chuang H-R (2006) Design of miniature planar dual-band filter using dual feeding structures and embedded resonators. IEEE Microw Wirel Compon Lett 16 (12):669–671
4. Xue W, Liang C-H, Dai X-W, Fan J-W (2007) Design of miniature planar dual-band filter with 0° feed structures. Prog Electromagnet Res 77:493–499

5. Yao Z, Wang C, Kim NY (2013) A compact dual-mode dual-band bandpass filter using stepped-impedance open-loop resonators and center-loaded resonators. Microw Opt Technol Lett 55(12):3000–3005
6. Fan JW, Liang CH, Wu B (2008) Dual-band filter using equal-length split-ring resonators and zero-degree feed structure. Microw Opt Technol Lett 50(4):1098–1101
7. Yang R-Y, Hon K, Hung C-Y, Ye C-S (2010) Design of dual-band bandpass filters using a dual feeding structure and embedded uniform impedance resonators. Prog Electromagnet Res 105:93–102
8. Song K, Zhang F, Zhunge C, Fan Y (2013) Compact dual-band bandpass filter using spiral resonators and short-circuited stub-loaded resonators. Microw Opt Technol Lett 55(6):1393–1398
9. Chuang M-L (2005) Cascaded dual band coupled-fed microstrip open-loop filter. Microw Opt Technol Lett 45(6):519–522
10. Weng M-H, Huang C-Y, Wu H-W, Shu K, Su Y-K (2007) Compact dual-band bandpass filter with enhanced feed coupling structures. Microw Opt Technol Lett 49(1):171–173
11. Hong JS, Lancaster MJ (2001) Microstrip filters for RF/microwave applications. Wiley, New York. ISBN 0-471-38877-7 (Hardback)
12. Wu G-C, Wang G, Liang JG, Gao X-J, Zhu L (2015) Miniaturised microstrip dual-band bandpass filter using novel symmetric double-spiral resonators for WLAN application. Electron Lett 51(15):1177–1178

New Fuzzy Logic-Based Methods for the Data Reduction

Reyhaneh Tati

Abstract The problem of finding effective data reduction algorithm was discussed in the paper. To improve the quality of initial data, the matrix and singular value decomposition and fuzzy reasoning were used, the method of selection of the most efficient reduction algorithm was proposed.

Keywords Data reduction · Singular value decomposition · Fuzzy logic

1 Introduction

Composition of a data mining model is a dynamic and repeatable process. Improving the quality of the data in data preprocessing is very important. The overhead data because of huge dimension leads to some useless data mining algorithms. One way to improve the quality of data is the data reduction. There are various noises, incomplete data, unused fields with redundancy, lost data or outlier data, etc. in the databases. Any unrelated and redundant data in the database should be removed by data processing. Data preparation is important to get the best results with maximum efficiency in data mining algorithms and to extract valuable knowledge from the data.

There are many data reduction methods. In most of these methods only attributes are reduced, but there are number of instances in the data set which should also be reduced after processing. On the other hand in many ways of data reduction the effectiveness of data in the data set is not considered. Thus there is the danger that some relevant data could be omitted and it will lead to incorrect results. In the method, proposed in the paper, the singular value decomposition and data sets are considered in the form of matrix. The matrix columns are attributes, and rows of the matrix are the instances of data set. The use of singular value decomposition reduced rank-*k* and the product matrix that is the result of multiplying of orthogonal

R. Tati (✉)
Department of Computer, Doroud Branch, Islamic Azad University, Doroud, Iran
e-mail: Tati.computer@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_65

matrix and diagonal matrix of reduced rank-*k* run the fuzzy reasoning. Multi-factor evaluation is one of the applications of fuzzy sets theory in decision-making process. With fuzzy reasoning, determining the weight of each row and column for selecting one among the equal rows and one among the equal columns provides the reduce of the instances of data sets and also and reduces attributes of data sets. The advantage of this method over single value decomposition method lies in the possibility to leave the relevant data but at the same time repeated attributes will be removed. Another advantage of this method—the priorities between instances and attributes are based on achieved weights. It created a way to find the best method the data reduction among another available methods using fuzzy reasoning. There are many data reduction methods using the parameters like: computation time, accuracy, the number of reduced attributes that can be computed. With fuzzy reasoning derived rules and parameters are considered in data reduction methods as factors, achieved to the final criterion for evaluating methods of data reduction based on all computable parameters. In these methods it is possible to achieve the score for each method of data reduction in any time according to existing data that should be processed.

What is very important today is not only shortage or absence of information, but also lack and/or absence of methods of accurate identification of good methods of knowledge extraction. The main topic of knowledge discovery in database is how to use the theories for extracting knowledge and information from large amount of data.

Some applications of data mining for increasing improvement of the quality of data are: data integration, data cleaning, data normalization, data reduction. Preprocessing on the data before data mining can improve the quality of the data.

A database or data warehouse may store terabytes of data. Complex data analysis may take a very long time to run on the complete data set. So, the data reduction problem arises. The data reduction obtains a reduced representation of the data set that is much smaller in volume but still produces the same (or almost the same) analytical contents. The most popular data reduction strategies include dimensionality reduction, e.g., remove unimportant attributes, wavelet transforms, principal components analysis (PCA), feature subset selection, feature creation, regression, histograms and clustering [1]. There is another way for data reduction called as singular value decomposition (SVD) since it is the foundation for new algorithms that is proposed in the presented paper.

2 The Fuzzy Singular Value Decomposition Method (FSVD)

The proposed fuzzy singular value decomposition method based on classical attributes and instances dimensionality reduction by singular value decomposition (SVD) and by Takagi-Sugeno (TS) fuzzy model.

At first stage SVD is used on initial dataset. At second stage TS fuzzy model is used to analyze SVD of data set.

2.1 The Takagi-Sugeno Fuzzy Model

The proposed fuzzy singular value decomposition method (FSVD) based on the Takagi-Sugeno fuzzy model and the model must be first considered.

Rule-based models of the Takagi-Sugeno type [2] are suitable for the approximation of dynamic systems. The Takagi-Sugeno rule l within the fuzzy inference system is written as follows:

$$\begin{matrix} \mathit{If}\ \hat{x}^1 \text{ is } B_l^1 \text{ and} \ldots \text{and } \hat{x}^{m_1}\ \mathit{is}\ B_l^{m_1} \\ \text{then } y_l = \varsigma_{l0} + \varsigma_{l1}\hat{x}^1 + \cdots + \varsigma_{lm_1}\hat{x}^{m_1} \end{matrix}, l = 1, \ldots, r, \tag{1}$$

where $\hat{x} = [\hat{x}^1, \hat{x}^2, \ldots, \hat{x}^{m_1}]^T$ is the input vector, y_l is the output vector of the l th rule, $B_l^{t_1}$, $t_1 \in \{1, \ldots, m_1\}$ are fuzzy sets defined in the antecedents space by membership functions $\gamma_{B_l^{t_1}}(\hat{x}^{t_1}) : \mathrm{R} \to [0, 1]$, ς_{lt_1}, $t_1 \in \{1, \ldots, m_1\}$ are consequent parameters, and r is the number of rules.

2.2 Singular Value Decomposition Method

The singular value decomposition (SVD) method firstly is the method for transforming correlated variables into a set of uncorrelated ones that better expose the various relationships among the original data items. Secondly SVD is a method for identifying and ordering the dimensions and data points exhibitions of most variation. Also SVD is the tools for analysis of discovered data and data reduction and finding the dependence among variables and mapping them in the area of small dimensions [3].

2.3 The FSVD Method for Data Reduction

Obtained from singular value decomposition reduced rank-k and the product matrix that is the result of multiplying the orthogonal matrix and diagonal matrix of reduced rank-k, will run the fuzzy reasoning. In the proposed fuzzy singular value decomposition method, after calculating the weight of each row and column the rows and columns with the same weights could be considered as one row and column.

So, there is a four-step procedure, where the matrix $\hat{X}_{n\times m_1} = [\hat{x}_i^{t_1}]$, $i = 1,\ldots,n$, $t_1 = 1,\ldots,m_1$ of the initial data is the input.

Step 1. A single value decomposition of the matrix $\hat{X}_{n\times m_1} = [\hat{x}_i^{t_1}]$ should be calculated according to formula

$$SVD(\hat{X}_{n\times m_1}) = U_{m_1\times m_1}\Sigma_{m_1\times n}V_{n\times m_1}, \tag{2}$$

The value of k, $k = i$, $1 \le i \le n$ must be chosen and get smaller dimensions of the analyzed matrix as follows:

$$\hat{X}k_{n\times m_1} = Uk_{n\times k}\Sigma k_{k\times k}Vk_{k\times m_1}, \tag{3}$$

Step 2. To get the weight of each row of the matrix $\hat{X}k_{n\times m_1}$, which represents the weight of each instance with its attributes consideration, the Takagi-Sugeno fuzzy model should be used. Then the fuzzy rules r will be transformed into fuzzy relationship. The number r of fuzzy rules in the rule base can be calculated according to formula

$$r = n(Rule) = n(\hat{X}_{n\times m_1})^{n(\hat{x}_i^t)}, \tag{4}$$

Step 3. Return a fuzzy rule to the fuzzy relation according to formula

$$w^l = \prod_{t_1=1}^{m_1} \gamma_{B_l^{t_1}}(\hat{x}^{t_1}), \tag{5}$$

where w^l is the compromise ability of the l th rule. The following method to defuzzification of performance [4] should be applied: using linear functions for the input–output for fuzzy sets that is used as the result of fuzzy rules and doing Steps 3 and 4 for rows and columns of matrix $\hat{X}k_{n\times m_1}$.

Step 4. Viewing the weights of rows to choose the rows that have equal weight. Thus there is reducing the instances. Also viewing the weights of columns of the matrix to choose the columns that have equal weight. By this method there is a reduction of attributes of data collection.

By fuzzy reasoning the weights for rows and columns are obtained. Then one row of the set of rows with the same weight should be selected and thereby to provide the data reduction of instances. The same process also should be applies to columns for reduction of attributes of data sets.

The advantage of this method over single value decomposition methods used for data reduction is that it does not remove data that is related to data collection. Only the attributes which are repeated are removed.

2.4 Properties of the FSVD Method

All attributes in existing instances are weighted by rules obtained in fuzzy reasoning. These attributes can be taken into account in all instances, because they are all relevant and dependent on existing attributes in instances. In SVD method a small difference between data is not considered. The values obtained from the SVD method are considered as the parameters for the Takagi-Sugeno fuzzy model and all fuzzy rules are produced according to membership functions for each value of instance in attribute. The same process will be for all these instances.

After evaluation of all instances and their attribute values the output values have to be combined together and after their defuzzification, weight of instances will be obtained. The process for all attributes should be repeated. In comparison with others this method uses the singular value decomposition so that each of attributes is considered according to the values of all instances in this attribute and their relationship and achieved criterion for each studied attribute. For all instances, that have the same weight, only one instance is used. This method will cause instances reduction. The instance with a greatest weight is worthy for the data set, because this amount of weight is obtained according to the investigation of instances value and considering all possible states that attributes can have, so the value is valid. The same process should be done for all attributes of data set.

The advantage of this method is a combination of instances reduction and also attributes reduction. Most data reduction methods consider only attributes reduction. Since in different situations such as data mining there are a lot of instances, the method chooses an instance among instances that has greater role in data set. With this method ability of choosing instances and attributes according to their values and role in data collection will be possible.

Another property of this method is that there will be priority for selecting attributes and instances according to their weights, and ranging them from large to small. Fuzzy logic does select data reduction for these studied attributes and available instances, receiving the validate value for making decision about selection and reduction of attributes. Another sensible advantage of this method is that the rules based on the input will be identified and modeled to keep instances and attributes that have an effective role in data collection. The proposed method has generated smaller dimension of data compared with other methods of data reduction.

3 An Illustrative Example

The Anderson's Iris data set is characterized in brief in the first subsection. The second subsection includes the results obtained from the proposed FSVD method of the data reduction.

3.1 The Anderson’s Iris Data Set

The performance of the proposed fuzzy singular value decomposition method should be explained by an illustrative example. For the aim the Anderson’s Iris data were selected.

The Anderson’s Iris data [5] is the most known database to be found in the pattern recognition literature. The data set represents different categories of Iris plants having attribute values. The attribute values represent the sepal length, sepal width, petal length and petal width measured for 150 irises. It has three classes Setosa, Versicolor and Virginica, with 50 samples per class. It is known that two classes Versicolor and Virginica have some amount of overlap while the class Setosa is linearly separable from the other two.

3.2 The Obtained Results

The MATLAB® software was used to show proposed method implementation. To evaluate the first proposed method, let’s first define input parameters in the software MATLAB®. Using of Iris data set for test proposed algorithm, attribute information was described as follows: 1. sepal length in cm, 2. sepal width in cm, 3. petal length in cm, 4. petal width in cm [5]. The matrix columns represent attributes and the rows indicate the data set instances. For reducing attributes defined in software MATLAB®, 4 parameters were considered which show instances numbers. For each instance: [−4 8]:

```
>> instance4 = [−4:.1:8]’;
>> instance2 = [−4:.1:8]’;
>> instance3 = [−4:.1:8]’;
>> instance1 = [−4:.1:8]’;
>> attribute = [−4:.1:8]’;
>> input = [instance1, instance2, instance3, instance4];
>> weight = [input attribute];
>> anfisedit;
```

After writing above subject in MATLAB and produce rules calculate k-reduced singular value decomposition for $k = 1$ then calculate uk, sk, vk.

So, $bk = uk * sk * vk$, after calculation of bk the amount of attributes for all instances in obtained rules and obtained weight for each attributes attention new method should entered. Also avoidance of message “out of memory” for large amount of data in MATLAB® considers each 4 instances together at final and it is necessary to repeat the process. In calculation of weight for each 150 instances, it is possible to consider one instance instead of several instances with the same weights, so it will give the reduction of instances. So the new method can recognize the priority of instances. Also with new method the amount of instances will be reduced to 65. The use of new method for attributes reduction receives the weights of each attribute and the priority of attributes that will be recognized. Then it will be used for arranging each sepal length, sepal width, petal length and petal width. That is why all attributes must be presented in the resulting matrix. So, there are reduction of dimensions of the data set, from $150 \times 4 = 600$ to $65 \times 4 = 260$.

On the other hand, two attributes such as sepal width and petal length was selected by using the clustering technique for attribute selection [6] and reduction of dimensions of the data set was achieved from $150 \times 4 = 600$ to $150 \times 2 = 300$. So, the results obtained from the FSVD algorithm seem to be appropriate in comparison with time clustering method which was proposed in [6].

4 A Technique for Selection of an Appropriate Method for the Data Reduction

A plan of the proposed technique is described in the first subsection. The second subsection deals with the results obtained from the proposed technique.

4.1 An Outline for the Technique

The proposed technique determines criteria for selecting the best method data reduction between available data reduction methods. The technique can be described as a seven-step procedure.

Step 1. Read the parameters X in the data reduction, where X are input variables of fuzzy set $x = (x_1, \ldots, x_n)$, which is variable of reasoning premise.

Step 2. A fuzzy set $A = (A_1, \ldots, A_n)$ with its membership function is determined for each X on the Step 1.

Step 3. Fuzzy rules should be produced according to A and X. The number of fuzzy rules can be determined according to a formula

$$r = n(Rule) = n(A)^{n(X)}. \tag{6}$$

The direct method of the fuzzy reasoning when rule number increases that makes rules rather boring.

Step 4. Return fuzzy rule to fuzzy relation as follows [7]:

$$w^i = \prod_{k=1}^{n} \mu_{A_k^i}(x_k), \tag{7}$$

where $\mu_{A_K^I}(x_k)$ is the value of membership of fuzzy set A_k^i for corresponding x_k. So, the number of fuzzy relations with a number of fuzzy rules is equal and r is the total number of rules, and the w^i is the compromise ability of the i th rule.

Step 5. Fuzzy output values are calculated as follows:

$$y^i = c_0^i + c_1^i x_1 + \cdots + c_n^i x_n,\ i = 1, 2, \ldots, r \tag{8}$$

where x is the input variables and c_k^i is the result parameter for the ith rule.

Step 6. The following method for defuzzification of the performance is used:

$$y = \sum_{i=1}^{r} w^i y^i / \sum_{i=1}^{r} w^i, \tag{9}$$

The linear functions are used for the input–output and it is replaced by fuzzy sets that are used in result of the fuzzy rules.

Step 7. The performance of each method is calculated as follows

$$\text{Performance} = 1 - \text{y}, \tag{10}$$

The rating of each method is obtained. So, performance for each method will be obtained, because each output value represents the rate of appropriate feature of each method of data reduction. Choosing the most appropriate method for data reduction among different data reduction methods according to the data set can be made.

4.2 The Obtained Results

There are several methods for data reduction, so there are two parameters which should be considered as the best criterions for the calculating the rating the best methods of data reduction. These two parameters are included: (1) time consuming of data reduction for each method and (2) amount of reduced attribute for each method. Using second proposed method will show the method of data reduction with best criteria.

Let's see some important for data reduction features:

T: time-consuming of method of data reduction as an input variable:

$$T \in [0\ \ 100], \tag{11}$$

A: number of reduced attributes as an input variable (accuracy):

$$A \in [0 \quad 1], \tag{12}$$

P: performance of method of data reduction as an output variable:

$$P \in [0 \quad 1]. \tag{13}$$

Considered fuzzy sets for input variables, time and accuracy A are defined as follows:

Fuzzy sets for input variable T: {excellent, good, bad}

Fuzzy sets for input variable number of attributes selected by each method: {low, unsuitable, very}.

According to fuzzy sets and the number of input variables according to Step 3, there are 9 fuzzy rules are created. According to Step 4 conversion of fuzzy rules to fuzzy relations, and after stage 5 Non-fuzzy value Y is achieved.

The required time is achieved by subtracting the end time from the start time. The number of reduced attributes is achieved by subtracting the number of obtained attributes from the number of attributes of the total data. Value obtained for the performance is score of each method in which the higher it is the more suitable is the data reduction method. Let us consider a subset of the image segmentation data that makes by University of Massachusetts which is presented in WEKA®Software [8].

After loading and doing data reduction by WEKA®Software [8], obtained values of time and the number of reduced attributes and their normalized values and their performance of 7 selected methods according to proposed method is shown in Table 1, and enter into MATLAB®Software such Fig. 1 and achieve performance of the method by new method.

A comparison of data reduction methods regarding to the parameters number of attribute selected and time consuming for data reduction is presented in Fig. 2. So, according to performance obtained by M3 method and according to available data,

Table 1 Values obtained by each data reduction methods which are implemented in the WEKA®Software and performance values

Number of method	Models of attribute evaluation	Method for searching	Number of selected attributes	Normalized selected attributes	Time used for attribute reduction	Performance
M1	Exhaustive search	CfsSubsetEval	3	0.60	8	0.9753
M2	Random search	CfsSubsetEval	2	0.55	9	0.9815
M3	Best first	ConsistencySubsetEval	1	0.60	8	0.9916
M4	Ranker	ReliefFAttributeEval	3	0.05	19	0.9700
M5	Ranker	PrincipalComponents	1	0.50	10	0.9876
M6	Ranker	SVMAttributeEval	25	0.05	19	0.7500
M7	Ranker	ChiSquaredAttributeEval	1	0.05	19	0.9900

Fig. 1 Performance of the fuzzy inference system for proposed method

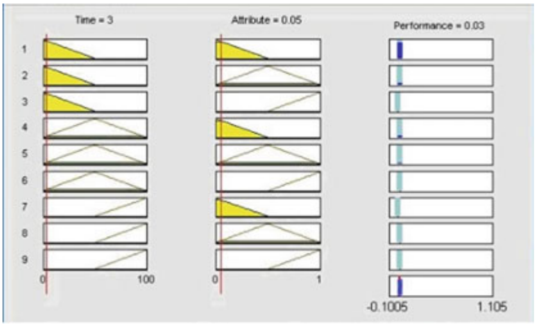


Fig. 2 Chart of methods of data reduction with obtained scores by proposed method

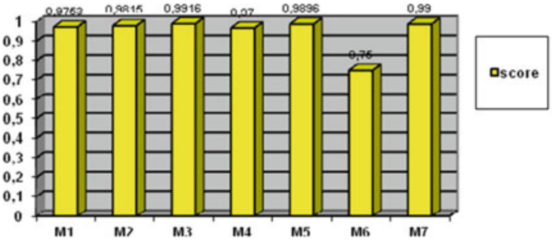
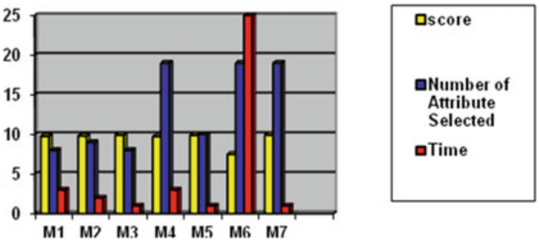


Fig. 3 Comparison of data reduction methods and each method scores obtained by proposed method



M3 is the best to apply for data reduction. In method according to number of selected attributes, method M3, M5, and M7 will have the same value and it is uncertain which one should be chosen, because both will reduce the same number of attributes neglecting the time consuming of each method.

Also, regarding to required time consuming for each method of data reduction the methods M1 and M3 will be in the same conditions and it will be doubt to choose one of them due to the same time consuming and there will be differences between number of selected reduced attributes, that could be neglected. M3 is the highest score in the second proposed method.

Resulting characteristics of different data reduction methods obtained from the proposed technique are presented in Fig. 3.

So, fuzzy logic and fuzzy reasoning allow to make a conclusion that among 7 studied methods for reducing desired data, the third method should be used.

By using fuzzy logic and fuzzy reasoning the criteria that has been achieved for evaluating the data reduction method and selecting the way with higher performance.

5 Concluding Remarks

The fuzzy singular value decomposition (FSVD) method for the data reduction is proposed in the paper. The method is based on singular value decomposition and fuzzy reasoning. The Takagi-Sugeno fuzzy model is used for fuzzy reasoning. Using the fuzzy reasoning, the value of weight is obtained for each instance and attributes. After comparison of obtained weights one weight among the same weights is chosen.

In the proposed method not only attributes reduction is considered like in other methods but also the reduction of the number of instances is considered. So, there are reduction of dimensions of data set. In addition, this method avoids removing related properties of data set that it is one of the important factors in data set. The result of computational experiment for Anderson's Iris data confirms the highest efficiency of the proposed method.

Also the paper shows the way to find the best method among different methods of data reduction considering the criteria by fuzzy logic. It will help to choose which method is suitable for data reduction according to existing data.

References

1. Han J, Kamber M, Pei J (2011) Data mining: concepts and techniques, 3rd edn, Morgan Kaufmann Publishers, Inc.
2. Sugeno M, Takagi T (1985) Fuzzy identification of systems and its application to modeling and control. IEEE Trans Syst Man Cybern 15(1):116–132
3. Baker K (2005) Singular value decomposition tutorial, 29 Mar, 24p
4. Belohlavek R, Sklenar V, Zacpal J (2004) Concept lattices constrained by attribute dependencies. In: Snásel V, Pokorný J, Richta K (eds) Proceedings of the Dateso'2004 (2004)—Annual International Workshop on DAtabases, TExts, Specifications and Objects (Desna, Czech Republic, 14–16 Apr 2004), pp 63–73
5. Anderson E (1935) The irises of the Gaspe Peninsula. Bull Am Iris Soc 59(1):2–5
6. Viattchenin DA (2009) An algorithm for detecting the principal allotment among fuzzy clusters and its application as a technique of reduction of analyzed features space dimensionality. J Inf Organ Sci 33(1):205–217
7. Sugeno M, Kang GT (1986) Fuzzy modeling and control of multi-layer incinerator. Fuzzy Sets Syst 18(3):329–345
8. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten I (2009) The WEKA data mining software: an update. SIGKDD Explor 11(1):10–18

A New Approach for Processing the Variable Density Log Signal Using Frequency-Time Analysis

Esmat Mousavi, Yousef Seifi Kavian and Gholamreza Akbarizadeh

Abstract Cementing is a common practice in drilling and completion of each well, and it is necessary to evaluate how to bonding the cement in this operation. So far, all the methods used to evaluate the extent and how to bonding the cement is eye-catching. Due to the fact that most of the signals are non-stationary, time-frequency analysis methods are among the best methods, the most important of which is the short-time Fourier transform and continuous wavelet transform. Due to the problem of the dispersion of the signal components in these two transformations, the conversion of synchrosqueezing by fixing this limit, it has achieved better results. In this paper, using these transformations, various modes of VDL sonic signaling, which is a criterion for evaluating the quality of cement bond, are investigated. The data used in this article is related to the South Yarran Field to a depth of 2294.5 m located in Ahvaz, Iran.

Keywords Cement bonding · VDL signal · Short time Fourier transform
Continuous wavelet transform · Synchrosqueezing transform

1 Introduction

Cementing is a common practice in the drilling and completion of each well, and it is necessary to evaluate the extent and how to bonding the cement in this operation. The main objectives of the cementing of the casing and the liner pipes in the oil and gas wells are the separation of the production layers, the prevention of fluid leakage

E. Mousavi · Y. S. Kavian · G. Akbarizadeh (✉)
Department of Electrical Engineering,
Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran
e-mail: g.akbari@scu.ac.ir

E. Mousavi
e-mail: e.mousavi@mscstu.scu.ac.ir

Y. S. Kavian
e-mail: y.s.kavian@scu.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_66

to low pressure layer, prevent the formation of unwanted fluid, support the external surface of the pipes and prevent their corrosion.

There are different methods for assessing the quality of cement bond, but what's more common is the CBL and VDL cement bonding tools. These tools are an array of sound-imaging tools that are easy to handle, but their data interpretation requires a lot of precision and skill.

Cement bonding loggings have been developing since about 50 years ago, and these tools are divided into two types of sonic and ultrasonic, and a wide range of techniques have been proposed for their evaluation. In [1], the measurement relations are presented for measuring cement bonding, which is still used to interpret the graphs. In [2], it has been shown that there are large variations in the measurement of domain values due to the variable gateway settings, and it is suggested that another form of waveform is needed to overcome this ambiguity.

In [3], it used two sets of sonic and ultrasonic tools to evaluate cement bonding. By combining CBL and ultrasonic data, small changes in cement are detected. In [4] and [5], by examining the wave, methods have been identified for a better examination of cement bonding. In [6], a new transmitter was used in [7, 8] indicate that the transmitter used on this device has a lot of advantages over devices that use only one transmitter. Nowadays more advanced tools for Checking the cement bond is used. But what we are looking for in this article is to use the CBL-VDL tool to examine the cement bond.

The signal under study in this article is like a lot of signals in our everyday life, is non-stationary signal, and this means that if the spectrum of these signals is drawn, then we see that this spectrum changes a long the time. The strongest and best way to analyze these signals is to use time-frequency display methods.

In [9] short time Fourier transform is proposed. In [10], it has been shown that continuous wavelet transformation has improved the proposed algorithms and spectral interpretation techniques. In [11], the instant spectrum, along with the experimental mode of decomposition, offered for seismic data. The conversion of synchrosqueezing was initially used to analyze the sonic signal, then it became a powerful tool for analyzing signals whose behavior vary over time [12–14].

In Part 2, the type of data is examined, in Sect. 3, the conversions used are explained, and in Sect. 4, the results are compared with the comparison.

2 VDL Data

The CBL/VDL tool is used to measure the cement bond in the drilled wells, whose structure is shown schematically in Fig. 1. This tool generally has a full-wave sonic transmitter and dual receiver. In the most common mode, a receiver is located at a distance of 3 ft (CBL) and another at a transmitter's 5 ft (VDL) distance. This tool should be thoroughly arranged in the center of the well and can not be placed in presence of the gas or gas bubbles into the well.

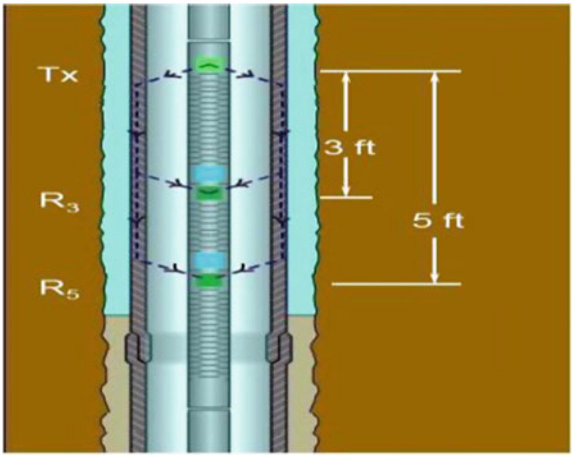


Fig. 1 The structure of the instrument CBL/VDL [16]

The CBL instrument transmitters typically work in the range of 15–30 kHz and 15–60 pulses per second, depending on the type of instrument and the company or server are different [12, 13]. The signal coming out from the sender can have several paths to reach the receiver, paths include: wells, walls, cements, and formations. The signal seen in VDL is a combination of these signals [15]. A view of the signal paths and their waveforms are shown in Fig. 2.

The cement placed behind the casing rise to a variety of situation, in this paper, we consider four modes which are show in Fig. 3.

Examples of cases include: (1) free pipe (2) full cement (3) when the bond in the vicinity of casing is not well, but it is good in the vicinity of the formation. (4) The last state is when the cement state in the vicinity of casing is good but it is bad in the vicinity of the formation. Figure 4 show a set of VDL waves.

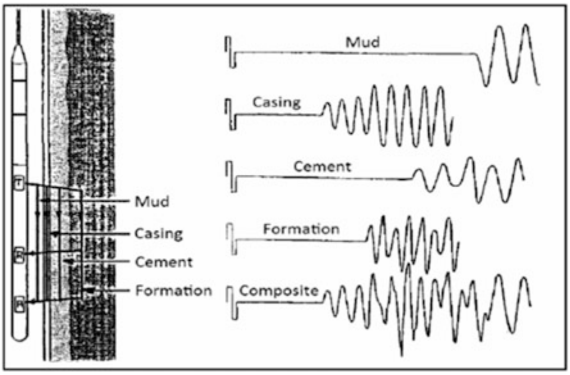


Fig. 2 Paths motion of Sonic signal [17]

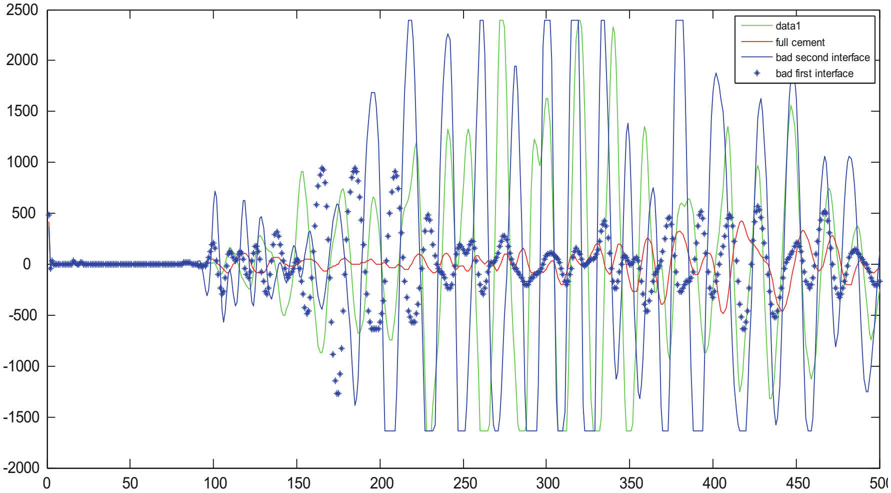


Fig. 3 Different modes of getting cement

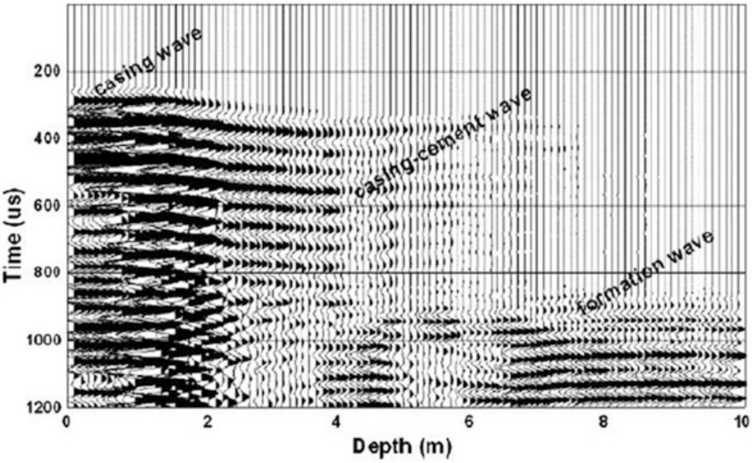


Fig. 4 A train of VDL waves [18]

3 Analysis of Time-Frequency Transformation Methods

One of the best methods for analyzing non-stationary signals, the signal model of which is not available, is to use analytical methods in the time-frequency domain. The main purpose of this analysis is to determine the energy density along the frequency axis at a constant time. Here we mention the most important ones.

3.1 Short Time Fourier Transformation (STFT)

If we consider a non-invariable signal in limited-length windows, it can be static and, by taking fourier formation to short time fourier transformation, that corresponds to Eq. 1. [20]

$$\mathrm{STFT}(t,\tau)=\int_{-\infty}^{\infty}\mathrm{s}(\tau).\mathrm{w}(\tau-\mathrm{t}).\mathrm{e}^{-\mathrm{j}2\pi\mathrm{f}\tau}\mathrm{d}\tau \tag{1}$$

$s(\tau)$ is signal and W(t) is Window function.

3.2 Continuous Wavelet Transform (CWT)

In transforming wavelet, signal converts to wavelets which are transferred form or changed scaled by a mother wavelet. The Eq. 2 is the wavelet transform relation, which x(t) is signal and $\varphi(t)$ is mother wavelet.

$$CWT(\tau,s)=\frac{1}{\sqrt{s}}\int_{-\infty}^{\infty}x(t).\varphi^{*}\left(\frac{t-\tau}{s}\right) \tag{2}$$

3.3 Synchrosqueezing Transformation (SST)

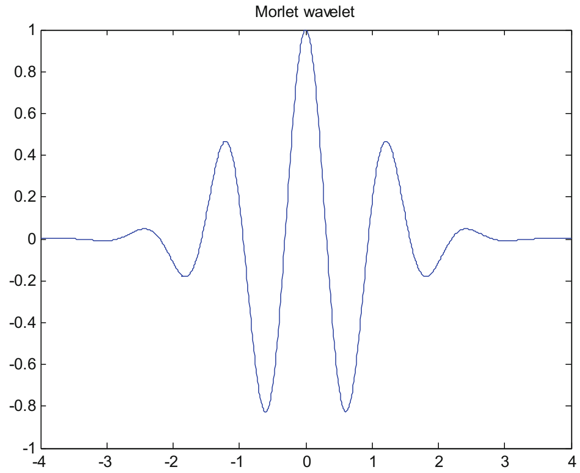
The synchrosqueezing transform draws by focusing on the energy spectrum along with the frequency and time display a frequency-time diagram [20], which is performed in three steps. First, we calculate the signal wavelet transform using Eq. 3.

$$\mathrm{W}(\mathrm{a},\mathrm{b})=\frac{1}{\sqrt{\mathrm{a}}}\int_{-\infty}^{\infty}\mathrm{x}(\mathrm{t}).\varphi^{*}\left(\frac{\mathrm{t}-\mathrm{b}}{\mathrm{a}}\right) \tag{3}$$

$\varphi(t)$ is mother's wavelet, which is used in conjunction with the proximity of the seismic signal to the Morellet wavelet as a mother wavelet, which is shown in Fig. 5

In the second step, given that the dispersion in the wavelet diagram is high, in order to solve this problem use the demodulation of the time-lapsed graph which is obtained in the previous step in the direction of the frequency with respect to Eq. 4 [19].

Fig. 5 Morellet wavelet



$$w(a,b) = -\frac{i\partial \mathrm{bW(a,b)}}{\mathrm{W(a,b)}} \tag{4}$$

And in the final step, using Eq. 5, synchrosqueezing transforms, we get the signal [19].

$$T(w,b) = \int W(a,b)a^{-3/2}\,da \tag{5}$$

Thus, each of the oscillatory components of the signal is centered on the time-frequency plane.

4 Results

In this section, using the STFT, CWT, and SST transformations, we consider various VDL signal modes, which include situations where cement is adjacent to the good and bad pipe, and the situations where the cement is in the vicinity of the good well and adjacent to the bad pipe and the results are compared together.

5 Conclusion

As you can see from Figs. 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, SST has a better resolution and resolution than two wavelet transforms and short-time Fourier transform. In areas where there is little or no cement bond, there is a small amount

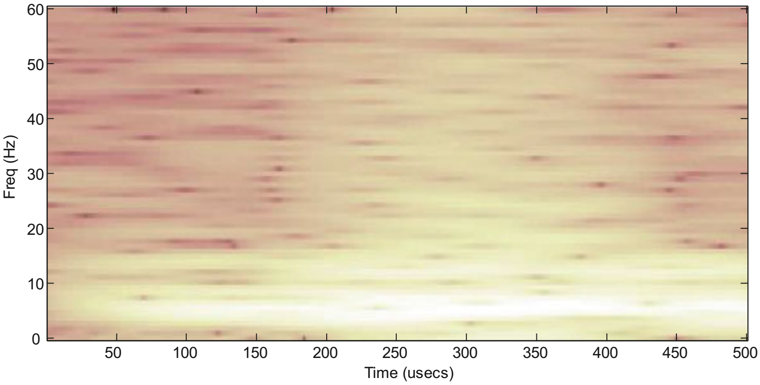


Fig. 6 The short-time Fourier transform of free pipe

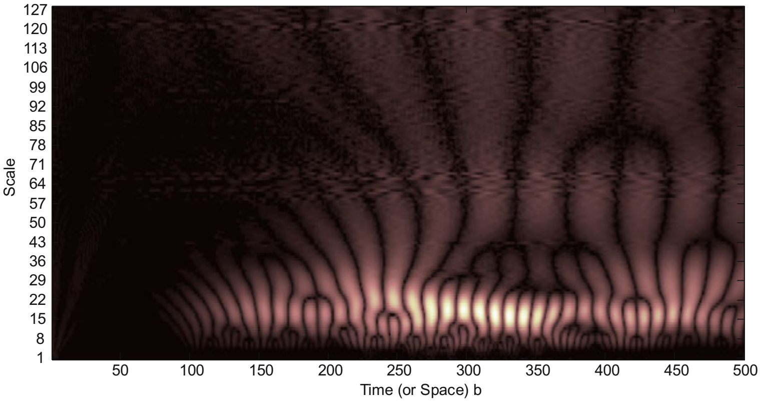


Fig. 7 Continuous wavelet transform of free pipe

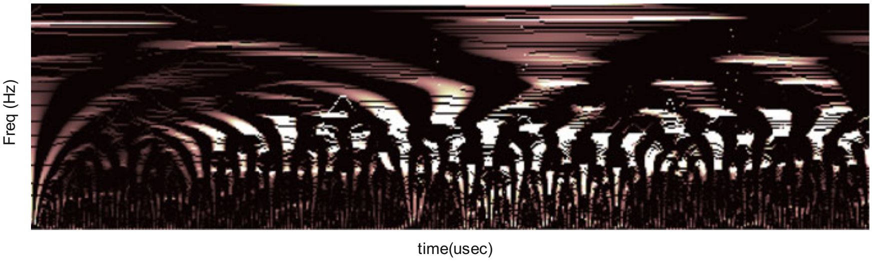


Fig. 8 Synchrosqueezing transform of free pipe

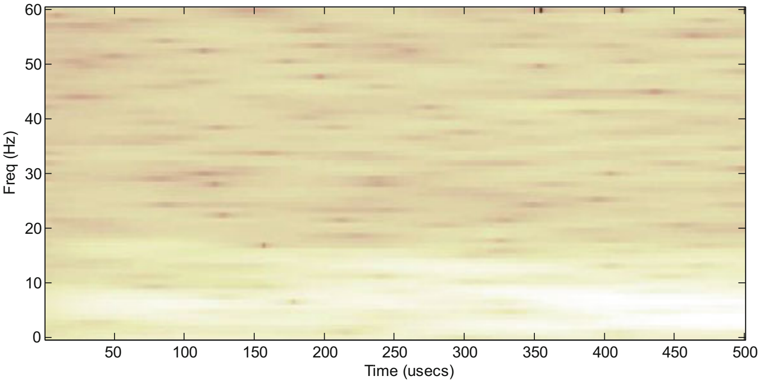


Fig. 9 The short-time Fourier transform of full cement

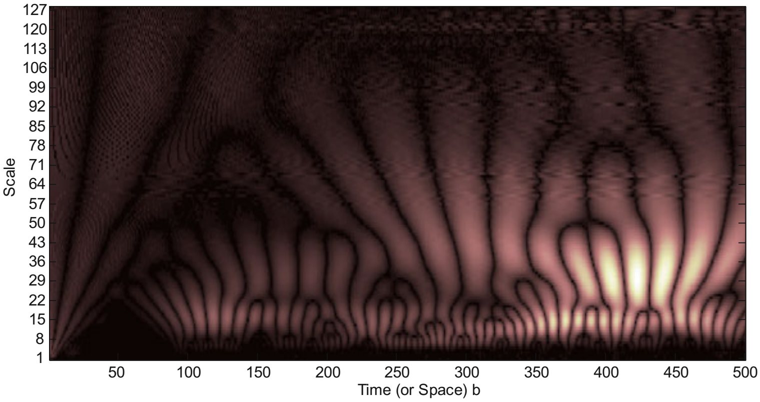


Fig. 10 Continuous wavelet transform of full cement

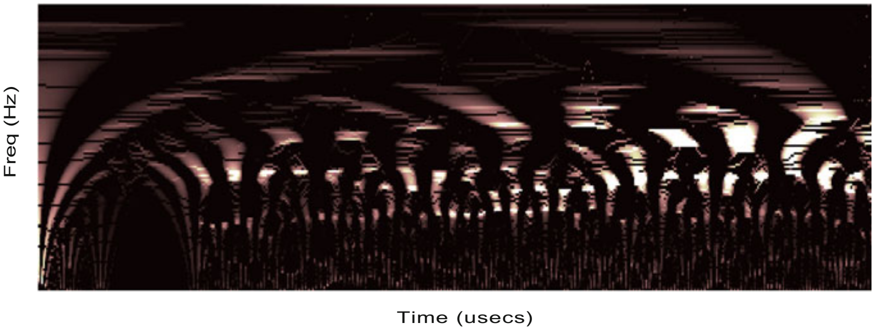


Fig. 11 Synchrosqueezing transform of full cement

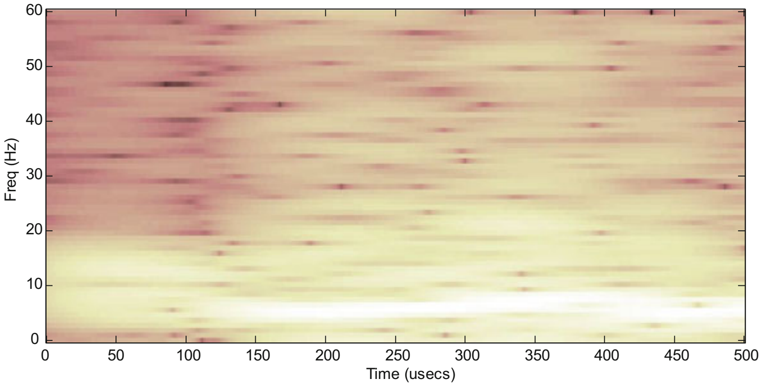


Fig. 12 The short-time Fourier transform of good casing and bad formation

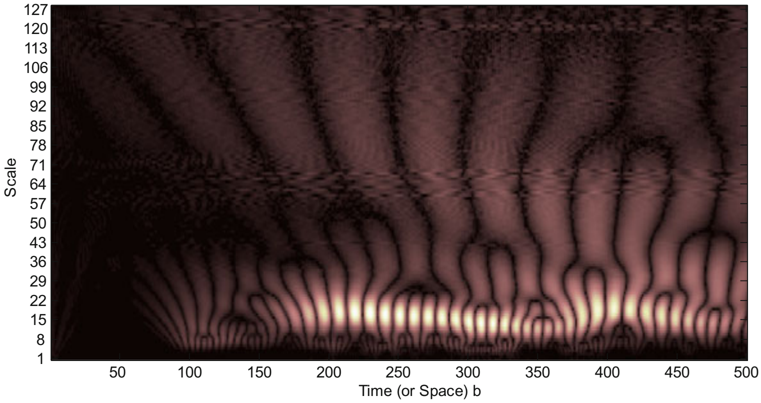


Fig. 13 Continuous wavelet transform of good casing and bad formation

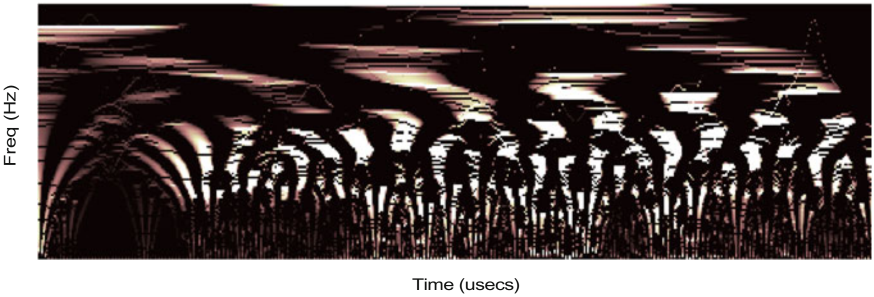


Fig. 14 Synchrosqueezing transform of good casing and bad formation

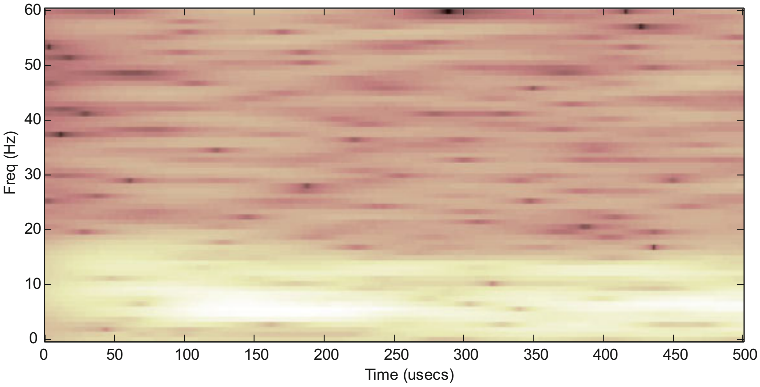


Fig. 15 The short-time Fourier transform of bad casing and good formation

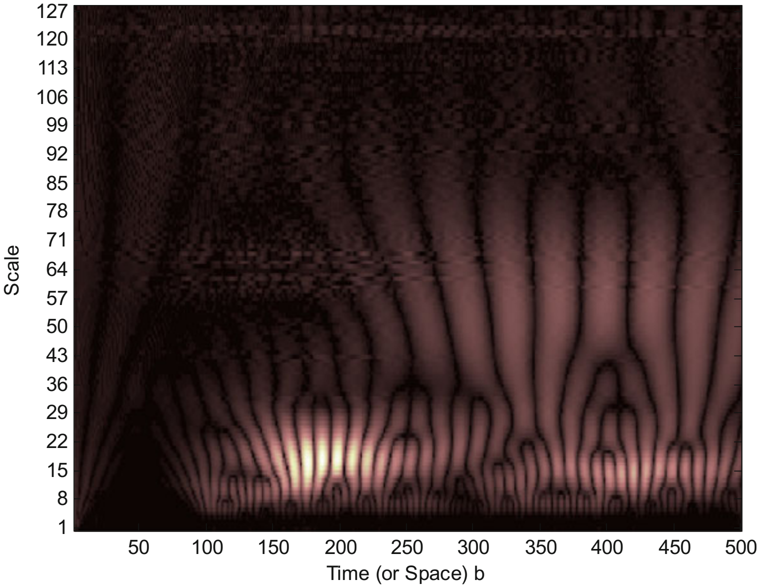


Fig. 16 Continuous wavelet transform of bad casing and good formation

of strapping, the signal transmitted from the transmitter without undue distortion reaches the receiver, which causes the state to indicate in the time-frequency transformations a more energetic state, These states are hardly detectable in Fourier transforms of short time, but in SST, as you can see, the resolution is much better.

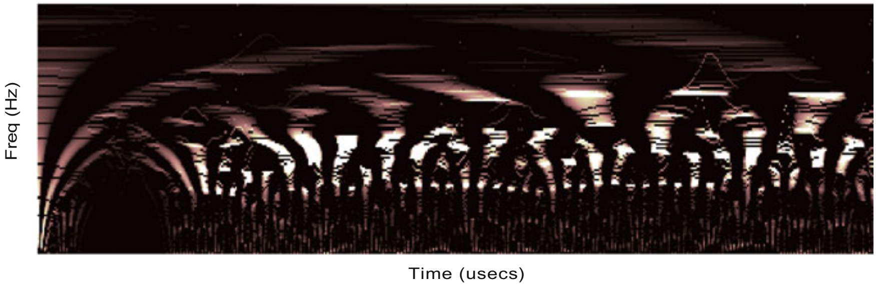


Fig. 17 Synchrosqueezing transform of bad casing and good formation

References

1. Padue GH, Morris RL (1963) Cement bond log-a study of cement and casing variables. J Petrol Technol:545–554. May 1963
2. Pickett GR (1963). Acoustic character logs and their application in formation evaluation. J Petrol Technol:15, June 1963
3. Frish G, Fox P (2005) Advances in cement evaluation tools and processing methods allow improved interpretation of complex cement. Society of Petroleum Engineers, (SPE), Technical conference and exhibition, 2005
4. Song RL, Liu et al (2012) Numerical simulation of sector bond log and improved cement bond image. Geophysics 77(4):95–104
5. He X, Chen H, Wang XM (2014) Ultrasonic leaky flexural waves in multilayered media: cement bond detection for cased wellbores. Geophysics 79(2):7–A11
6. Qiao WX, Ju XD et al (2006) Downhole acoustic arc array transmitter with controlled azimuthal directivity. China Patent, 2006
7. Che XH, Qiao WX (2009) Numerical simulation of an acoustic field generated by a phased arc array in a fluid-filled borehole. Petrol Sci 6(3):225–229
8. Che XH, Qiao WX, Wang RJ et al (2014) Numerical simulation of an acoustic field generated by a phased arc array in a fluid-filled cased borehole. Petrol Sci 11(3):385–390
9. Partyka G et al (1999) Interpretational applications of spectral decomposition in reservoir characterization. Lead Edge 18(3):353–360
10. Chakraborty A, Okaya D (1995) Frequency-time decomposition of seismic data using wavelet-based methods. Geophysics 60(6):1906–1916
11. Han J, van der Baan M (2013) Empirical mode decomposition for seismic time-frequency analysis. Geophysics 78(2):1–19
12. Daubechies I, Males S (1996) A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models. In: Wavelets in medicine and biology. CRC Press, Boca Raton. pp 527–546
13. Daubechies I et al (2011) Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool. Comput Harmon Anal 30(2):243–261
14. Thakur G et al (2013) The synchrosqueezing algorithm for time-varying spectral analysis: robustness properties and new pale climate applications, Signal Process 93(5):1079–1094
15. Bigelow EL (1989) A practical approach to the interpretation of cement bond logs, Society of Petroleum Engineers, SPE conference, 1989
16. Cased Hole Log Interpretation Principles/Applications, (1989) Schlumberger, 1989
17. Nayfeh TH et al (1984) The fluid compensated cement bond log, Society of Petroleum Engineers (SPE), vol. 1, Aug 1984

18. McGhee BF, Vacca HL (1980) Guidelines for improved monitoring of cementing operations. In: Society of petro physicists and well log analysts, 21th logging. Symposium, July 1980
19. Seryasat OR, Habibi M, Ghane M, Taherkhani H (2014) Fault detection of rolling bearings using discrete wavelet transform and neural network of SVM. Adv Environ Biol 8(6):2175–2183
20. Thakur G et al (2013) The synchrosqueezing algorithm for time-varying spectral analysis: robustness properties and new pale climate applications. Signal Process 93(5):1079–1094

Detection of Malicious Node in Centralized Cognitive Radio Networks Based on MLP Neural Network

Zeynab Sadat Seyed Marvasti and Omid Abedi

Abstract The cognitive radio network (CRNs) has been developed in recent years for the optimal use of Available vacuum in the frequency spectrum. In this network, Cooperate Spectrum Sensing (CSS) is used to combine the observations of all users. In CRNs, security is one of the most important problems spectrum sensing data falsification (SSDF) attack is one of major challenges for CSS in CRNs, in which Malicious user are among honest users trying to change the information sent to the fusion center and thus make the fusion center's wrong decision. In this paper, a method for defense against SSDF attack is proposed using MLP-based neural network. In this scheme, the weights of secondary users were constantly updated and finally the sensing results were combined in the fusion center based on their trusted weights. Simulation results show that the proposed scheme can significantly reduce the effects of Spectrum Sensing Data Falsification (SSDF) attack even percentage of malicious users are more than trusted users.

Keywords Cognitive radio · MLP · Malicious user · SSDF attack

1 Introduction

Cognitive radio (CR) technology is suggested to improve the frequency spectrum utilization by opportunistic access to the free space of the licensed frequency bands in the presence of the licensed primary users. In CRNs, the spectrum allocated to

Z. S. S. Marvasti
Department of Computer Engineering, Yazd Branch,
Islamic Azad University, Yazd, Iran
e-mail: zeinab.seiedmarvasty@gmail.com

O. Abedi (✉)
Department of Computer Engineering,
Shahid Bahonar University of Kerman, Kerman, Iran
e-mail: oabedi@uk.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_67

primary users is not used fully and secondary users could utilize the idle spectrum as long as primary users are absent [1].

One of the most important challenges in CRNs is reliable spectrum sensing with the aim of finding the vacant spectrum band [2]. Spectrum sensing procedure can be affected by shadowing, multi-path fading and hidden terminal problem. So, Spectrum sensing by a user may not be reliable. To mitigate these effects, Cooperate Spectrum Sensing (CSS) has been proposed in recent years [3, 4].

In CSS schemes, data fusion and decision making can be done centralized or distributed. In the distributed mode, since there is no fusion center for decision making, secondary users exchange the spectral sensing results between themselves. But in centralized mode, as shown in Fig. 1, all spectrum sensing operations are managed and coordinated by a central controller called the fusion Center (FC). The FC collects and processes the information of presence or absence of primary users from secondary users and then it decides itself on whether or not the primary user signal is present [4, 5]. FC is responsible for the coordination between secondary users and the final decision.

Even though the Cooperate Spectrum Sensing (CSS) is a solution to improve the detection precision, but it is vulnerable to spectrum sensing data falsification attack (SSDF attack). In an SSDF attack, some malicious users intentionally report incorrect local sensing results to the FC and disrupt the global decision, which will result in interference or inappropriate Throughput. The structure of the SSDF attack is shown in Fig. 2.

There are three different types of SSDF attack, classified based on the strategy in sending the false sensing data, as follows:

- Always Yes Attacker: this type of attackers always declares that the primary user is active without sensing spectrum. The motivation of Always-Yes attacker is to prevent other SUs from accessing the spectrum, which gives him the chance to use the spectrum alone. Thus, it causes the optimal non-use of the spectrum.

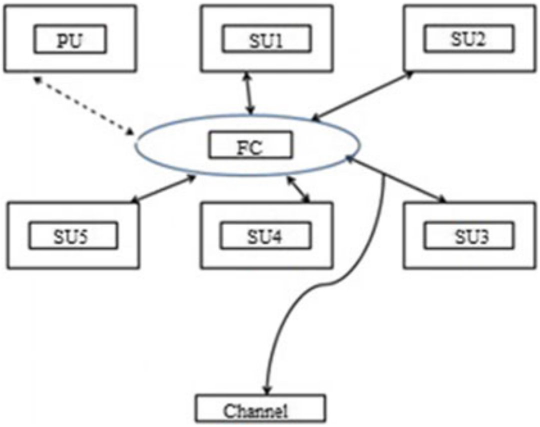
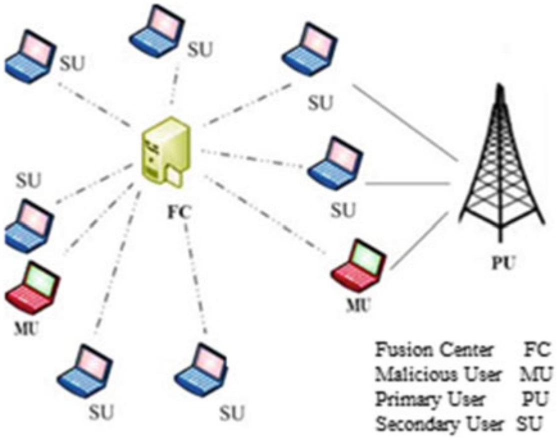


Fig. 1 The centralized structure of Cognitive radio networks

Fig. 2 Spectrum sensing data falsification attack



- Always No Attacker: this type of attackers always reports an absence of primary signal. This will cause reduce the performance of CRN and primary users.
- The third type of attacker that called smart is the one who always reports the opposite of its local spectrum sensing result. Thus, they mislead the system once in a while, but they behave correctly during the rest of the time.

In [6], authors have used the ARC classification algorithm, which uses the history of nodes to identify the cluster. In this way, the attacker's detection is somewhat feasible, but the initial number of the cluster plays an important role in the performance and efficiency of the algorithm, and even the results may not be as expected.

In [7], a study based on authentication-based cryptography was performed to detect SSDF attacks. In this way, the false alert rate is lowered for primary user identification.

A study was also conducted in [8] to Detection of SSDF attack using the SVDD algorithm, which uses a data descriptor vector to detect an attack. The SVDD algorithm distinguishes malicious users from trusted users and removes them from the decision-making phase. Simulated results indicate that this algorithm works well in detecting malicious users. But it is not able to detect "smart" attackers.

In [9], the Wald's Sequential Probability Ratio Test (WSPRT) was proposed to detect the SSDF attack. In this method, each node wants to sense the spectrum, collect local reports from the neighboring nodes. In this way, if the report of sending data and the final decision is the same in FC, the weight of that second user is added to one unit and otherwise the weight of one unit is reduced. The WSPRT consists of two main stages. First, each node initially has a zero-credit value with a local-area report right, the value of credit will increase to one. The second step is the actual testing hypothesis of the WSPRT, which is based on the probability sequence test. In this method, only two attack models "always yes" and "always good" are examined.

Also, in [10], researchers used Dixon Test statistical method to detect malicious users and compared with other statistical methods such as Grubb Method1 and Grubb Method2. It was concluded that the result was Dixon's method is better than the other two methods. this research also has limitations in the number of malicious users.

In the studies that have been done so far, the two first types of attackers have been investigated and the third one, which is smart attacker, has been less studied. So, we need a method that, beside the detection of "always Yes" and "always No" attackers, can detect the smart attacker with a lower error rate. according to this, in this paper, it has been utilized by the intelligence feature and the ability to learn of radio cognitive network and its similarity to the neural network.

The rest of the paper is organized as follows. In Sect. 2, the basic theory will be expressed. In Sect. 3, the system model will be described. The proposed scheme is introduced in Sect. 4, and the numerical results are depicted in Sect. 5. Section 6 concludes the remarks.

2 Basic Theory

To describe the function of radio cognitive, two parameters of the probability of Missed Detection and the probability of a False alarm are introduced.

2.1 Missed Detection

Missed detection occurs when the user does not detect the primary user activity in the spectrum sensing, or, in other words, occupy a frequency spectrum as an Open Spectrum. In fact, P_m Indicates the degree of interference that a secondary user can make for primary user based on his own decision.

$$P_m = 1 - P_d = P\,\{\text{decision} = H_0|H_1\}. \tag{1}$$

H_0 represents the absence of the primary user and the H_1 is considered to be the primary user on its dedicated channel.

2.2 False Alarm

It indicates that an empty frequency band that can be used as an Open Spectrum can be mistakenly considered as occupied by a frequency band. This is shown with P_{fa}. In fact, P_{fa} represents the percentage of Open Spectrum that is not optimally used.

It is important to note that a false alarm does not have a detrimental effect on the activity of primary users and only reduces the use of the maximum useful spectrum capacity for secondary users.

But in the event of a false detection, we will see interference between the primary and secondary users.

3 System Model

In this paper, have been used the neural network to minimize the interference of the secondary users with the primary user and to cover all the Malicious user's states. Unlike the proposed proposals [6–10], which require continuous training, neural network training is designed to detect a potential attacker only once in an offline mode when the observation process is static. When the neural network is trained, the complexity of computing in online mode is significantly reduced.

In this paper, two sets of different data are considered for training and validation of neural networks:

(1) The training set, which is used to create the model, and network parameters including weights and thresholds are determined.
(2) Validation set, which is used to evaluate the network performance by keeping the parameters maintained during training.

The structure for this neural network is intended to be implemented includes 50 entries, a hidden layer with 10 neurons and 1 output. The overall structure of this network is shown in Fig. 3.

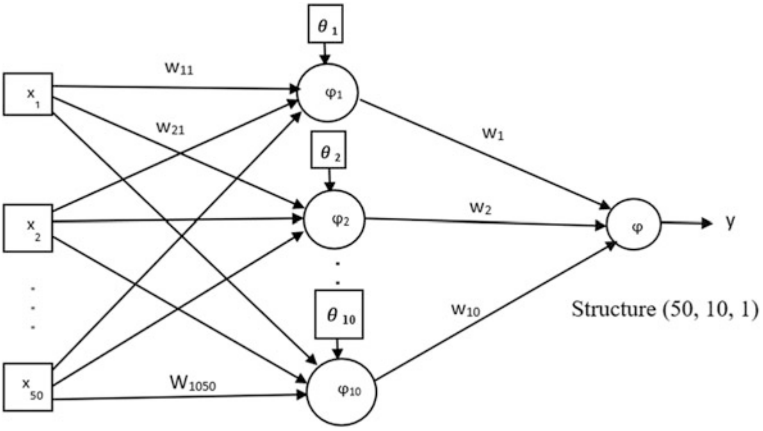


Fig. 3 Structure of the MLP network

4 Proposed Scheme

4.1 Database and Basic User Traffic

To simulate, a single-channel cognitive radio system with a primary user and 50 secondary users is considered.

Assuming a slotted system, users in the network work with the time slot model, and at the beginning of each slot, the secondary user sense the primary user activity for a short time and send the result to the control center he does. The control center decides on the results received from all secondary users.

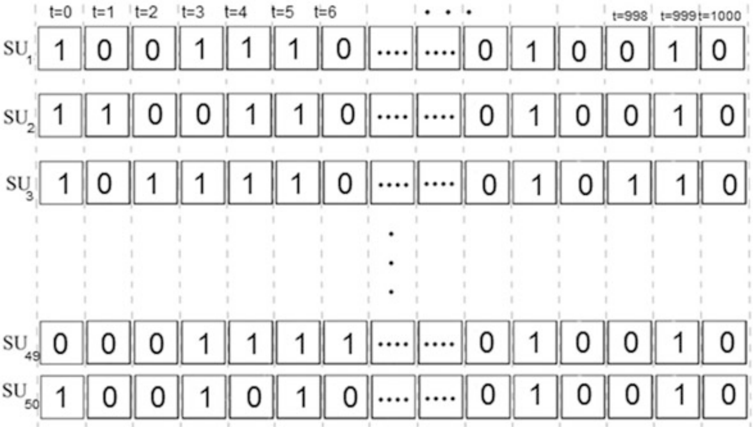
To model the primary user and secondary user traffic, the binary series generated randomly with the Round (rand ()) command in the Matlab environment for 50 secondary users over 1000-time periods has been used as a database.

In Fig. 4, an example of a database is shown. 700 slots of these 50 users are used to learn the neural network and determine its weights. 50-time slots for Validation and 250 other slots have been used to test the MLP's neural network.

In order to detect a Malicious user in a cognitive radio network, we consider the traffic conditions to be static. In this way, the average user traffic is constant over a channel and the primary user traffic will not change over time.

4.2 Normalization and de Normalization of Data

Because the inputs are in the form of a signal of 0 (The absence of the primary user) and 1 (Primary user presence). Input zeroes will result in zero output and eliminate



	t=0	t=1	t=2	t=3	t=4	t=5	t=6		• • •				t=998	t=999	t=1000
SU_1	1	0	0	1	1	1	0	....	....	0	1	0	0	1	0
SU_2	1	1	0	0	1	1	0	....	....	0	1	0	0	1	0
SU_3	1	0	1	1	1	1	0	....	....	0	1	0	1	1	0
								⋮							
SU_{49}	0	0	0	1	1	1	1	....	....	0	1	0	0	1	0
SU_{50}	1	0	0	1	0	1	0	....	....	0	1	0	0	1	0

Fig. 4 An example of the database used

the effect of this input on the output and it does not allow the right decision to the fusion center, normalizing the data is done before the start of the training phase.

Therefore, we first normalize the inputs in accordance with relations (2) in the interval [1, 1.5].

$$x_{scaled} = x_{real} \times s + o. \tag{2}$$

In which, o and S are obtained from relations (3) and (4).

$$S = \frac{H_i - L_o}{Max - Min} = \frac{1-0}{1+1} = \frac{1}{2} \tag{3}$$

$$O = \frac{Max \times L_0 - Min \times H_i}{Max - Min} = \frac{1 \times 0 - 1 \times -1}{1+1} = 1 \tag{4}$$

$$x_{scaled} = \begin{cases} 0 \times \frac{1}{2} + 1 = 1 \\ 1 \times \frac{1}{2} + 1 = \frac{3}{2} \end{cases} \tag{5}$$

According to Eq. (5), input 0 and 1 are normalized to outputs 1 and 1.5. The output after training is De normalized in accordance with Eq. (6).

$$\mathrm{x_{real}} = \frac{\mathrm{y_{scale}} - \mathrm{o}}{\mathrm{s}} \begin{cases} \text{Output} = 1 & \rightarrow \quad \text{Real} = 0 \\ \text{Output} = \frac{3}{2} & \rightarrow \quad \text{Real} = 1 \end{cases} \tag{6}$$

To prevent tend of a neural network to a very large negative or positive number; the weights are limited between zero and one $0 \leq w_i \leq 1$.

4.3 Training Phase

In the training phase, input include 50 secondary users in 700 times slots. The target is also a 1 × 700 matrix and is based on sensing the majority of users. If the majority of users detect the channel occupancy, the target is "1" and otherwise the target is "0".

At first, weights will be considered the same and equal to 0.5 for all users. Because at first FC does not have any documentation for detecting attackers from an honest Secondary user. The more time it passes, the users whose reports are like the real status of the primary user will have a higher weight and the users whose reports are opposite of the real status of the primary user will have a less weight.

The thresholds and weights between the inputs and the hidden layer and the hidden layer and output are obtained by network training. In this paper, the learning factor is Considered equal to 0.5 (μ = 0.5) to prevent the rapid convergence of the algorithm, as well as to cover the entire search space. Also, the convergence condition, or, in other words, the condition for stopping the change algorithm is considered less than 0.002 Validation output in 6 consecutive repetitions. With Considering this stop condition, the outputs 1.002 and 0.998 are considered as “1” and the outputs +0.002 and −0.002 are considered as “0”.

4.4 Test Phase

At this point, the input contains 50 users in 250 times slots. No information is available about the honest or attacker percentage of users. Weights, thresholds and functions are from the training phase, and the output is obtained with the parameters of the previous step and the same training phase.

4.5 Detection of “Always Yes” and “Always No” Attacker

In this article, an algorithm has been added to the neural network that “Always Yes” and “Always No” Attacker can be detected. The algorithm is described below in order to detect both of these attackers:

```
    Suoutput = 0;
    finaloutput = -1;
    For k = 0:1:4
        Suoutput = sum(suArray);
        fcoutput = sum(fcArray);
      end
      if suoutput ==5 && suoutput > fcoutput
        finaloutput = 1;
      elseif suoutput ==0 && fcoutput ∼=0
        finaloutput = 0;
      else
        finaloutput = -1;
      end
end
```

As shown in the above algorithm, in a period of 5 s the sum of the numbers sent from each of the secondary users is calculated. If the calculated number for each secondary user is equal to 5 and larger than the number verified by the FC, that user will be referred to as the “always yes” attacker and if the calculated number is zero and the sum of the numbers verified by the FC is Opposite zero, then the user will be referred to as “always No”.

4.6 Detection of Smart Attacker

At this point, the input contains 50 users in 250 times slots. No information is available about the honest or attacker percentage of users. Weights, thresholds and functions are from the training phase, and the output is obtained with the parameters of the previous step and the same training phase.

In detecting "smart" attacker to ensure that the noise is not the cause of the decision was wrong, this is done within a 5 s time interval, and the sum of the Absolute value of the difference between each input and the output is divided into 5. If this difference is greater than 0.5 for each user according to Eq. (7), it is known as an attacker. But if, according to (8), it is less than 0.5, then the user is honest. The closer the number is to 0.5, indicating that shadowing or fading has occurred or the channel noise is high and the secondary user can not recognize the channel.

$$\frac{\left|\sum_{t=5} Input - \sum_{t=5} Output\right|}{5} > 0.5 \rightarrow Malicious\,\text{User}. \tag{7}$$

$$\frac{\left|\sum_{t=5} Input - \sum_{t=5} Output\right|}{5} < 0.5 \rightarrow Honest\,Secondary\,User. \tag{8}$$

5 Simulation Results

In Fig. 5, the weights are updated for an honest second user and in Fig. 6 weighted values for an attacker in the training phase are shown.

As shown in Figs. 5 and 6, the initial weight for both the secondary user and the attacker is initiated from 0.5 because at the start of the algorithm, there is no correct recognition of the honest or attacker user of the secondary user. As the training process, the weight of the honest user will increase and close to one and the weight of the attacker will decrease and reaches zero. Neural network with Intelligent makes it possible in during the several of epochs, honest secondary users will gain more weight and have a greater share in proportion to the attacker in deciding at the FC center.

The simulation results of the users' data with the MLP neural network are shown in Fig. 8 in three phases of training, validation and testing in a time slot.

As shown in Fig. 7, testing of user data with trained weights and thresholds in the training procedure converges during the 421rd epochs. In other words, changes less than 0.002 Validation output in 6 consecutive repetitions.

The Probability of correct sense is given by:

$$P_D = \frac{M_D}{M_R} \times 100. \tag{9}$$

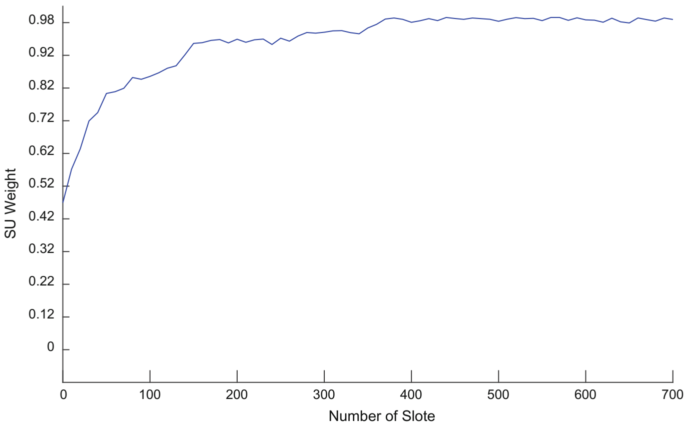


Fig. 5 Updated weights for an honest secondary user

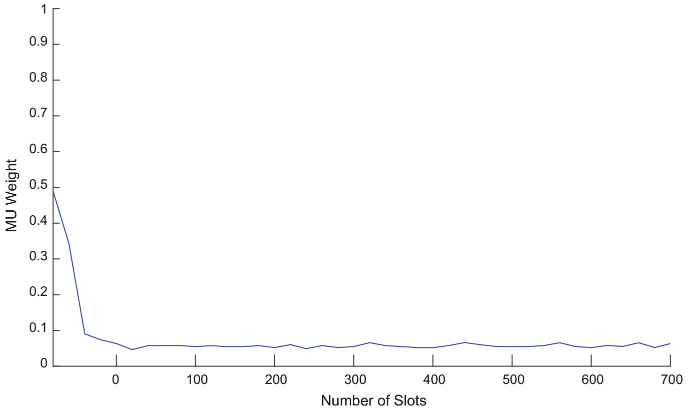


Fig. 6 Updated weights for an attacker

where M_D the number of malicious users is detected, M_R is the number of real malicious users and P_D is The Probability of correct sense.

In Fig. 8, the probability of correct sense is shown for the MLP algorithm and the SVDD algorithm in [8].

As shown in Fig. 8, as the number of time slots increases, the probability of correct sense of spectrum remains almost constant by the SVDD algorithm. But in the MLP algorithm, as already mentioned, by increasing the number of slots and according to weight updates, honest users will have more weight and their effect on the decision on the spectrum will increase, so the probability of correctly sense increases.

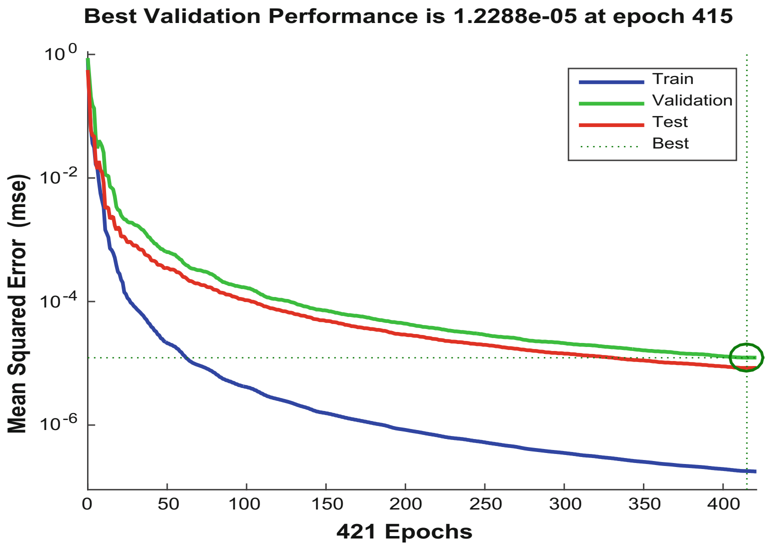


Fig. 7 Results of simulation of secondary users' data with the MLP neural network in two stages of teaching and testing for a time slot

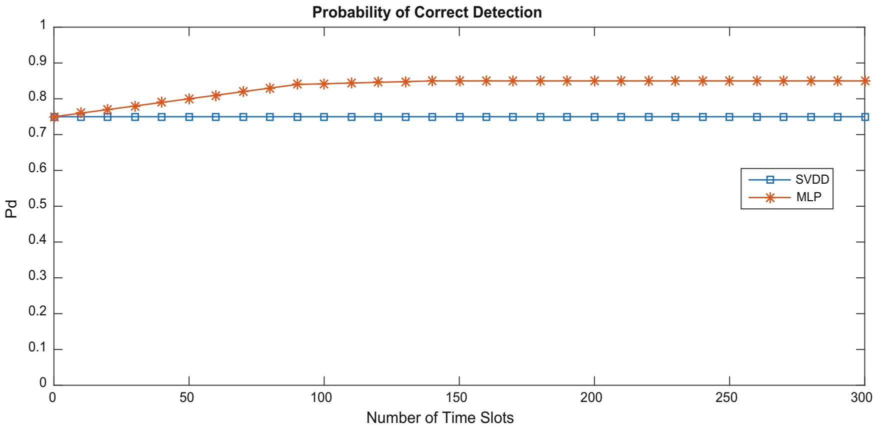


Fig. 8 The Probability of correct sense of spectrum by the SVDD algorithm and the MLP algorithm

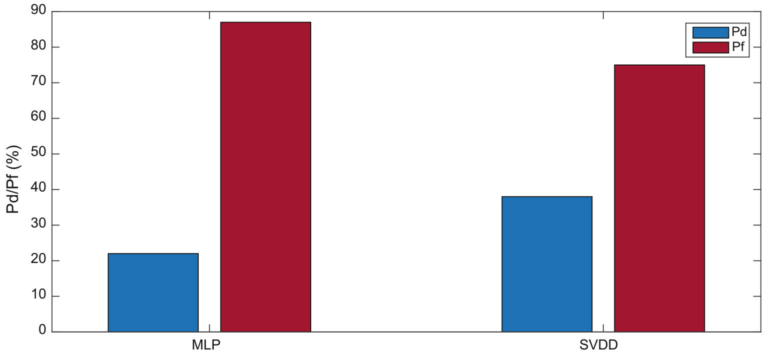


Fig. 9 False alarm percentage by two SVDD and MLP algorithms in FC

Another important parameter is the possibility of False alarm and modified as:

$$P_{FA} = \frac{S_M}{S_T} \times 100. \tag{10}$$

where S_M the number of secondary users is detected as malicious users and S_T is Total number of secondary users.

Figure 9 shows the comparison of the False alarm percentage with two MLP and SVDD algorithms in FC.

As shown in the diagram, the MLP algorithm has fewer false alarms than the SVDD algorithm.

The higher the number of slots when used to train the MLP neural network, the better weights and thresholds are taught, and the percentage of false alarms is reduced.

The higher the number of slots which is used to train the MLP neural network, the lower the percentage of false alarms.

The SVDD algorithm produces a lot of false alarms Due to the lack of detection of the smart attacker.

6 Conclusion

In order to Spectrum Accurate Sensing, decrease computing time and also identify detecting all three types of attackers in a SSDF attack, in this paper, the MLP neural network is proposed for decision at the FC center. The MLP's neural network reduces online computing due to network training offline so it's faster. The simulation results show that with using the MLP algorithm, the probability of false alarms decreases and the Probability of correct sense increases. This algorithm can have a good result even when the attackers in the cognitive radio are more than half the total number of users.

Reference

1. Zhang T, Safavi-Naini R, Li Z (2013) ReDiSen: reputation-based secure cooperative sensing in distributed cognitive radio networks. In: IEEE International Conference on Communications (ICC), Budapest, 2013, pp 2601–2605
2. Meghanathan N (2013) A survey on the communication protocols and security in cognitive radio networks. Int J Commun Netw Inf Secur (IJCNIS) 5(1)
3. Akyildiz IF, Lo BF, Balakrishnan R (2011) Cooperative spectrum sensing in cognitive radio networks: a survey. Phys Commun 4(1):40–62
4. Chen Q, Motani M, Wong W-C, Nallanathan A (2011) Cooperative spectrum sensing strategies for cognitive radio mesh networks. IEEE J Sel Topics Signal Process 5(1):56–67
5. Zou Y, Yao Y-D, Zheng B (2011) A selective-relay based cooperative spectrum sensing scheme without dedicated reporting channels in cognitive radio networks. IEEE Trans Wireless Commun 10(4):1188–1198
6. Amutha Priya R, Nandhakumar S (2015) Attack prevention for spectrum sensing data falsification attacks in cognitive radio networks using ARC. IJARSET 2(3)
7. Dharani D, Sharmila A (2015) A robust collaborative spectrum sensing decision against SSDF attack in cognitive radio network. Int J Sci Prog Res (IJSPR), 11(01)
8. Farmani F, Jannat-Abad MA, Berangi R (2011) Detection of SSDF attack using SVDD algorithm in cognitive radio networks. IEEE
9. Chen R, Park JM, Bian K (2008) Robust distributed spectrum sensing in cognitive radio networks, INFOCOM 2008. IEEE 27th Conference on Computer Communications. pp 1876–1884
10. Kalamkar SS, Roychowdhury A, Banerjee A (2012) Malicious user suppression for cooperative spectrum sensing in cognitive radio networks using Dixon’s outlier detection method, IEEE

FIR Filter Realization Using New Algorithms in Order to Eliminate Power Line Interference from ECG Signal

Akbar Farajdokht and Behbood Mashoufi

Abstract Biomedical recordings are often contaminated by power line interference (PLI). Notch filtering is one of the most common filtering that suppresses the major PLI as well as its harmonics in the electrocardiographic recording. A finite impulse response (FIR) filter is one of the fascinating methods to filter interference so that obtain a relatively pure signal. In this paper, the authors focus on providing new algorithms (HSTAE, MPSO) for reducing the hardware complexity. Using these algorithms, in addition to reducing the hardware volume, the frequency response and quality of the output signal will be improved. The simulation results showed that the proposed algorithms remarkably reduce the hardware complexity of the FIR filter for eliminating PLI from the electrocardiogram (ECG) signal.

Keywords ECG · PSO · PLI noise · FIR filter · Filter realising

1 Introduction

Electrocardiography (ECG) is a common diagnostic medical test that records the electrical activity of the heart by placing small electrodes on the skin. These electrodes can detect any small electrical changes on the skin. In a standard ECG, 12 electrodes are positioned on the surface of the chest and limbs of the patient to record the total magnitude of the heart's electrical potential from 12 different angles. In this technique, the overall magnitude and direction of electrical depolarization of the heart during the cardiac cycle can be captured [1].

The voltage versus time graph provided by this noninvasive technique is known as an electrocardiogram. An ECG signal is shown in Fig. 1. The first wave in the ECG is called the P wave which is generated as a result of sequential activation (depolarization) of the right and left atria. The duration of atrial contraction should not be more than 0.11 s and the amplitude should not be more than 3 mm. The most

A. Farajdokht (✉) · B. Mashoufi
Microelectronics Research Laboratory, Urmia University, Urmia, Iran
e-mail: akbar.farajdokht.1992@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_68

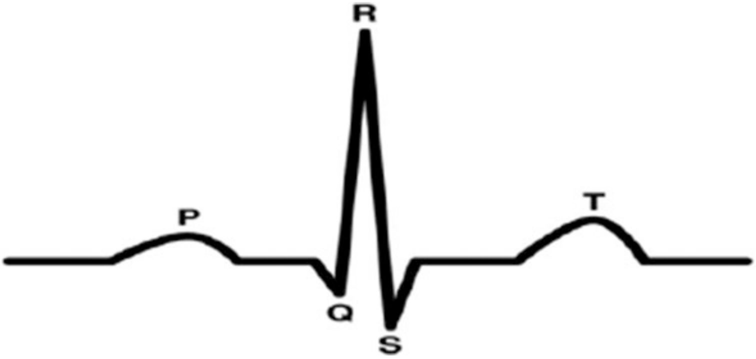


Fig. 1 Typical ECG signal in time domain

important complex in the ECG signal is the quantron resonance system (QRS) which shows right and left ventricular depolarization. The duration (QRS interval) which is measured from the beginning of the QRS complex to its end is normally 0.05–0.10 s. The T wave represents the repolarization of the ventricles. The height of the T wave should not be greater than 5 mm. Finally, the U wave which reflects repolarization of the papillary muscles [2] is seen as a different section of an ECG signal.

The ECG signals are weak in nature and sensitive to even small noise. These signals are vulnerable to corrupt by several noises including baseline wander, power line interference (PLI), electromyographic noise, and electrosurgical noise [3]. In the clinic, the ECG plays a fundamental role in the primary diagnosis and monitoring of the heart health, and each part of its signal conveys the information about a specific part of the heart. Therefore, it is important to create some conditions to receive a rather pure signal without any noise. Filtering is an approach to remove the undesired noises from the signal. For each noise, depending on the frequency of its occurrence, researchers have to design a specific filter. The PLI as the most major noises polluting the ECG signal is defined in Sect. 2. So far, several studies have been conducted to eliminate the noise of the PLI from the ECG signal [4–7]. In this paper, we focused on PLI elimination using digital filters. In the Sect. 2 PLI is defined and in the Sect. 3 finite impulse response (FIR) notch FIR filter for filtering ECG is defined. In the Sects. 4 and 5 PSO and its equation and various methods of realising the FIR filter are defined respectively. Subsequently in the Sect. 6 we introduce our algorithms then in Sect. 7 simulation result are shown. Discussion and conclusions are provided in the Sect. 8. The simulations in this paper carried out using MATLAB and used ECG signals [8].

2 Power Line Interference (PLI)

The PLI of 50/60 Hz is the source of interference and it contaminates the ECG recordings which are:

(a) Electromagnetic interference by power line.
(b) Electromagnetic field by the equipment which are located nearby. The signal component holds harmonics with different amplitudes and frequencies. The harmonics frequency is an integral multiple of the fundamental frequency for example 50 Hz.
(c) Stray effect of the alternating current fields owing to loops in the cables.
(d) Inappropriate grounding of the ECG machine or the patient.
(e) Electrical equipment such as air conditioners, elevators and X-ray units draw heavy power line current, which induce 50 Hz signals in the input circuits of the ECG machine [9].

Identifying QRS complex is the most important point in ECG signal processing. However, this signal becomes distorted by PLI, making it unable to make right process on this signal. A noisy ECG signal is shown in Fig. 2, in which the

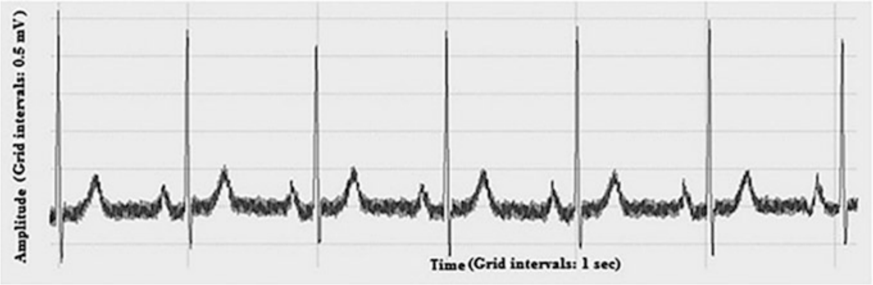


Fig. 2 ECG signal corrupted with 60 Hz PLI

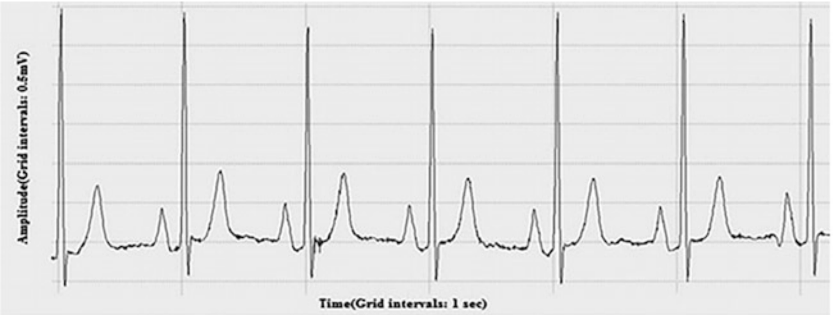


Fig. 3 Denoised ECG signal using an FIR equiripple filter

frequency spectrum of an ECG signal corrupted with PLI. The occurrence of 60 Hz PLI in Fig. 3 is clear. The notch filter which cuts 60 Hz frequency of corrupted ECG signal is the way to eliminate PLI from ECG.

3 FIR Notch Filter

There are plenty of ways to design notch FIR filter which their cut frequency is 50/60 Hz [10, 11]. In this paper we decided to choose FIR filter because of the following advantages:

1. Stability (there is no feedback).
2. Linear phase design.
3. The hardware complexity is less.
4. Unlike IIR filters, only main signal samples are used.

The main disadvantage of FIR filters is that considerably more computation power in a general purpose processor is required compared to an IIR filter with similar sharpness or selectivity, especially when low frequency (relative to the sample rate) cutoffs are needed. However, many digital signal processors provide particular hardware features to make FIR filters approximately as efficient as IIR for many applications. For designing the FIR filter, again we have various choices such as windowing method, LMS method and so on. We tried to design FIR filter with following features:

1. Filter with a less amount of hardware complexity
2. Output signal and frequency response of filter was considerable better.
3. SNR of designed filter was acceptable.

The specifications of the filter are designed as follows:

- Order 14
- SNR of output signal 8.36
- Filter Coefficients will be produced by the Kaiser—window method.

By using this filter with less order (realisable with algorithms), peak of the signal was easily detectable. In order to have a clear signal and better SNR, we should have a filter with higher order, but due to manufacture limitations, reduced power consumption and better signal detection, we need to use a smaller order filter. Hence a filter with above properties is the desire one.

4 Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) technique combines social psychology principles in socio-cognition human agents and evolutionary computations. Generally, PSO is considered as simple in idea, easy to implement, and computationally efficient. Differing from other heuristic methods, PSO is a flexible and well-balanced mechanism to enhance the global and local exploration capabilities, making PSO less vulnerable to getting trapped into local optima unlike GA, and simulated annealing etc. [12]. Two solutions are verified in PSO. One is the personal best (pbest), which represents the best solution achieved by a single particle, corresponding with the personal experience of bird. While the global best (gbest) is the best solution achieved by all the particles, corresponding with the genius of the whole flock and always regard as the best solution. Namely, each particle tries to modify its position using the following information:

- The distance between the current position and the pbest
- The distance between the current position and the gbest.

A simple alternative of the PSO algorithm works through having a population (called swarm) of candidate solutions (called particles). These particles are moved around in the search-space based on a few simple formulae [13]. The movements of the particles are directed by their individual best known position in the search-space plus the whole swarm's best known position. When adjusted positions are being revealed, these will then come to guide the movements of the swarm. The process is repeated and by doing so it is hoped, but not guaranteed, that a reasonable solution will ultimately be revealed. In PSO, each particle represents a potential solution to a problem, and searches around in a multi-dimensional search space [14]. A particle i is described by its position xi and velocity vi. During the search, each particle adjusts its location according to its own experience and the experience of its all adjacent particles, making use of the best position encountered by it (pi) and all the others (pg). Mathematically, velocities of the particle vectors and the searching point in the solution space may modify by (1) and (2):

$$\mathrm{v}i(t+1) = wvi(t) + c1r1(pi(t) - xi(t)) + c2r2(pg(t) - xi(t)) \tag{1}$$

$$xi(t+1) = vi(t+1) + xi(t) \tag{2}$$

where w is the inertia weight; $vi(t)$ is the velocity of particle ith vector; $C1$ and $C2$ are the positive weighting factors; r1 and r2 are the random numbers between 0 and 1; xi(t) is the current position of ith particle vector pi(t) is the personal best of the ith particle; $p_g(t)$ is the group best of the group. In the Eq. (1); $\mathrm{w} = \mathrm{w}_{\max} - (\mathrm{w}_{\max} - \mathrm{w}_{\min}) \times (\frac{iter}{iter\,max})^{\alpha}$, w_{max} is the maximum inertia of weight and w_{min} is the minimum inertia of weight, iter is the iteration number at the current time step, $iter_{max}$ is the maximum number of iterations, α is the nonlinear regulation coefficient during the iterations. Large values of w facilitate exploration with increased diversity, whereas small values promote local exploitation dynamic inertia weight and mutation mechanism. As for the velocity, it is the key point to control the proceeding process; many researchers have risen different types of adjustments to improve modified PSO by changing the velocity equation, specially, the inertia weight w later. By satisfying the stability conditions in PSO, the speed and precision of the convergence are improved [15]. FIR filter designed by Remez algorithm results in so many ripples in stopband and requires more coefficients for sharp cutoff. To overcome this problem PSO algorithm use to minimize error between the ideal and Remez filter [16]. The error equation is:

$$\mathrm{error} = \mathrm{E(w)} \ \times [\mathrm{H_i}(\mathrm{e^{jw}}) - \mathrm{H_d}(\mathrm{e^{jw}})] \tag{3}$$

where $H_i(e^{jw})$ is frequency response of filter using MPSO and $H_d(e^{jw})$ is frequency response of filter by Remez algorithm. In this paper we focused on inertia weight and how to improve it in order to achieve best solution of the particle.

5 FIR Filter Realising

After a filter is designed, it should be realised by developing a signal flow diagram that describes the filter in terms of operations on sample sequences. A given transfer function may be realised in many ways. Filter realising is a vital step in practical cases which directly deals with the amount of hardware. Main goal of realization can be introduction of some methods by researchers in order to save the hardware. The first and simple way to realise FIR filters is direct realization [17, 18], which exactly traverses the expression:

$$y = \sum_{k=0}^{N-1} h[k] \cdot x[n-k] \tag{4}$$

In above equation $x[n]$ is input; $y[n]$ is output and $h[n]$ is impulse response of the filter. In Eq. (4) there is no use of symmetric between coefficients. One way to realise filter is known as use of symmetric between coefficients which considerably realised number of multipliers. We know FIR filter has symmetric coefficients. It means we can express following expression [17]:

$$h[k] = h[n-k-1] \tag{5}$$

where $h[n-1]$ is the last coefficient. Hence by using symmetric between coefficients, number of multipliers decreases to 50%. However, amount of other hardware (delay and adder) does not decrease. To use less hardware, another method was introduced, multiplier free realization, [19–21] which does not use multipliers and had its own advantages. In addition, systolic realization method was introduced [22–25] which had its own definitions. Several methods with the purpose of realizing FIR filters by less hardware were introduced [26–28]. Researchers can sometimes modify the special characteristic of the FIR filter to achieve better output with less hardware. Some researches tried to increase the hardware efficiency of digital systems by means of bit-stream signal processing (BSSP). For example in [29], BSSP is applied on some digital systems which reduced the large number of logic gates and data lines in a hardware implementation because of simple arithmetic circuits and one routing line for one signal. In another study [30] the authors focused on FIR realising using hardware simplification by setting a coefficient to zero (HSSCZ) and they did not pay attention to frequency response of filter. In the current study there are no complicated mathematical equations. There is one more difference between our method and previously optimization algorithms. Also in this paper, we tried to improve frequency response of filter in addition to reducing the hardware complexity.

6 Proposed Algorithms

6.1 First Algorithm Hardware Simplification by Trial and Error (HSTAE)

In this case we set coefficients to zero to save a multiplier block and the adder of that step of the filter. To do this aim, we look at the amount of coefficients of the FIR filter and divide them into negative and positive groups. The next step was to move the positive and negative coefficients to zero. First, for each positive coefficient, we reduced the coefficients by 5% of its value and instead add 5% for negative coefficients. In the next step, we looked at the amount of SNR. We define an acceptable percent of SNR decreasing. If the SNR did not decrease more than an acceptable percent, in the next step we looked at the similarity between the output signal of our filter and the output signal of the equiripple filter. If their similarity was acceptable, we held the amount of that coefficient to zero. If the SNR value was not acceptable, we increased the percentages and continue until we reached the target range. By this way, we tried to save the quality of the output signal and realise the filter with less hardware. In addition to this purpose, we should check the frequency response of filter. Flowchart of this algorithm is shown in Fig. 4. If we want to introduce the steps, after set the absolute value of coefficient we divided the coefficients to positive and negative. For negative coefficients 5% of its initial value for each coefficient was added, and instead for positive coefficients, we reduced each factor by 5% to its original value. In this work, we set simultaneously coefficients to zero. After doing that we looked at the SNR of the new output signal (new SNR), if the SNR did not decrease or if the decrement of the SNR was acceptable (before doing algorithms, we should define the acceptable amount for our case as SNR (acp)), we will look at the similarity between the two signals (these two signals are the output signal of our filter and the output signal of the equiripple filter). If it was acceptable too, then we held the amount of that coefficient equal to zero. However, if the SNR decrement was more than the defined acceptable amount or if the similarity between the two signals was not satisfying, we had to change the percentage. First approach was increasing the percentage and again checks the new SNR. If SNR changed and if similarity between two signals were acceptable we did

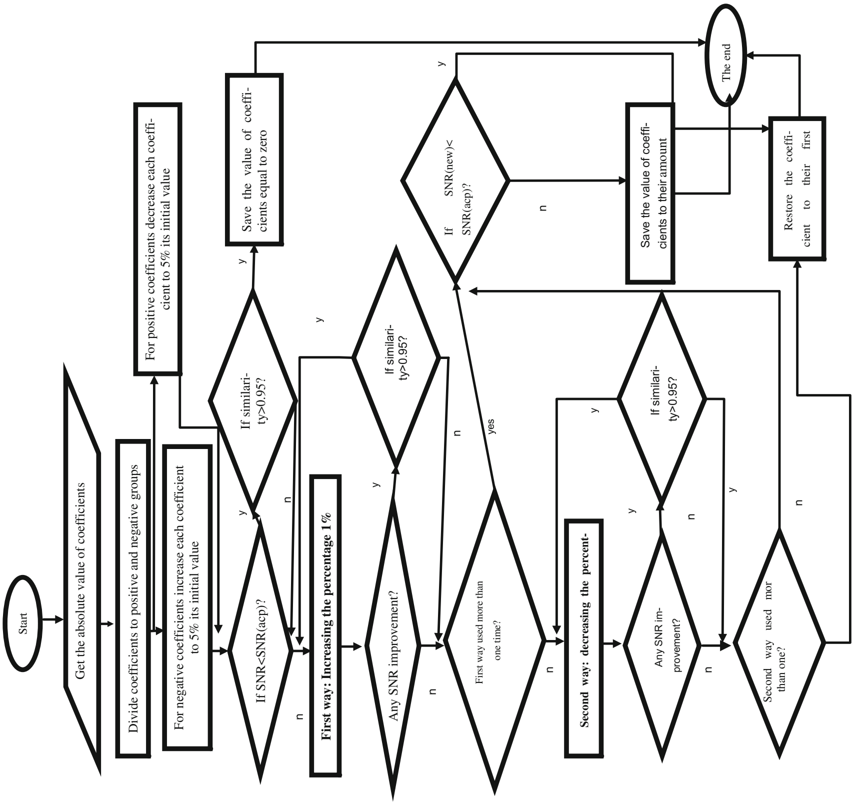

Start
Get the absolute value of coefficients
Divide coefficients to positive and negative groups
For negative coefficients increase each coefficient to 5% its initial value
For positive coefficients decrease each coefficient to 5% its initial value
If SNR<SNR(acp)?
y
n
If similarity>0.95?
y
n
Save the value of coefficients equal to zero
First way: Increasing the percentage 1 %
Any SNR improvement?
y
n
If similarity>0.95?
y
n
First way used more than one time?
yes
n
If SNR(new)< SNR(acp)?
y
n
Second way: decreasing the percent-
Any SNR improvement?
y
n
If similarity>0.95?
y
n
Second way used mor than one?
y
n
Save the value of coefficients to their amount
Restore the coefficient to their first
The end

Fig. 4 Flowchart of first proposed algorithm (HSTAE)

this way again. However, if the increase did not have a significant effect on SNR, then the second method is to reduce the percentage and again check the SNR and similarity between two signals. If none of these actions satisfy SNR amount or the similarity between signals, we restore the amount of all coefficients to their first amount. This means in that filter setting the amount of that coefficient to zero is not possible because it does not satisfy our demands.

In this algorithm, the similarity between two signals is measured by the cross correlation coefficient. If we have one dataset $\{x_1, \ldots, x_n\}$ containing *n* values and another dataset $\{y_1, \ldots, y_n\}$ containing *n* values, correlation is defined as:

$$correlation(\boldsymbol{r}) = \frac{\boldsymbol{n}\left(\sum \boldsymbol{x}_j\boldsymbol{y}_j\right) - \left(\sum \boldsymbol{x}_j\right)\left(\sum \boldsymbol{y}_j\right)}{\sqrt{\left[\boldsymbol{n}\sum \boldsymbol{x}_j^2 - \left(\boldsymbol{x}_j\right)^2\right] - \left[\sum \boldsymbol{y}_j^2 - \left(\sum \boldsymbol{y}_j\right)\right]^2}} \tag{6}$$

One distinction to emphasize is that in probability and statistics the definition of correlation always includes a standardization factor in such a way that correlations have values between −1 and +1. +1 for most similar signals and −1 for two inverse signals. In other word, as much as the amount of cross-correlation was close to +1, our signals are more similar. In this algorithm, if the coefficient of the last step sets to zero, the order of the filter decreases. It means alongside decrement in the number of multipliers and adders, we thrift the delay blocks.

6.2 Second Algorithm: Improved Modified PSO (MPSO)

As previously mentioned in the Sect. 4, we can improve PSO algorithm by changing the inertia weight. The objective of this paper was to design an adaptive inertia weight considering the stability we use Eq. (1) of optimization methods. In this case, we tried to independent the equation on iteration. To speed up the discrete environment we defined a new w:

$$\mathrm{w} = \mathrm{w}_{\min} + (\mathrm{w}_{\max} - \mathrm{w}_{\min}) \times \mathrm{Ps}\,(\mathrm{i}, \mathrm{t}) \tag{7}$$

where w_{min} is minimum amount of inertia weight and w_{max} is maximum value of inertia weight.

$$\mathrm{M_s}(\mathrm{i}, \mathrm{t}) = \begin{cases} 1 & if\,(gbest(t) - xi(t)) < pi(t) \\ 0 & if\,(gbest(t) - xi(t)) > pi \end{cases} \tag{8}$$

$$\mathrm{Ps}\,(\mathrm{i,t}) = \frac{1}{n}\sum_{i=0}^{n} \mathrm{Ms}\ (\mathrm{i}, \mathrm{t}) \tag{9}$$

In above equation gbest is the best position of coefficients and x_i is the initial position of each coefficient. pi(t) is the personal best of the ith particle. n is number of coefficients. None of the existing inertia weight adjustment strategies have paid attention to the stability of PSO and balance between gbest and initial position. As the inertia weight dynamically changes the searching ability and stability of conditions dynamically change. By satisfying the stability conditions in PSO, the speed and precision of the convergence are improved. In this algorithm, in contrast to Eq. (1), the relation does not depend on maximum iteration and the decisive factor for increasing the speed. In this case; it is depends on minimum distance between the best position of the particles gbest and pi(t). Since the PSO algorithm is a convergent algorithm, coefficients are expected to decrease with increasing the convergence speed. Flowchart of this algorithm is shown in Fig. 6. Stop condition, as previously mentioned, is SNR. We introduce steps. After setting the value of coefficients, we used Eq. (1) and initialized for each particle. Next step is updating gbest and pbest. As we said in MPSO algorithm, our stop condition is SNR and minimum distance between gbest and personal best (pi). Before doing algorithms, we had to define the acceptable amount for our case as SNR (acp). If SNR was acceptable we saved the amount of coefficient and stop. If SNR was not acceptable we should update the velocity and position. In this case, we applied our algorithm as new equations and speeding up the velocity (Fig. 5).

As it shown in Eq. (8) changing of the speed is depend on the best particle position. When the convergence rate increases particles get to their best position faster. In this step, we remove the filter coefficients that have the maximum distance with the best known position (gbest). We selected the coefficients with the least distance from the global optimal as the representative. The main purpose of using MPSO is to improve the frequency response of the filter and SNR. Since hardware complexity also needs to be reduced, based on the initial coefficients of the filter, we need to change the coefficients. One of the methods is changing the w. The other way is to choose some of best position of the particles as the representative which has minimum distance with gbest. Also we can use first algorithm to reduce the amount of coefficients.

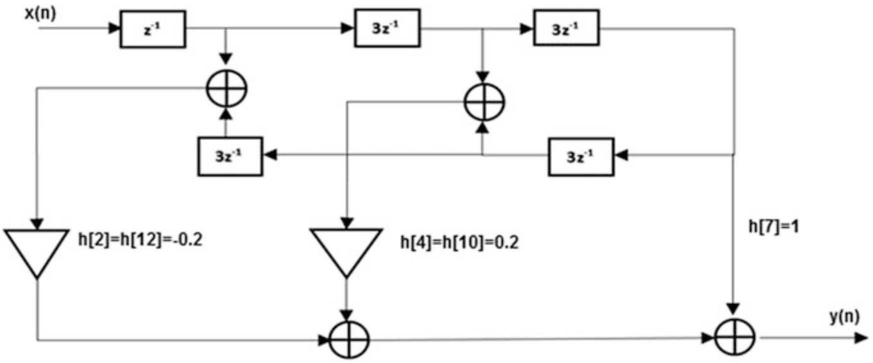


Fig. 5 Realised FIR filter using (HSTAE) algorithm

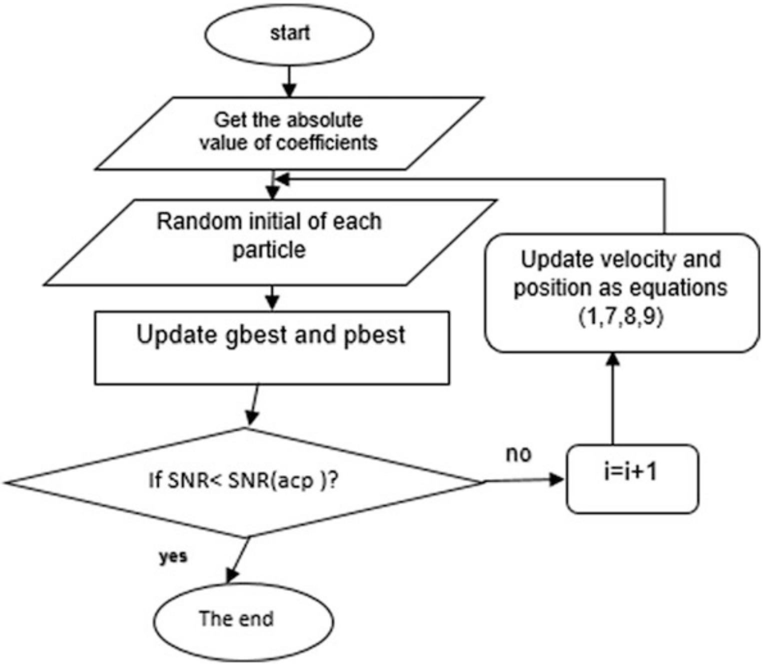

start
Get the absolute value of coefficients
Random initial of each particle
Update gbest and pbest
If SNR< SNR(acp)?
no
i=i+1
Update velocity and position as equations (1,7,8,9)
yes
The end

Fig. 6 Flowchart of second proposed algorithm (MPSO)

In this method, as in the first proposed method, we try to find the best known positions with respect to the stop condition and ignore small coefficients, (coefficients which have maximum distance with gbest). Besides the quality of the output signal, the filter output frequency response is also important. Hence we tried to investigate magnitude error, ripple band stop and its changes. The main reason for using the second algorithm (MPSO) was to improve frequency response of the filter in addition to improving the quality of the output signal.

7 Simulation Results

If we want to design an FIR filter using the equiripple method to eliminate the PLI from the ECG signal, our filter should have the following parameters:

Fpass1 = 41 Hz
Fstop1 = 58 Hz
Fstop2 = 60 Hz
Fpass2 = 85 Hz

Filter designed by above parameters has the following coefficients:

H[0] = −0.042170565559434589
H[1] = −0.11853363342094889
H[2] = −0.07411007446189434
H[3] = 0.060386911905517239
H[4] = 0.14442352449239027
H[5] = 0.079302062846230717
H[6] = −0.074037764845576615
H[7] = 1.0494790780874323
H[8] = −0.074037764845576615
H[9] = 0.079302062846230717
H[10] = 0.14442352449239027
H[11] = 0.060386911905517239
H[12] = −0.07411007446189434
H[13] = −0.11853363342094889
H[14] = −0.042170565559434589

With the above coefficient we need 14 delay blocks, eight multiplier and 14 adders to design equiripple filter. Through this FIR filter we can remove 60 Hz PLI noise from ECG signal. SNR of output signal in this filter is 8.36. By applying our proposed algorithms, we have two new coefficients. For the first algorithm (HSTAE) our coefficients are

H[0] = 0
H[1] = 0
H[2] = −0.2
H[3] = 0
H[4] = 0.2
H[5] = 0
H[6] = 0
H[7] = 1
H[8] = 0
H[9] = 0
H[10] = 0.2
H[11] = 0
H[12] = −0.2
H[13] = 0
H[14] = 0

By this coefficients filter realised as shown in Fig. 5. As shown in Fig. 5, the hardware complexity of the filter is considerably reduced. Comparison between both filter is given in Table 1. We increased the particle convergence rate to find the best known positions with respect to the stop condition and ignore the small coefficients. In the second algorithm (MPSO), using the proposed model the filter coefficients changes as following:

H[0] = 0
H[1] = −0.1
H[2] = 0
H[3] = 0.1
H[4] = 0.1
H[5] = 0.1
H[6] = 0
H[7] = 0.56
H[8] = 0
H[9] = 0.1
H[10] = 0.1
H[11] = 0.1
H[12] = 0
H[13] = −0.1
H[14] = 0

Block diagram of realised FIR filter using second proposed algorithm shown in Fig. 7 in which it is clear that delay blocks, adders and multipliers are significantly reduced by algorithms.

As we said we should consider similarity between equiripple filter and output filter. In this paper our c.correlation was 98.063% for HSTAE and 98.179% for MPSO. This coefficient shows that our filter is more similar to the equiripple filter and our cut frequency for both after using algorithms also was 60 Hz.

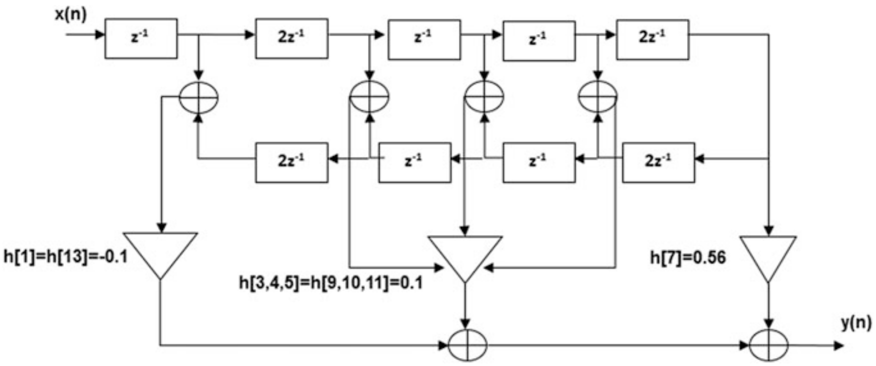


Fig. 7 MPSO algorithm to realise FIR filter

Table 1 Comparison of the number of different blocks and SNR of output signal (HSTAE)

	Designed filter by (equiripple) method	HSTAE	Alter (%)
No. delay blocks	14	13	−8
No. adders	14	4	−71
No. multipliers	8	2	−75
SNR	8.36	16.42	+68

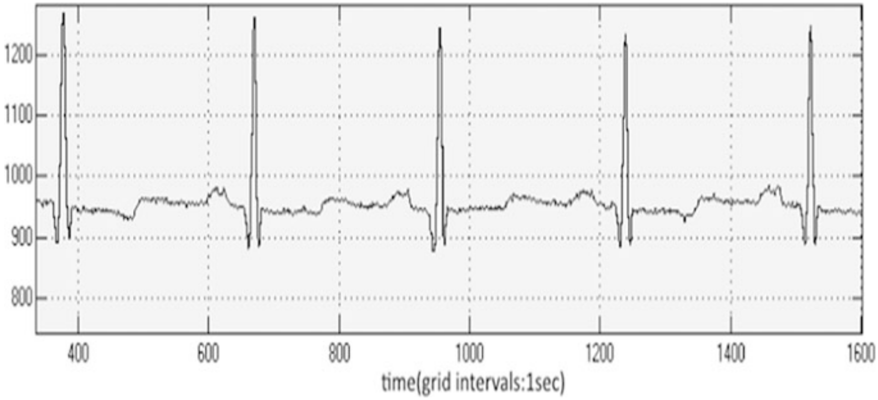


Fig. 8 Denoised ECG signal using HSTAE algorithm (amplitude (mv)/time (s))

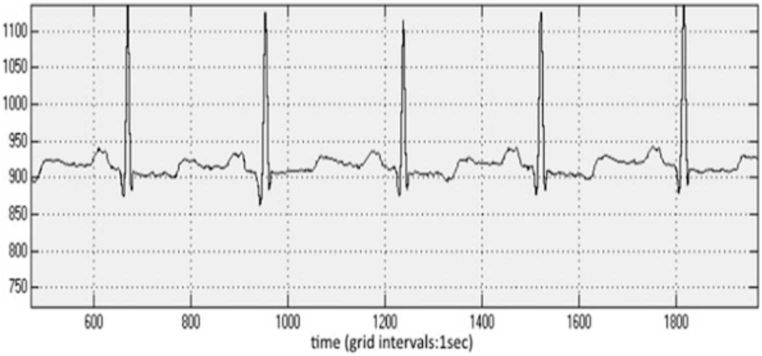


Fig. 9 Denoised ECG signal using MPSO algorithm (amplitude (mv)/time (s))

Table 2 Comparison of the number of different blocks and SNR of output signal (MPSO)

	Designed filter by (equiripple) method	MPSO	Alter (%)
No. delay blocks	14	13	−8
No. adders	14	6	−68
No. multipliers	8	3	−68
SNR	8.36	18.22	+74

As shown in Figs. 8 and 9, the filter which is realised by the proposed algorithms, the desired output signal has been achieved. ECG signal denoised and peak of the signal is easily detectable. Applying the proposed algorithms to a typical filter decreases its hardware. In Tables 1 and 2, we compare the number of different blocks of filters. As shown in Fig. 10, after using algorithms frequency response of

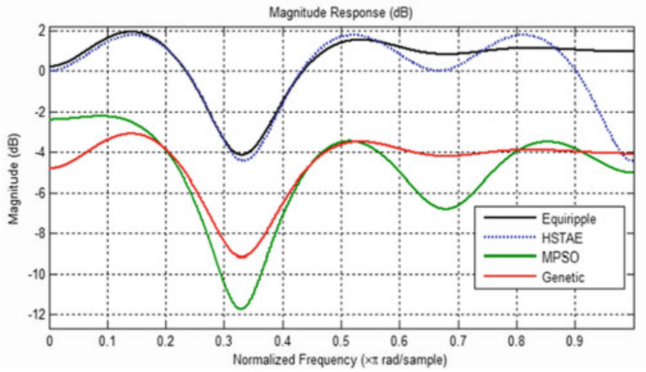


Fig. 10 Comparing of frequency response of proposed algorithms

Table 3 Comparison of magnitude error and band stop ripple (normalized frequency) order = 14

	Magnitude error (normalize)	Min. bandstop ripple	Max. bandstop ripple	Average bandstop ripple
PSO	1.356452	0.02966	0.01565	0.02087
HSTAE	1.788494	0.01378	0.01028	0.01245
MPSO	1.264297	0.014777	0.012561	0.013170
GA	1.6594201	0.01415	0.00645	0.010613

filter has been improved and the strength of signal in MPSO algorithm was better than the genetic and HSTAE algorithms.

As mentioned in Sect. 4; FIR filter designed by Remez algorithm results in so many ripples in stopband and requires more coefficients for sharp cutoff. To overcome this problem PSO algorithm was used to minimize error between the ideal and Remez filter. We used Eq. (3) to calculate the magnitude error. Moreover, to comparison between our work and previously work we compared our work with conventional PSO and genetic algorithm (GA). Results are provided in Table 3 in which MPSO has less magnitude error and band stop ripple. Simulation result shows that new algorithms have less hardware complexity and better frequency response. Also output signal has a better quality than equiripple filter that was our main goal of this study.

8 Conclusion

In this study we found that using the proposed algorithms helps to save hardware where the quality of the output signal does not decrease. In addition to this advantage, we consider frequency response of filter and we found that in the second algorithm, the signal attenuation has a perfect advantage to even the most famous genetic algorithms. The great achievement of our study was that using these two algorithms increases the amount of SNR. The number of delay blocks decreases by 8%, the number of multiplier blocks decreases by 75% and the number of adder blocks decreases by 71% for HSTAE. For MPSO, 8% for delay blocks, 68% for multipliers and 68% for adders. However, compared to the hardware volume, increasing of SNR for the HSTAE and MPSO was 68 and 74% respectively. In this work, in addition to reducing the hardware complexity, the filter found a better quality in terms of frequency response. Another important result in this article is providing a specific method for increasing the speed of the PSO algorithm, which requires an increase in speed over the number of iteration. Finally, by these algorithms, we achieved our targets. We saved the quality of the signal (in addition to quality of magnitude response) and decreased the hardware of the filter. Achieving the first target and saving the quality of the filtered signal are proof enough for the good performance of the realised filter.

References

1. ECG—Simplified. Aswini Kumar M.D., Life Hugger. Accessed on 11 Feb 2010
2. Thalkar S, Upasani D (2013) Various techniques for removal of power line interference from ECG signal. Int J Sci Eng Res 4(12):12–23
3. Shetty P, Bhat S (2014) Analysis of various filter configurations on noise reduction in ECG waveform. Int J Comput Commun Instrum Eng (IJCCIE) 1(1):88–91
4. Mishra S, Das D, Kumar R et al (2015) A power-line interference canceler based on sliding DFT phase locking scheme for ECG signals. IEEE Trans Instrum Meas 64(1):132–142
5. Shirbani F, Setarehdan SK ECG power line interference removal using combination of FFT and adaptive non-linear noise estimator, ©2013 IEEE. 978-1-4673-5634-3/13/$31.00
6. Butt M, Razzaq N, Sadiq I et al (2013) Power line interference removal from ECG signal using SSRLS algorithm. In: 2013 IEEE 9th international colloquium on signal processing and its applications, Kuala Lumpur, Malaysia, 8–10 Mar 2013
7. Biswas U, Maniruzzaman, Md (2014) Removing power line interference from ECG signal using adaptive filter and notch filter. In: International conference on electrical engineering and information and communication technology (ICEEICT)
8. Physionet. Retrieved from http://www.physionet.org/physiobank/database/mitd
9. Panda R, Pati UC (2012) Removal of artifacts from electrocardiogram using digital filter. In: 2012 IEEE students' conference on electrical, electronics and computer science
10. Sadiq I, Zuberi AM, Zaman I et al (2012) Adaptive removal of power-line interference from high resolution ECG. Adv Biosci Biotechnol 3:324–326
11. Chavan, MS, Agarwala RA, Uplane MD (2008) FIR equiripple digital filter for reduction of power line interference in the ECG signal. In: Proceedings of seventh WSEAS international conference on signal processing, robotics and automation (ISPRA'08), University of Cambridge, UK, 20–22 Feb 2008. ISSN: 1790-5117 1 4 7; ISBN: 978-960-6766-44-2
12. Kennedy J, Eberhart RC (1995) Particle swarm optimization. In: IEEE conference neural networks, Nov 1995, pp 1942–1948
13. Zhang Y, Wang S, Ji G (2015) A comprehensive survey on particle swarm optimization algorithm and its applications. Math Prob Eng 1:1–38

14. Kennedy J, Eberhart RC (1997) A discrete binary version of the particle swarm algorithm. In: IEEE international conference, Oct 1997, pp 4104-4108. 0-78034053-1/!l7/$10.OO 1997 IEEE
15. Taherkhani M, Safabakhsh R (2015) A novel stability-based adaptive inertia weight for particle swarm optimization, © 2015 Published by Elsevier B.V. 1568-4946
16. Yuanhai, X (2013) An improved particle swarm optimization algorithm for FIR filter design. Department of Electrical Engineering, King Fahd' University of Petroleum and Minerals, Dhahran, KSA, ©2013 IEEE. 31261 978-1-4799-2452-3/13/$31.00
17. Milivojevi Z (2009) Digital filter design. MikroElektronika
18. Parhi KK (1999) VLSI digital signal processing systems: design and implementation. Wiley, New York
19. Krishnan R, Vijayakumar S (2014) Multiplierless FIR filter design using global valued numbering and architecture. In: 2014 international conference on ICCCNT, Hefei, China, 11–13 July 2014
20. Jovanovic-Dolecek G, M-Alvarez M, Martinez M (2005) One simple method for the design of multiplier less FIR filters. J Appl Res Technol 3(2):125–138
21. Ye WB, Yu YJ (2012) Design of high order and wide coefficient wordlength multiplierless FIR filters with low hardware cost using genetic algorithm, ©2012 IEEE. 978-1-4673-0219-7/12/$31.00
22. Chiper DF (1999) A systolic array algorithm for an efficient unified memory-based implementation of the inverse discrete cosine transform. In: IEEE conference image process, Oct 1999, pp 764–768
23. Yin H, Du W, Hu YH et al (2012) A novel flexible foldable systolic architecture FIR filters generator, ©2012 IEEE. 978-1-4673-1295-0/12/$31.00
24. Uma R, Ponnian J (2012) Systolic FIR filter design with various parallel prefix adders in FPGA: performance analysis. In: 2012 international symposium on electronic system design, ©2012 IEEE. 978-0-7695-4902-6/12$26.00. https://doi.org/10.1109/ised.2012.45
25. Lee J-J, Song G-Y (2004) Implementation of a bit-level super-systolic FIR filter. In: 2004 IEEE Asia-Pacific conference on advanced system integrated circuits (AP-ASlC 2004), I 4–5 Aug 2004, IEEE. 0-7803-8637-X/04/$20.00 ~ 2004
26. Winod AP, Chang CH, Meher PK et al (2006) Low power FIR filter realization using minimal difference coefficients: Part I—Complexity analysis. In: Proceedings of IEEE conference APCCAS 2006, Singapore, pp 1549–1552
27. Ye WB, Yu YJ (2014) Bit-level multiplierless FIR filter optimization incorporating sparse filter technique. IEEE Trans Circ Syst I Reg Pap 61(11):3206–3215
28. Kyung-Saeng K, Lee K (2003) Low-power and area efficient FIR filter implementation suitable for multiple tape. IEEE Trans VLSI Syst 11(1):150–153
29. Fujisaka H, Kurata R, Sakamoto M et al (2002) Bitstream signal processing and its application to communication systems. IEEE Proc Circ Devices Syst 149:159–166
30. Meidani M, Mashoufi B (2016) Introducing new algorithms for realising an FIR filter with less hardware in order to eliminate power line interference from the ECG signal. IET Sig Process 10(7):709–716

Providing a Proper Solution to Solve Problems Related to Banking Operations Through the ATM Machines to Help the Disabled, the Elderly and the Illiterate People

Farhood Fathi Meresht

Abstract This article aims at solving the current problems of individuals with disabilities and illiteracy in using the ATM machines to carry out banking operations (e.g. paying and receiving payments or paying bills or any other operations that an ATM can do), so that these people, like other people, are able to do their banking operations without the involvement of another person.

Keywords ATM for the disabled · ATM equipped with sound
The elderly · The illiterate

1 Introduction

According to our research in this field (i.e. on ATM), we noticed that no team or person had done this before, and no similar work has been conducted. Therefore, we sought to come up with designing a device with regard to the problems faced by the disabled, the illiterate and the elderly people, so that these people do not experience much hardship when using the ATM machines that makes them ignore using them.

2 Speech Processing

2.1 Definitions and Applications [1, 2 and 3]

Speaker recognition is the process of automatically detecting the identity of the speaker based on the unique information available in the sound waves of his voice.

F. Fathi Meresht (✉)
Software Engineering, Lahijan Azad University, Lahijan, Iran
e-mail: farhood_fathi@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_69

This technique enables the identification of the speaking person and thus to control his/her access when using the services such as voice dialing, telephone banking, phone shopping, services of accessing to databases, information services, voice mail, security control for logging in confidential information domains and remote access to computers. In addition to the above-mentioned issues that generally deal with the computers and their users, this technology also has its own applications in legal matters.

2.2 Types of Speaker Recognition Systems [4] and [5]

In terms of application methods, speaker recognition systems are generally classified into two categories: speaker authentication systems and speaker identity recognition systems.

In a speaker authentication system, generally, by selecting or entering the name of one of the system's specific users, the individual claims that he is the same registered user of the system. In this case, the system is tasked to compare the audio characteristics of the claimant with the audio characteristics stored by the concerned registered user and, using the obtained result, accepts or rejects the claim of the individual.

In a speaker identity recognition system, the speaker does not claim the identity of a particular registered user, and this is the system that is tasked for recognizing it among the registered users of the system or to recognize that his/her audio features are not consistent with any of the registered users.

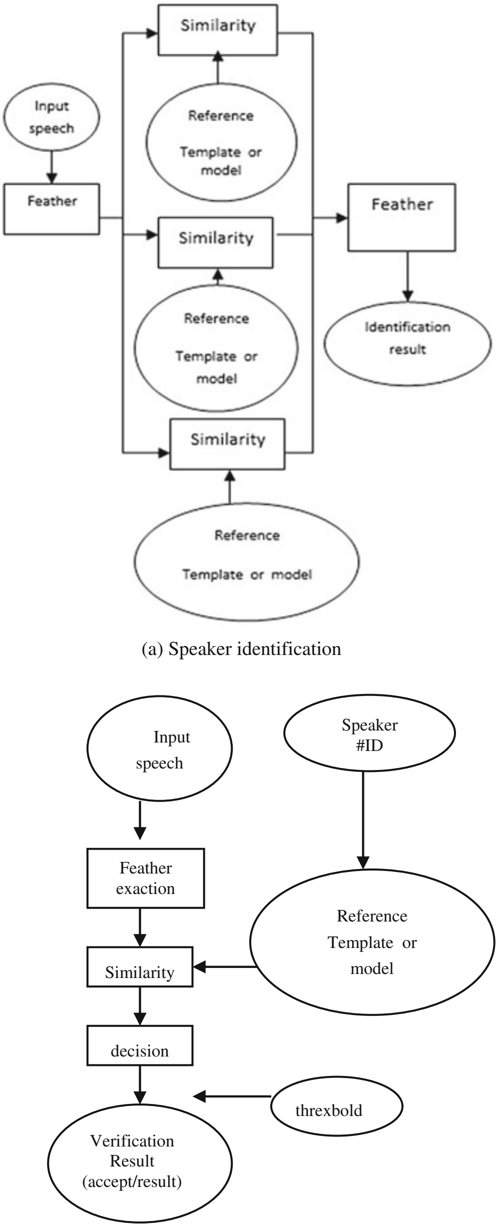
It seems that in the future, the use of the latter systems in multi-user large systems is greater than the applications of the former systems. Though, basically, these two systems are not much different.

Figure 1 depicts the basic structure of these two types of speaker detection systems.

From another perspective, speaker recognition systems are divided into two categories: text-dependent speaker recognition systems and text-independent speaker recognition systems. The first method requires that the speaker expresses the key words or certain sentences both in the learning phase and in the detection tests, while the latter is not dependent on a specific sentence or word.

Both methods have the same problem, that is, one can use the recorded voice of registered users to log in and easily deceive the system. There are some solutions to overcome this problem. For example, a small series of words such as digits can be used as keywords, and each time, the user is asked randomly to express a sequence of them. Even this method is not quite reliable, because it can be deceived by advanced electronic equipment with the ability to generate the sequences of phrases. Systems with the latter structure are known as text annotation (machine generated text) speaker detection systems.

Fig. 1 The basic structure of speaker identity recognition and identity verification systems


Input speech
Feather
Similarity
Reference Template or model
Similarity
Reference Template or model
Similarity
Reference Template or model
Feather
Identification result
Input speech
Speaker #ID
Feather exaction
Reference Template or model
Similarity
decision
threxbold
Verification Result (accept/result)

(a) Speaker identification

(b) Speaker verification

2.3 Implementation Techniques

In almost all identification systems, using a process called pattern recognition the similarity of each sample pair is scored. The use of this method requires the existence of a number of unique and comparable characteristics extracted from the selected feature as the input of the system.

Physical features of individuals such as the structure of the vocal organs, the size of the nasal cavity, and the characteristics of the vocal cords are unique, and can be extracted through signal processing algorithms as characteristic parameters or a set of extractable features. This fact is the basis of the speech recognition systems implementation methods.

The most important bottleneck of speaker detection systems (and consequently because of being of the same origin, the most important bottleneck of speech recognition systems) is how they work in places with conditions different than the laboratory conditions, one of the main features of which is the presence of noise in the system. Normalization methods are used to overcome this problem, which also are of different types and can be found in existing commercial systems.

Among the available algorithms for voice recognition, we can refer to hidden Markov model which is described below.

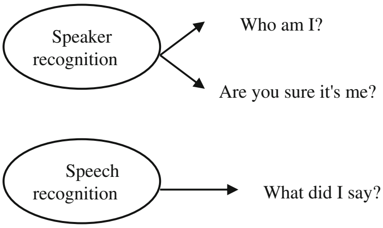
3 The Importance of Signal Modeling

In fact, computer speaking detection involves two basic types of cognitive actions: speech recognition and speaker detection. By analyzing an audio wave, it is possible to estimate the characteristics of vocal organs of the speaker, and these characteristics provide a solution to identify and authenticate him/her by biometrics. In contrast, speech recognition systems are trying to understand the concept of the mentioned sound wave. Most of the current studies on speech recognition technology are concerned with creating speaker independent systems with the ability to convert the speech of all speakers. While the objectives of the two types of systems look completely different, both are deeply fed by signal processing algorithms to extract features. In both fields, the efforts are continued to find a set of features that are sustainable in the face of environmental changes. This section will review the feature extraction algorithms used in both fields that includes a brief assessment of various algorithms for signal modeling with small diagnostic tests.

3.1 Understanding Signal Modeling

The speaker recognition systems are aimed at recognizing the features of the speech organs and the mode of speaking using the speaker's voice for the purpose of

Fig. 2 Different tasks



identification. The structure of the vocal organs, the size of nasal cavity, and the characteristics of the vocal cords can all be estimated using signal analysis. The speaker recognition is a general term used to refer to the identification and verification of the speaker. For detection, the estimated speaker's attributes are compared with the attributes existing in a database of registered users to find the closest comparable attributes. To confirm an identity, the speaker's claimed identity is accepted or rejected on the basis of his biometric signature (Fig. 2).

Speech recognition attempts to convert a spoken audio signal into words. Human beings express words by moving the vocal organs to a series of predictable places. If these sequences are extracted from the signal, the spoken words can be detected. Many speech recognition applications require speaker independent systems. These products can detect every speaker's speech.

Although these two goals look quite different, both of them apply pattern recognition to the spoken data. Some of the existing systems, such as the server 6 Nuance, apply both speech recognition and speaker authentication, simultaneously. Because of this similarity in approach, both of these technologies have the same weakness: severe performance degradation caused by the differences between training and testing environments. In summary, the efficiency of these technologies is heavily dependent on the environment in which they develop and therefore, the real world noisy conditions drive them to lower efficiency than optimal performance. The algorithms used for speech processing products are based on the audio model of vocal region and the ear canal. The next section clarifies the importance of extracting the features with an overview of pattern recognition and then, continues by describing the algorithms commonly used for the most commonly used products.

3.2 Pattern Detection

A pattern recognition system consists of two components: a feature extractor and a classifier. Ideally, when the data is transferred to the features data space, it is drawn to a class closest to itself, and rejected by different classes.

When the classifier is trained to distinguish between classes in this transferred space of features, a detection system only requires to transfer the input data through the same feature extraction system, and to determine in which class a new observation occurs. There are two important problems in applying this approach to speech processing. The first is that there is no requirement for comparability between the training environment and the testing environment. The use of a different microphone, background noise and transmission channels can degrade the performance dramatically (An essential criterion for judging a set of features, is its stability against such channel changes). The second problem is that there is a high overlapping between the classes in the feature space. Zhao presents some diagrams to illustrate this overlapping in two categories of spoken data collected through the telephone network. To overcome this overlapping problem, speech recognition engines use powerful statistical processing to unify the language model, which goes beyond the scope of this paper.

3.3 Signal Modeling Algorithms

The objective of signal modeling (often referred to as features extraction) is transferring the audio data to a space where the observations of one class are placed in a single group, and the observations of the different classes are separated from each other. These transfers are selected based on biological studies of vocal systems and human vocal organs. For example, vocal organs cannot be moved from one place to another place in less than about 5 ms. Therefore, practical systems can sample the spectrum 100 times per second, while the precision of operation is reduced by only a very small amount. Speech is a dynamic signal; therefore, we are interested in testing a small range. The continuity duration of a frame is defined as the length of time in which a set of parameters is valid. Although the frames do not overlap, we usually use an overlapping analysis window to consider more signal samples for each spectrum measurement. The direct application of spectral analysis to such a small amount of data is equivalent to applying a sharp rectangular window to the signal that causes spectral distortion. The frequency response of a rectangular pulse is a sinc function (sinc x = sin x/x) which has a curved passband and lot of rippling in the stopband. Different shapes for windows are achieved by applying a weight function. Hamming window with the equation w(n) = (a − (1 − a) cos (2p/[n − 1])/b

$$W(n) = \left(a - (1-a)\cos\left(\frac{2p}{[n-1]}\right)\right)$$

is a special example of the Hamming window with a = 0.54 (Pi number is 3.1415). The parameter b for normalization is chosen in such a way that the signal energy remains unchanged during the test. Hamming window shape provides a spatial analysis with a smoother passband and a stopsband without any significant distortion that both of these characteristics are important for obtaining the variable parametric estimates. Most today's systems use a frame size of 10 ms and a window size of 25 ms.

A feature extracted from signal is the absolute energy of the signal. Another category is the spectral measurement of the energy of certain frequencies. These measurements are similar to those of the initial conditions of human vocal system (hair cells in the cochlea are used to achieve the same objective). There are three ways to achieve these audio measurements: direct application of a digital filter bank to the time domain, using Fourier transform and linear predictive analysis. In terms of computing efficiency, both the latest methods are more common in today's systems.

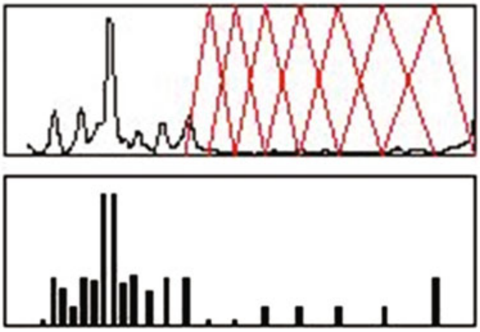
Because human hearing is not equally sensitive along a linear size, we project the spectrum to an understandable frequency size. Experiences about human perception have shown that frequencies with certain bandwidths and nominal frequencies known as critical bandwidths cannot be detected individually. Mel size is a simpler approximation that captures the observable pitch of a sound to a linear value. In 1940, Stevens and Folkman experimentally determined a mapping between the Mel size and the real frequencies. The difference in size is hardly below 1000 Hz, linearly and higher than 1000 Hz, logarithmically (Fig. 3).

The filter banks based on simple Fourier transform designed for the final features result in the desired frequency accuracy based on the MEL scale. To implement this filter bank, using the Fourier transform, the speech data window is transferred to the frequency domain. In the frequency domain, the coefficients of the range of each filter bank are found by applying a linear combination of the spectrum and the frequency response of the desired filter. In practice, triangular filter banks with overlapping are used where the central frequency of a filter is used as the endpoints of two adjacent filters. Therefore, the domain coefficients of each filter bank indicate the average value of the spectrum in the filter channel:

$$S_{avg}(F) = \frac{1}{N_s}\sum_{a=0}^{N} w_{FB}(\prod)_{S(F)} \tag{1}$$

Here is Eq. (1) where N(s) is the number of samples used to achieve the mean value and W(n) is the weighing function (similar to the triangular function

Fig. 3 Filter banks with triangular MEL space



described previously) and S(f) is the frequency response value calculated by the Fourier transform.

Linear predictive analysis is a tool for obtaining a smooth spectral coverage P(w) from an all-pole model of power spectrum. The predictor linear coefficients are directly correlated with the logarithmic region proportions, which are geometric parameters of defect tubular model for speech production. The filter bank domains are obtained by sampling the linear predictive spectral model at the appropriate bank frequencies. This can be done by direct evaluating of the LPC model, but in practice, Fourier transformation is applied to the predictor coefficients. Since the number of LPC coefficients is less than that of the audio samples, this method is computationally efficient. Just as the coefficients of the filter bank range were obtained from range obtained by the Fourier transform, they are obtained from a linear predictive range.

A homogeneous system has been used for speech processing, because it provides a method for separating the disturbance signal from the shape of the audio region. A space with this feature is a cepstrum obtained by inverse discrete Fourier transform of the energy logarithm. The cepstral coefficients are obtained by calculating the filter bank domains using the following equation:

$$c(n) = \tfrac{1}{N_s}\sum_{k=0}^{N_s} \log\left|S_{avg}(k)\right| e^{j\frac{2\prod}{N_s}\quad kn}, \qquad 0 \le n \le N_s - 1 \tag{2}$$

The Eq. (2) where S(avg) is the average value of the signal in the kth channel of the filter

In practice, discrete cosine transformation is used for its computational efficiency. The cepstral coefficients are often weighed to minimize the changes that do not lead to the creation of information that this process is referred to as liftering. It is interesting to note that in speech detection literature, the characteristics of the speaker are removed as non-data creating changes, but the speaker detection systems also use liftering. By the derivation of the basic features with respect to time, both speech recognition and speaker detection systems provide short-term local information. For example, a vowel can be detected by finding its commands in the spectrum, while a consonant sound is modeled by spectrum transitions. The values of first-order derivatives of characteristics are called delta coefficients, and their second-order derivatives are called acceleration or delta-delta coefficients. The time derivative is approximated using a regression relation that takes a frame set before and after the current frame.

Speaker detection systems also use a feature selection module in the pattern recognition framework. For speech recognition, all signals must be written to a text display, while the speaker detection system does not need to work under this obligation. Therefore, the feature selection module only stores the attributes of the vowel sounds. Vowel sounds directly satisfy the linear predictive modeling assumptions and are less affected by audio noise.

3.4 Digital Methods for Saving Sounds

In designing a digital audio system, there are two questions to be answered: (1) How much does the sound need to look good? (2) What is the tolerable data rate? The answer to these questions often leads to one of these three choices: First is music with high fidelity, where the sound quality is the most important thing, and almost any data rate is acceptable. The second is a phone call that requires looking as a natural talk and a low data rate to reduce the cost of the system. The third is a compressed speech where data rate reduction is very important and unnatural sound quality seems tolerable to some extent. This includes military communications, cell phones, and the speech stored digitally for voice mail audio or multimedia applications. Figure 5 shows the interactions in the selection of each of these three methods.

While music requires a bandwidth of 20 kHz, a talk that seems natural would only require a bandwidth of 3.2 kHz. However, although the bandwidth is limited to 16% of the initial value, but only 20% of the initial information is lost.

Telecommunications systems often use a sampling rate of about 8 kHz that allows the speech to be transmitted at a natural quality, but if it is used to transfer music, its quality will be lost to a high degree. You're probably familiar with the difference between these two rates. The FM radio stations broadcast with a bandwidth of about 20 kHz, while AM stations are limited to 3.2 kHz. Speech and the usual sounds on the second type of stations look natural, while the music is not so.

Table 1: Audio data rate versus sound quality. The sound quality of a digital audio signal depends on its data rate, which is equal to the product of its sampling rate by its number of bits per sample, divided into three parts: music with high fidelity (706 kbps), speaking with the phone talking quality (64 kbps) and compressed speech (4 Kbps).

Systems that deal only with sound (and not with music) can reduce the precision from 16 to 12 bits, without significant loss of precision. By selecting an unequal size for the step of size switching, this can be reduced to 8 bits per sample. A sampling rate of 8 kHz with an ADC precision of 8 bits per sample leads to a data rate of 64 kbps. This is an ultimate limit for speech to sound natural. Note that

Table 1 Difference between sounds quality

Sound quality required	Band with	Sampling rate (kHz)	Number of bits (bit)	Date rate (bits/s) (k)
High fidelity music (concept dick)	5 Hz–20 kHz	44.1	16	706
Telephone quality speech (with companding)	200 Hz–3.2 kHz 200 Hz–3.2 kHz	8 8	12 8	96 64
Speech encoded by liner predictive coded	200 Hz–3.2 kHz	8	12	4

speaking requires a data rate equivalent to 10% of music data rates with high-fidelity.

The data rate of 64 kbit/s represents the ultimate application of sampling theory and the choosing values for audio signals. Methods of reducing data rates are much further based on compressing the data flow by eliminating the inherent iterations of speech signals. One of the most effective methods available is LPC with various types and subcategories. Based on the required quality of the speech signal, this method can reduce the data rate 2–6 kbps.

4 Voice Recognition Algorithm

4.1 Hidden Markov's Model

Hidden Markov's model was first described in a series of statistical papers by Leonard E. Baum and other writers in the second half of the 1960s. One of the first applications of HMM was speech recognition, which began in the mid-1970s. In the second half of the 1980s, HMM was introduced into the field of analysis of biological sequences, specifically DNA. Since then, its application has expanded in bioinformatics.

A Markov model is a statistical model in which the modeled system is assumed to be a Markov process with unobserved (hidden) states. A hidden Markov model can be considered as the simplest dynamic Bayesian network.

In the normal Markov model, the state is directly visible by the observer and therefore, the state transition probabilities are the only parameters available. In a hidden Markov model, the state is not directly visible, but the output is visible depending on the state. Each mode has a probability distribution on the possible output symbols. Therefore, the sequence of symbols generated by a hidden Markov model provides some information about the sequence of states. Note that the 'hidden' attribute refers to the sequence of states that the model passes over, not to the model parameters; even if the model parameters are accurately defined, the model still remains 'hidden'.

Hidden Markov's models are mostly known for their use in pattern recognition such as recognition of voice and handwriting, recognition of gestures and movements, part of speech tagging, bioinformatics, etc.

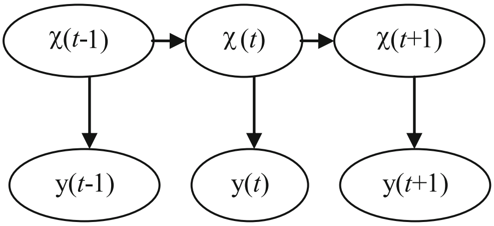
4.2 Architecture of Hidden Markov's Model

The following figure shows the overall architecture of an HMM sample. Each ellipse shape is a random variable that can assume any numeric value. The random variable x(t) is a hidden state at time t and the random variable y(t) is an observation at the time t. Arrows mean conditional dependencies. It is clear from the figure that for all the t times, the conditional probability distribution of the hidden variable x(t) gives a value for x that is dependent only on the value of the hidden variable x (t − 1). The values at the times t − 2 and earlier have no effect. This characteristic is called Marco. Similarly, the observed variable y(t) depends only on the value of the hidden variable x(t) (both at the particular time (t)). In standard mode, the hidden Markov model considered here is the state space of the hidden discrete variables, while observable variables can be discrete or continuous (from the Gaussian distribution).

In the hidden Markov model, there are two types of parameters: the probability of displacements (between states) and the outputs probability (or observations). The displacement probability controls the transferring mode from t − 1 to t.

The hidden state space contains N possible values for states. At time t, the hidden variable can assume any of these values. The transferring probability is the probability that, at the times t and t + 1, we are in states k (one of the possible states) and k1, respectively. Therefore, generally, there are N^2 possibilities of displacements (The sum of the probabilities of moving from one state to all other states is 1). The output probability determines the probability of occurrence of each member of the observational set for any hidden state possible, which can follow a probability distribution. The number of members of the observational set depends on the observable variable nature.

If the number of members of the observable variables is equal to M, then the total number of output probabilities is NM.



4.3 Discussion About Automatic Speech Detection

Speech can be examined from two aspects.

1. From the aspect of speech production
2. From the aspect of speech understanding and perception.

Hidden Markov Model (HMM) is an attempt to statistically model the speech production system, and therefore, it belongs to the first category of speech detection methods. Over the past few years, this method has been used as the most successful method for speech detection. The main reason for this is the fact that the HMM model is able to define the speech signal characteristics very well in an understandable mathematic format.

In an HMM-based ASR system, a feature extraction step is performed before HMM training. Therefore, the HMM input is a discrete sequence of vector parameters. Feature vectors can be trained in one of the two ways: introducing quantized vectors or continuous quantities to the HMM model. The HMM model can be designed to receive each of these types of inputs. The key issue is how the HMM model will adapt to the random nature of the feature vector values.

5 Presentation of the Proposed System

5.1 Presentation of Existing Problems

People with disabilities, the elderly and the illiterate have many problems with using the ATM machines that may make them disappointed, so that they may not use this device or may get help from someone else to do their banking operations. If such individuals want to do banking in a confidential manner, they cannot do so.

Among the people who are having trouble using an ATM are:

A. People using wheelchairs. Due to the long distance between the ground and ATM card slot, these people cannot enter their cards into the machine and inevitably, they cannot use the device. According to statistics about these people, they can raise their hands up to 150 cm from the ground level on average, while the card slot of these devices is located at the average height of 1.60 cm.
B. Blind people. People with this type of disability can never use an ATM.
C. The visually impaired people. These people may be able to access the ATMs, but they face a lot of difficulties for doing their banking tasks, e.g. light intensity, which, if it is a bit too high, they will be in trouble, or if they respond the device with delays due to their impaired vision, the device itself is interrupted.
D. Persons with disabilities in upper limb. Even if these people are able to reach the machine, they cannot do their own banking operations. These people may temporarily or permanently suffer from this physical disability, but in either case, they will not be able to perform their banking operations.

The people mentioned above were referred to as the disabled, but they are not the only community struggling with ATM problems. There are other individuals suffering the same problem, cited in the following.

(A) The illiterate people. It's better to first define illiteracy and then mention to their problems. Illiterate refers to someone who does not have the ability to write or read the alphabets. Although, some of these people also have little ability to read, but in general, all the people within this class cannot write, understand, express or use the numbers very well.

Therefore, this class of society that is significant in number (according to statistics, there is almost an illiterate person among each family), cannot generally do their own banking tasks on their own. Even when they refer to the banks, they must fill a special banking form to do anything; therefore, they are unable to do their own banking tasks.

(B) The elderly people. Because of their age, these people usually have hearing difficulties, vision problems, or muscle problems that suffer from long standing or long waiting, and for the same reasons, they do not refer to ATMs or banks.

Furthermore, disabled war veterans who suffer from physical disabilities cannot perform their personal tasks using ATMs.

5.2 Proposed Solution

By understanding the problems of disabled veterans, the elderly, the illiterate, and the disabled people who shall be honored by the whole community, and can serve as examples for the young people, some efforts must be made to achieve better perspectives that these people do not suffer hardships.

Here, the strategy and idea mentioned in this article are presented with a complete example:

For example, if one of the mentioned people wants to refer to a bank to perform his banking tasks, he/she will be encountered with many problems, such as filling out the forms, waiting in the queue, signing, etc. Therefore, it's better for him/her to use an ATM.

These people also have some problems in using an ATM, which we mentioned to. Instead of doing their banking tasks like ordinary people, they should interact with the bank clerk verbally, and the clerk in charge can fill the forms for him/her. If this attending clerk is supposed to be externally available, still, these people are having troubles. This operator is better to be outside the banks. If these assumptions were feasible, the problems would be greatly reduced, but unfortunately, none of the banks do such a thing.

There are some ATMs outside the banks that perform almost 80% of the client's tasks. However, due to the mentioned problems, the abovementioned people cannot communicate with these devices, verbally.

Therefore, if there is an ATM that can communicate with the clients verbally, there is no problem using ATMs for these people.

This is a good solution, but this type of ATM does not exist, however it can be designed and produced. If a microphone is embedded into the machine, we can practically have an ATM device with which we can communicate interactively, and this will solve the problems.

The type of microphone embedded in these devices is slightly different than the normal microphones. Their difference is in a small electronic piece that prevents unrelated sounds from entering into the system (i.e., a kind of audio noise filter).

For example, after installing this microphone, the disabled, the elderly or the illiterate people by referring to it and saying a word such as "Start" or saying a certain number, such as "One", inform the machine of the start of their operation. In the next step, the system will play a sentence like "Enter your card number". By telling the numbers on his/her card and saying a word like "Finished" to the system, the person announces that he/she has entered his/her card number. After checking the number and its validity, if it is valid, i.e. the number of expressed digits is right, the system prompts the customer to enter a password by playing a sentence like "Enter your password number". By telling his/her password (e.g. if the password is 123456, the individual first says "One", then "Two", etc. to the system) and saying a word like "Finished", the person tells the system that he has entered his password number. After authenticating the password, if the password is valid, the system announces to the client to choose his/her desired action.

The selection of banking tasks can be done in one of the following two ways:

(A) The customer declares to the verbal system about what he/she intends to do. For example, the customer says "Receive" and the system will perform the customer's intended action after the word is detected.
In this case, the voice detection issue is raised that using different algorithms (mentioned above), the system processes the spoken words and responds to the client's requests.

(B) Using a number, the system will tell the customer what it can do for him/her. For example, by playing a sentence such as "1—Receiving 2—Paying 3—Paying a bill 4—Charge, etc.", the system announces its services, and the customer declares the intended type of transaction by the announcement of the transaction associated number.

For example, if the customer wants to pay a bill, by telling "bill payment" in model A, or by saying "Three" in model B, he tells the system that he/she intends to pay a bill. In the next step, the system announces to the customer that "Enter the billing ID". The customer declares the numbers on the bill and says a word like "Finished" to the system showing that he/she has declared the number completely. In the next step, the system will again notify the customer that "Enter the payment ID". The customer declares the numbers on the bill and says the word "Finished" to the system, showing that he/she has declared the number completely. After authenticating the number, the system will prompt the customer about the bill cost and type of the bill, so that if the client confirms the accuracy of the bill, the machine can deduct the amount from the client's account. Finally, if the client does not request another task, by the announcement of words such as "Finished", he/she tells the system that he/she has no task to do and to close his/her account.

If the customer declares a wrong password or card number for three consecutive times, the system closes the client account.

5.3 Advantages of the Proposed System

This method does not require a physical card, but only a card number and a password known by the individual will suffice. The main advantages of this method are: (1) No need for physical card (2) Saving the budget spent on issuing cards (3) No need for the time when the person awaits for the card to be issued, which in most banks will take up to 3 weeks. (4) If the card is broken or physically damaged, the cardholder cannot use it, anymore. But this does not require a card to enter into the device and its associated interferences, and this problem is also solved.

5.4 The Drawbacks of the Proposed System

One of the questions that the reader might think of is the following: "How does the proposed system maintain the security?"

In answer, it should be noted that these types of ATMs should be placed within their own booths (The description of the cabin is as follows: first of all, this type of booth should have a framework for the ATM to be placed inside) (Photos 4 and 5).

It must be mentioned that these booths should only fit an individual.

There's a sensor inside the booth that if there are more than one person behind the system, it will tell the client that someone else besides him/her is at the back of the booth.

These booths also have a door that is closed when the customer enters in. The door can be opened and closed automatically or manually. After these people enter into the cabin, the door close to prevent the client's voice from spreading out of the cabin.

Fig. 4 The photo of the cabin in which the proposed ATM is supposed to be located

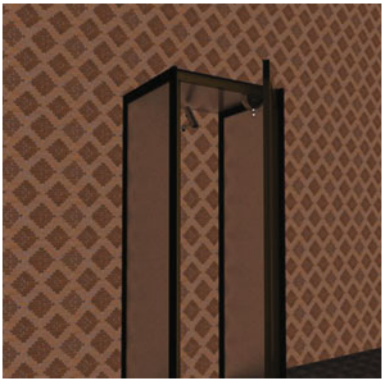


Fig. 5 The photo of the cabin in which the proposed ATM is supposed to be located

6 Conclusions and Recommendations

According to the cases mentioned in the article, it can be concluded that the production of ATMs communicating with the clients by sound via a microphone is necessary to solve the problems of the mentioned people.

The authors of this article are proposing the production of this type of device and, if possible, this could be achieved by updating the software and hardware of the current systems in the banks. Therefore, it is expected that in the near future, this project can be realized by the help and support of the relevant organizations to be carried out in our beloved country.

Acknowledgements I want to sincerely thank Mr. Hadi Roshan and Mr. Shahrooz Rajabzadeh Khanafi who helped me prepare and present this article.

References

1. Furui S NTT human interface laboratories. A text-independent speaker recognition method robust against utterance variations. IEEE international conference on acoustics, speech, and signal processing, Toronto, Ontario, Canada, 06 August 2002
2. Smith SW The scientist and engineer's guide to digital signal processing, from chapter 22: audio processing, 1997–1998
3. Hira ZM, Gillies DF A review of feature selection and feature extraction methods applied on microarray data. Adv Bioinform Vol 2015, 18 May 2015
4. Vogt T, Andre E Comparing feature sets for acted and spontaneous speech in view of automatic emotion recognition. IEEE International Conference on Multimedia and Expo, 24 October 2005
5. Esfandiari A (2002) Design and implementation of an audio passphrase control system

Presenting a New Clustering Algorithm by Combining Intelligent Bat and Chaotic Map Algorithms to Improve Energy Consumption in Wireless Sensor Network

Masome Asadi and Seyyed Majid Mazinani

Abstract One of the major challenges that wireless sensor networks face is the limited energy of nodes which reduces network's life time. Clustering is a popular approach to overcome this problem. Also, it is a particular energy efficient mechanism within large scale wireless sensor networks. Most problems of the computer systems like wireless sensor network could not be solved by linear solutions and there is not any deterministic solution for most NP-hard problems and the result of such problems is always optimizing. To solve these problems, applying evolutionary algorithms is recommended. The bat algorithm could find the shortest path between member nodes of the cluster and cluster head. In this paper, to reduce the energy consumption in wireless sensor nodes and also select the suitable cluster heads, the capabilities of combining the evolutionary bat algorithm and chaotic map is used. Applying chaotic map instead of some particular and random parameters in the bat algorithm improves the clustering. The results obtained from the implementation of the proposed method in MATLAB and their comparison with the existing methods such as GA, GAPSO, LEACH and LEACH-T represent significant impact in energy consumption improvement, network lifetime increase and also the number of live nodes increase within different rounds of algorithm execution.

Keywords Clustering · Bat algorithm · Chaotic map · Energy efficiency
Wireless sensor network

M. Asadi (✉)
Department of Computer, Khorasan Razavi Science and Research Branch,
Islamic Azad University, Neyshabur, Iran
e-mail: m.asadi@iau-neyshabur.ac.ir

S. M. Mazinani
Electrical Engineering Department, Imam Reza International University,
Mashhad, Iran

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_70

1 Introduction

The short lifetime of wireless sensor networks is considered as one of the most important challenges of these networks. Using clustering and routing protocols is one of the best solutions to overcome lifetime challenge [1]. Solving the energy problem in wireless sensor network also reduces costs, since for instance, instead of annual usage of new nodes; we could do this every two years! And this means that the costs have been halved. In this paper, an evolutionary algorithm is used to perform clustering in order to improve life time within wireless sensor network. Many evolutionary algorithms have been noticed with various applications in recent years. Some of these algorithms such as particle swarm, imperialist competition, shrimp and bat algorithms have appropriate results in optimizations. The proposed method applies the combination of evolutionary bat algorithm and chaotic map algorithm and also exploiting the collective intelligence of bat groups to select the cluster head and perform clustering [2–4]. Selecting the appropriate cluster head for generations to run algorithm is one of the most important issues in the proposed algorithm. In selecting the cluster head, the coverage preservation and inter-cluster distance are two major and effective factors. Appropriate selection of cluster head leads to reduction of energy consumption in each cluster and as a result the nodes' life time increases and finally it results in increasing the life time of the wireless sensor network. In this paper, comparisons such as the number of dead nodes, the network's remaining energy at different times of algorithm run and also network lifetime are performed.

2 Related Works

In this section, we study the literature inspired of evolutionary algorithms to optimize energy consumption in wireless sensor network. The studied algorithms include ant colony algorithm, genetic algorithm, particle swarm algorithm, hybrid genetic and particle swarm and LEACH-T algorithms.

In ant colony optimization [5], three types of routing optimization have been presented based on evolutionary ant algorithm. In these methods, the density of clusters results in increasing of network lifetime. Also, in another research on the application of ant algorithm [6], the optimal search is performed based on the ant colony algorithm and factors such as distance and remaining energy of nodes. The benefit of this method is the optimal cluster head selection among the nodes of a cluster which leads to network lifetime increase and load balancing in such networks. In [7], the back path selection is applied. In this paper, the ant colony algorithm is used for clustering. In this method, the ant colony algorithm is applied to find the optimal path among cluster heads. Here, the main idea is that the clusters closer to the base station are smaller than the other clusters. This method increase the sensor network's lifetime significantly. The old being of this algorithm is considered as its drawback.

The genetic algorithm has also been used to improve the network lifetime. In [8], the genetic algorithm is used to perform clustering. In this method, the network lifetime has significant increase in comparison with LEACH algorithm. In this algorithm, the crossover and mutation are used to generate better generations. In [9], the clustering is performed at each round of the algorithm by using the genetic algorithm with two fitness functions based on nodes' distance and energy. This method selects a cluster head with high energy and low distance to the base station. Also, in this algorithm, it is important to select elitism chromosomes for crossover in order to have smarter and better generations [10]. The genetic algorithm has been used to increase the lifetime of wireless sensor network. In this method, the focus is on selecting the next generation of the population. Here, the amount of coverage has considerable improve over previous methods. The high cost and computational overhead and also old being are the drawbacks of this algorithm.

The optimal particle swarm algorithm is inspired of the behaviors of birds and fish and it is an appropriate method to optimize energy consumption and suitable selection of cluster head in wireless sensor network. In [11], the criterion of cluster head selection is the remaining energy of each node. The aim of cluster head selection is to find the best path to the base station. The equality of node numbers at each cluster and selecting the best cluster head are the major goals of this method. One of the important advantages of this algorithm is to find the optimal roots of proper scheduling.

In LEACH-T algorithm [12], the clustering mechanism is used to reduce energy consumption. In this algorithm, nodes are deployed randomly in the environment and after clustering, a cluster head is selected for each cluster. In this method, the combination of TLBO[1] algorithm with LEACH algorithm has been used to perform clustering and cluster head selection. In [13], applying the optimization algorithm and TLBO learning in wireless sensor network include two training and learning phases. In the training phase, the aim is to promote the training, so that the result is improved significantly. In the learning phase, each member of the network must enhance its own learning capability and learn from other members. The computational overhead and random selections in this algorithm have led to reduction of final efficiency.

In GA algorithm, the features of genetic algorithm such as crossover and mutation are used to perform clustering and cluster head selection. In this algorithm, chromosomes include information and the best chromosomes are selected within the population by particular operations such as inheritance, mutation, and crossover and then the next generation is produced by using fitness function. PSO algorithm is inspired of the natural behavior of birds flying to find food. In this algorithm, there are numerous particles where the more transparent particle is searched by the fitness function and selected as the optimal. GAPSO algorithm is formed of the combination of genetic and particle algorithms. In [12], PSO has drawbacks such as falling into local optimum and high rate of information stream among particles

[1] Teaching Learning Based Optimization.

which makes similar particles and as a result missing diversity. Applying the characteristics of the genetic algorithm like crossover and mutation could be effective in reaching the convergence. In fact, the combination of these two algorithms increases convergence significantly.

The general bat algorithm presented in 2016 [14] has an objective function with parameter of total energy amount minus the energy used in nodes. In this method, the distance is an important factor in packet transmission which is computed by Euclidian formula. One of the benefits of this algorithm is to reduce the lost packets due to the appropriate selection of cluster heads. The random selection of effective parameters in making diversity in this algorithm is considered as a drawback.

3 The Proposed Algorithm

3.1 Bat Algorithm Introduction

The bat algorithm is inspired of the echolocation behavior of bats. This algorithm is a bio metaheuristic algorithm [15].

3.1.1 Studying Bat's Behavior

Most bats have the ability of echolocation. They use this ability to find prey, detect obstacles, the entrance and exit from the nest hole in the dark. Bats produce very high sound pulses and listen to its return from the surrounding. Bats also perform 3D image processing of its prey's motion, direction, speed and type by using reflection delay, reflection detection, reflected loudness, time difference between their two ears. Bats propagate different pulses according to the location, type and features of their prey. Each bat generates about 10–20 pulse per second which reaches 200 per second at the hunting time. Each pulse has 25–150 kHz sound frequency [16].

3.2 Studying the Evolutionary Bat Algorithm

The evolutionary bat algorithm is one of the most efficient evolutionary algorithms currently be used for optimization. In this algorithm, multi criteria functions are applied to find the global optimum. The echolocation capability of bats is an important factor to identify the target and prevent collisions with obstacles and also discover the path in the dark. To formulate the bat algorithm properly, the following assumptions are addressed in terms of bats' echolocation.

1. The distance of targets is always sent to bats by using positioning system. This case has the capability of detecting various targets even in the dark.

2. Bats fly randomly with V_i velocity, Fmin frequency within X_i state and wavelength with λ oscillation. Also, the loudness value to find the target or food is $A_0 - A_{min}$. Either the wavelength or frequency could be changed automatically based on the proximity of bats to the target as the pulse dispersion rate r ∈ [0, 1] changes. The amount of changes in the dispersion parameters is varied between the maximum A_0 and minimum A_{min}.

In the proposed bat algorithm, at first all the variables of echolocation system are initialized. In the standard form of the bat algorithm, the values of these frequencies to generate pulse and dispersion are produced randomly. By repeating this step, bats attempt to find the best solution in such a way that they try to obtain the best existing solution from their initial location per iteration. Each solution is updated automatically to find the best solution. At last, as the best solution is observed based on the given criteria, the process terminates.

Firstly, the initial solution is generated randomly for the bats population with chaotic function and later the new solution could be generated based on bats motion and by using the following functions.

$$f_i = f_{\min} + (f_{\max} - f_{\min})\beta \tag{3.1}$$

$$v_i^t = v_i^{t-1} + (x - x^*)f_i \tag{3.2}$$

$$x_i^t = x_i^{t-1} + v_i^t \tag{3.3}$$

where β is the random vector used to generate a uniform distribution in β ∈ [1, 0]. This parameter is determined by the chaotic theory to obtain better results. In the following, we present how to use this chaos instead of this parameter.

x* is the current global optimum which is obtained after comparing all solutions across bats.

f_i is a frequency generated uniformly from [f_{min}, f_{max}].

There is a need to random step with a direct development in order to generate the best solution which is generated by the following equation.

$$x_{new} = x_{old} + \varepsilon A^t \tag{3.4}$$

ε is a random number in [−1, 1] and r_i is the estimation rate.

For each bat, after finding the food its dispersion amount and pulse dispersion rate increases.

The mathematical formulation of dispersion and pulse is as the following:

$$A_i^{t+1} = \alpha A_i^t \tag{3.5}$$

$$r_i^{t+1} = r_i^0[1 - \exp(-\gamma t)] \tag{3.6}$$

$$A_i^t \rightarrow 0 \; and \; r_i^t \; \rightarrow {}^{0}_{i} \; as \; t \rightarrow \; \infty \tag{3.7}$$

where

$$\begin{aligned} \alpha : &\quad constant \quad 0<\alpha<1 \\ \gamma : &\quad constant \quad \gamma > 0 \end{aligned}$$

In the standard bat algorithm, the values of α and γ are constant. However, in the proposed algorithm, the chaos theory has been used to initialize these parameters which led to increase the algorithm's efficiency in optimizing energy consumption within these networks.

3.3 Advantages of Bat Algorithm

It can be said that the particle swarm optimization and harmony algorithm are the particular case of bat algorithm by considering simplification assumptions. If a random parameter is replaced instead of f_i and we have $r_i = 1$ and $A_i = 0$, then the bat algorithm is converted to the standard PSO algorithm. In bat algorithm, if v_i is not used and also the values of r_i and A_i are constant, the simple harmony search is obtained.

- The bat algorithm has the ability of fast convergence in the initial steps.
- This algorithm is an applicable for those problems try to find the solution quickly.
- It is suitable for large scale problems and has acceptable results.

3.4 Applications of the Bat Algorithm

Since the creation of the bat algorithm, it has been used almost in all optimization fields such as classification, scheduling, data mining, image processing, feature selection and etc.

Some of the algorithms presented upon the bat algorithm include the extended bat algorithm, the fuzzy logic of bat algorithm, multi-purpose bat algorithm, zero and one bat algorithm, improved bat algorithm, k-means bat algorithm and also chaotic bat algorithm.

3.5 *Chaotic Bat Algorithm*

3.5.1 Chaotic Theory

Using chaotic theory and replacement of some particular and random parameters in algorithms, such as bat algorithm lead to an improvement of the search results. The chaotic theory has been used instead of some particular algorithms in several algorithms such as particle swarm optimization [17], ant and bee colony [18], imperialist competition [19] and firefly [20].

The obtained results represent that using the chaotic map instead of particular and random parameters has improved the results. The reason why using the chaotic map instead of particular parameters in these algorithms has led to improve the conditions is not well known yet. However, the performed researches show that the chaotic map could have high level of integration and replacing a constant or random parameter with the chaotic map could make higher mobility and diversity in optimization and search. Therefore, it is expected that applying the combination of chaotic map with evolutionary bat algorithm could have acceptable results.

3.5.2 Introducing Chaotic Maps

Some of the parameters are initialized either with random functions or constant and uniform parameters in most evolutionary algorithms. Using chaotic theory leads to higher mobility and distribution in the obtained results. In the following, we will study some of the chaotic maps, briefly.

3.5.2.1 The Chebyshev chaotic map is defined as the following

$$x_{k+1} = \cos(k\ \cos^{-1}(x_k)) \tag{3.8}$$

3.5.2.2 The iterative chaotic map is defined as the following

$$x_{k+1} = \sin\left(\frac{\alpha\pi}{x_k}\right) \tag{3.9}$$

In this function, the suitable parameter for variable a is a ∈ (0, 1).

3.5.2.3 The Sinusoidal chaotic map is defined as the following

$$x_{k+1} = ax_k^2\ \sin(\pi x_k) \tag{3.10}$$

In this function also the appropriate values for the variables are a = 2.3 and x_0 = 0.7.

3.5.2.4 The Sin chaotic map is defined as the following

$$x_{k+1} = \frac{a}{4} \sin(\pi x_k) \tag{3.11}$$

where the proper value for variable a is $0 \leq a \leq 4$.

The Sin chaotic map is used in the proposed method. This mapping generates better results at the implementation and also it has more suitable format in MATLAB coding. The results obtained by using chaotic map are diverse and depend on the previous and primary runs and values. Therefore, the obtained results represent the average of several algorithm runs. In some parameters of the proposed algorithm, the chaotic map could be used instead of them. In the following, we would present the parameters replaced with chaotic map.

In the chaotic bat algorithm, the frequency equation could be modified and optimized by the chaotic function.

$$f_i = f_{\min} + (f_{\max} + f_{\min})\beta \tag{3.12}$$

$$f_i = f_{\min} + (f_{\max} - f_{\min})CM_i \tag{3.13}$$

In the standard bat algorithm, β is a random number between 0 and 1 (Eq. 3.14) and its value is selected by the chaotic function $\mathrm{CM_i}$ (Eq. 3.15) in the proposed algorithm.

The selection of parameter λ_i from the velocity equation in the bat algorithm could be modified by the chaotic function.

$$v_i^t = v_i^{t-1} + \left(x_i^t - x^*\right)f_i \tag{3.14}$$

$$v_i^t = v_i^{t-1} + \left(x_i^t - x^*\right)CM_i f_i \tag{3.15}$$

In the bat algorithm, λ equals to a value for normal frequency between 0 and 1 (Eq. 3.14) and it could be modified by the chaotic function $\mathrm{CM_i f_i}$ (Eq. 3.14).

The pseudo code of chaotic bat algorithm:

1. Start
2. First phase: initialization
3. Set the count of the generation to 1: t = 1
4. Initialize the bat NP value randomly and consider each bat as a potential solution.
5. Set the following cases.

$\mathrm{A_i}$: Acoustic loudness
$\mathrm{Q_i}$: Frequency pulse value (by using chaotic map)
$\mathrm{V_{(i=1,2,\ldots,NP)}}$: initial velocity (by using chaotic map)
$\mathrm{r_i}$: Pulse rate

6. Second phase: iteration loop
7. Continue until the loop end condition such as "determined round number", "network remaining energy", "the number of live nodes" and "the number of given generations" is obtained.
8. Generate new solutions and more optimal clusters and cluster heads at each loop run (updating clusters).
9. Select one of the solutions as the best solution, if the generated chaotic number is more than the pulse rate.
10. Generate a new random solution.
11. Consider a new solution, if the generated chaotic number is less than the acoustic loudness and the new solution is better than the global optimum.
12. Increase the pulse rate value and decrease the acoustic loudness.
13. Select the best node within each cluster.
14. Loop end
15. **Third phase:** represent the simulation results.

4 Simulation and Results of the Proposed Algorithm

The simulation of the proposed method is performed by using 4 types of different implementation in wireless sensor network. These implementations are done in MATLAB.

In Table 1, parameters such as network number, network area and the number of nodes are presented and in Table 2, different locations considered for site are represented.

In this section, we discuss on simulation environment. As above mentioned, in the proposed method, the bat optimization algorithm and chaotic map algorithm are used to optimize the energy consumption in wireless sensor network. The used parameters in the performed implementations could be seen in Table 4.

Table 1 Parameters of the wireless sensor network

The number of nodes	Simulation environment size	Number
100	100 * 100 m^2	A
500	100 * 100 m^2	B
1000	100 * 100 m^2	C

Table 2 Sink location in simulation

50 * 50	Middle	A
0 * 100	Corner	B
50 * 150	Outside	C

The implementation has been done in environments with different sizes and this led to more accurate evaluation of the proposed method. The numerical analysis and efficiency evaluation have been used to compare the proposed method with the standard bat algorithm and knowledge boundary methods.

4.1 Simulation

The proposed method has been simulated by MATLAB with 100 m × 100 m size. 100 nodes have been placed in this environment. In Table 3, parameters such as network area, the number of nodes and also the applied parameters for energy radio model have been represented. These parameters are standards obtained based on usual technical and mathematical tools like sensor and amplifier within the wireless sensor network. In Table 4, the parameters of the chaotic bat algorithm have been shown. These parameters have been addressed based on the best obtained efficiency. The location of the base station has been considered in all types of implementations in the simulation environment.

Table 3 Radio model parameters

Value	Parameter
100 * 100	Simulation environment size
100	The number of nodes
0.5 J/node	Initial energy (E_0)
0.0001 J	Minimum energy
6400 bits	The packet size sent from the cluster to the base station
200 bits	The packet size from node to cluster head
50 nJ/bit	Electronic transmitter (Eelec)
50 nJ/bit	Electronic receiver (Eelec)
50 nJ/bit	The amount of energy used to collect data

Table 4 Parameters of the bat algorithm

Value	Parameter
In terms of the number of nodes	Population size
100	Generation number
0.5	Acoustic loudness
0	Minimum frequency
1	Maximum frequency

4.2 Performance Analysis

The main idea of the proposed method is based on nodes' energy amount and physical distance of inner clusters among nodes and their cluster heads. In the proposed method, the nodes with higher energy are selected as cluster head. The clustering is addressed by using the bat algorithm in terms of fitness function based on internal distance of clusters.

The implementation and studying the performance of the proposed algorithm represent high optimization in wireless sensor networks. The performance of the proposed method is observable and analyzable numerically in the following table and figures. The presented method is compared with GA, GAPSO, LEACH-T, LEACH and bat algorithms. The results of this comparison and also numerical analytics of criteria such as the number of dead nodes, total energy used in wireless sensor network and network lifetime represent high and optimal efficiency of the proposed method.

In Table 5, the numerical analysis of nodes' lifetime is presented in wireless sensor network. This table shows the number of rounds for death of the first node, death of 50% nodes and death of the last node in the network. The top round for death of the first node in wireless network represents that the proposed method has high performance. This happens due to usage of chaotic map, since using chaotic map leads to higher diversity in cluster head selection and this fact results in load balancing at energy consumption reduction and consequently avoid early death of some nodes and nodes' energy reduction would be happened uniformly. According to the performed implementations in wireless sensor network and the obtained results, the first death occurs at rounds 158, 226, 573, 340 and 565 in LEACH, GA, GAPSO, LEACH-T and Bat algorithms, respectively, while this occurs at round 627 in the proposed algorithm. For 50% of the existing nodes, this happens at rounds 541, 960, 889, 540, 1180 and 1208 for LEACH, GA, GAPSO, LEACH-T, Bat and the proposed method, respectively. As it can be seen, the proposed algorithm has more improvement at higher rounds and this represents that this algorithm would give better results at high rounds of network development. The death of the last node or maximum network lifetime equals to 1047, 1726, 1901, 1627, 1890, 1913 number of rounds in LEACH, GA, GAPSO, LEACH-T, Bat and the proposed method. This fact represents that the network lifetime has increased significantly in the proposed algorithm in comparison with the other algorithms. In Table 5, a detailed comparison of mean death time of the first node, half of the nodes and all nodes in the proposed algorithm with the competitive methods is presented. One of the most important reasons of increasing network lifetime is to remove random parameters and increase the intelligence of the evolutionary algorithm. The appropriate selection of cluster heads and balance in energy consumption in all of the nodes has increased the performance of the proposed algorithm and consequently lifetime of the wireless sensor network (Figs. 1, 2 and 3).

To generalize the results, a number of constant rounds are considered to study the collective number of dead nodes in terms of rounds. This affects the number of

Table 5 Comparison of mean death time of the first node, half of the nodes and all of the nodes in the proposed algorithm and LEACH, GA, GAPSO, LEACH-T and Bat algorithms in 100 × 100 m^2 simulation environment with different sink locations

Chaos BAT	Bat	LEACH-T	GAPSO	GA	LEACH	Algorithm
Sink at the middle of the simulation environment						
627	565	340	673	246	158	The death round number of first node
1208	1180	540	889	960	541	The death round number of half of the nodes
1913	1670	1627	1901	1726	1047	The death round number of all of the nodes
Sink at the corner of the simulation environment						
581	521	321	625	223	147	The death round number of first node
1195	1023	530	871	953	533	The death round number of half of the nodes
1890	1540	1589	1854	1678	972	The death round number of all of the nodes
Sink at outside of the simulation environment						
572	530	312	632	213	144	The death round number of first node
1150	1085	521	863	932	521	The death round number of half of the nodes
1840	1590	1532	1832	1630	932	The death round number of all of the nodes

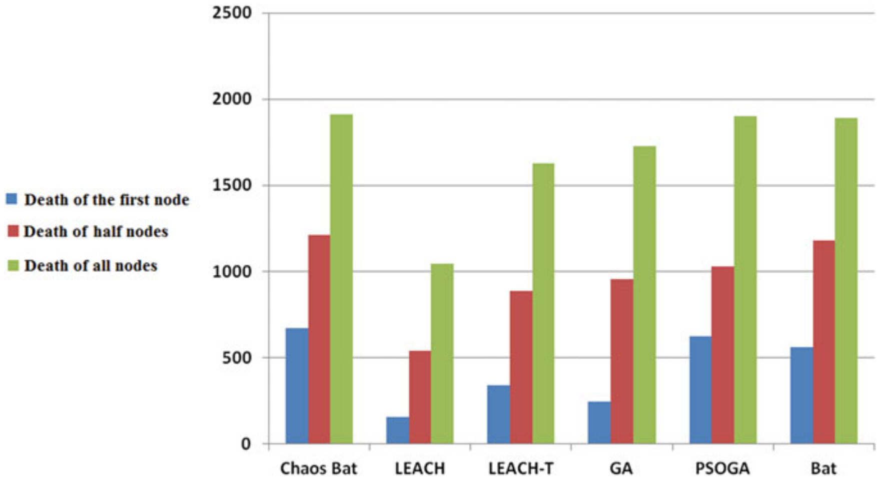


Fig. 1 The mean lifetime of wireless sensor network's nodes (100 × 100 simulation environment with 100 sensor nodes) and sink at the middle of the simulation environment

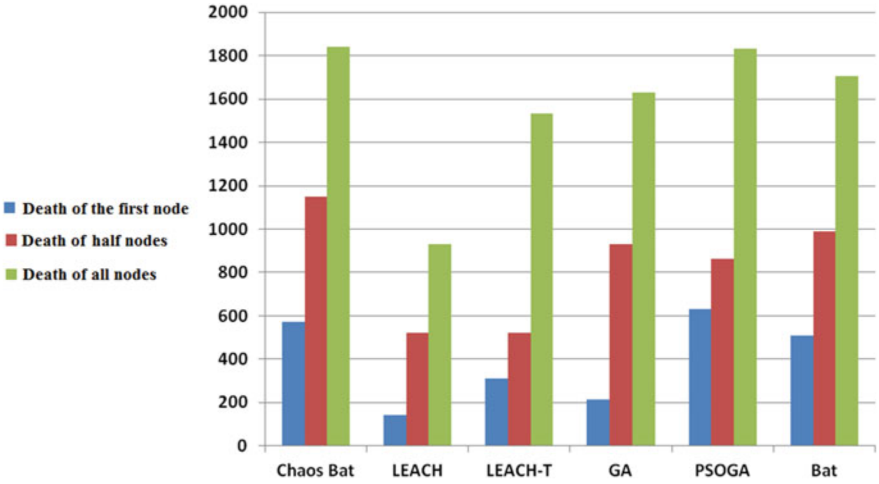


Fig. 2 The mean lifetime of wireless sensor network's nodes (100 × 100 simulation environment with 500 sensor nodes) and sink at the corner of the simulation environment

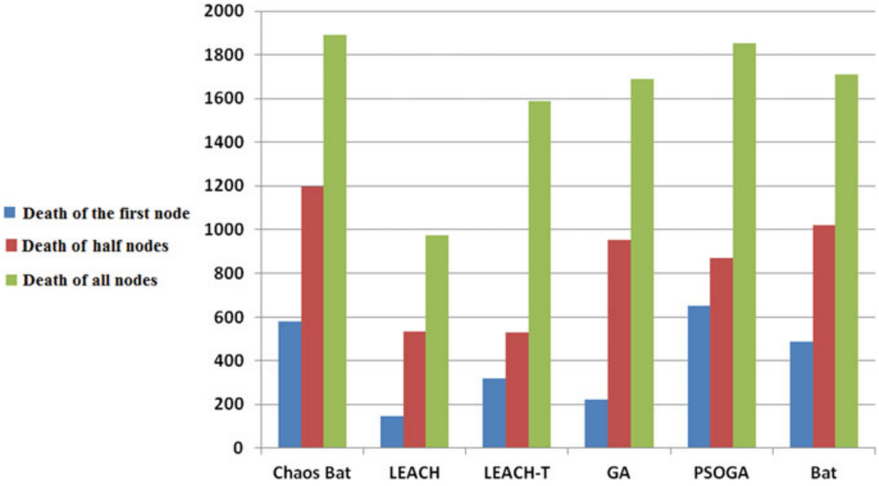


Fig. 3 The mean lifetime of wireless sensor network's nodes (100 × 100 simulation environment with 1000 sensor nodes) and sink at outside of the simulation environment

live nodes and therefore, influences network lifetime. Round 500 is selected for our tests. Table 6 represents the number of dead nodes at round 500.

To evaluate the performance, the performance criteria are also considered.

The number of dead nodes represent the lifetime of wireless sensor networks and improvement of used energy in these networks.

Table 6 Mean number of dead nodes at round 500 of running algorithms in different simulation environments

Mean number of dead nodes at round 1500	Mean number of dead nodes at round 1000	Mean number of dead nodes at round 500	Algorithm
100	100	40	LEACH
95	80	15	LEACH-T
91	82	8	GA
85	63	0	PSOGA
88	50	2	Bat
80	55	0	Choas BAT

Death of the first node: this factor shows the time interval for wireless sensor network implementations until reaching the dead node. This time interval represents the coverage amount in wireless sensor networks.

Death of 50% of nodes: this factor shows the time interval for the death of 50% of nodes.

Death of all nodes: this factor shows the time interval during which the last node dies. It also represents the wireless sensor network's lifetime.

Total energies of nodes versus rounds: this factor shows the amount of consumed energy in wireless sensor network's lifetime. The existing figures represent the performance factors studied for 4 implementations.

4.3 The Number of Dead Nodes per Round

The simulated results represent the lifetime of dead nodes in terms of rounds number. Results show that the proposed method gives promising results. The death of the first node occurs later over other methods. The death of 50% of nodes represents the amount of consumed initial energy. The death of all nodes computes the network lifetime. Figure 4 shows a comparative graph for the number of death versus the number of rounds in terms of wireless sensor network lifetime for the proposed method and other methods. For different implementations, the proposed method show high performance and stability. This method could be applied in most environments with the vast range of different implementations with high stability and scalability.

4.4 Total Energy of Nodes per Round

The comparative analysis on the proposed method represents that this method increases the energy stability and network lifetime. The proposed method could

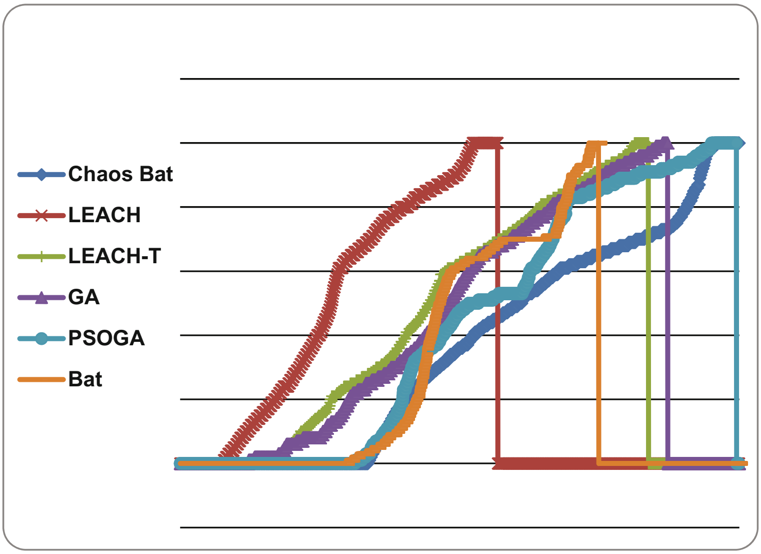


Fig. 4 The mean number of dead nodes in GA, GAPSO, LEACH, LEACH-T, Bat and the proposed algorithm for wireless sensor network implementation with 100 × 100 m^2 size and 100, 500 and 1000 sensor nodes

obtain the global optimum in network lifetime in terms of optimal selection of cluster head nodes which improves the general energy consumption (Fig. 5).

Based on the obtained results, we could conclude that the proposed algorithm is more suitable over the presented methods, since the network lifetime and information transmission amount have been increased. Also, due to proper load distribution among network nodes, the amount of energy consumption has also been reduced. As a result, it can be said that the proposed algorithm could increase the lifetime of wireless sensor networks.

5 Conclusion

In this paper, in addition to introduce the wireless sensor networks, we have studied the structure of these networks' nodes and also the addressed challenges in this field. Then, we have investigated some of the algorithms which have selected the appropriate cluster heads by different methods to increase network lifetime and finally we have presented our proposed algorithm. The presented algorithm performs clustering based on combining the intelligent evolutionary algorithm and chaotic map in the nodes of wireless sensor network and improves the algorithm's performance by replacing some of the particular and random parameters with chaotic map. Simulations have been done by MATLAB. Diagrams, numerical

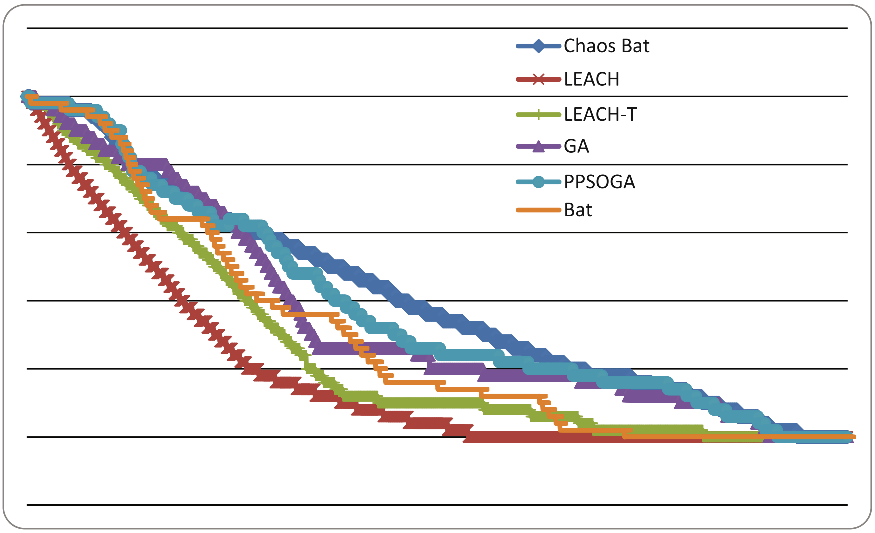


Fig. 5 The remaining energy of nodes in GA, GAPSO, LEACH, LEACH-T, Bat and the proposed algorithm for wireless sensor network implementation with 100 × 100 m^2 size and 100, 500 and 1000 sensor nodes

analytics and performance criteria show that the proposed method has more optimal energy consumption over the previous methods. As a result, this algorithm overcomes the existing energy problems in hierarchical clustering. Experimental results represent that the proposed algorithm could improve network lifetime, the number of live nodes at different rounds of algorithm run and the remaining energy in network. As a future work, the packet compression algorithms could be used to send information to the cluster head, simultaneously and the fitness function could be optimized in an evolutionary manner.

References

1. Liu JL, Ravishankar CV Member, IEEE (2011) LEACH-GA: genetic algorithm-based energy-efficient adaptive clustering protocol for wireless sensor networks. Int J Mach Learn Comput 1(1)
2. Gandomi, AH, Yang X-S, Chaotic bat algorithm. J Compu Sci
3. Talatahari S, Farahmand Azar B, Sheikholeslami R, Gandomi AH (2011) Imperialist competitive algorithm combined with chaos for global optimization. Commun Nonlinear Sci Numer Simul
4. Adnan Md, Akhtaruzzaman MA, Razzaque IA, Isnin IF (2013) Bio-mimic optimization strategies in wireless sensor networks: a survey. Sensors 14(1):299–345
5. Okafor FO, Fagbohunmi, GS (2013) Energy efficient routing in wireless sensor networks based on ant colony optimization. West Afr J Ind Acad Res 8(1):102–109

6. Li Y, Wang J, Qu Y, Wang M, Qiu H (2013) A new energy-efficient transmission scheme based ant colony algorithm for wireless sensor networks. In: 2013 8th international ICST conference on IEEE communications and networking in China (CHINACOM), pp 473–478
7. Du J, Wang L (2011) Uneven clustering routing algorithm for wireless sensor networks based on ant colony optimization. In: Proceedings of the 3rd IEEE international conference on computer research and development, pp 67–71
8. Liu J-L, Ravishankar CV (2011) LEACHGA: Genetic algorithm-based energy-efficient adaptive clustering protocol for wireless sensor networks. Int J Mach Learn Comput 1(1):79–85
9. Delavar AG, Baradaran AA (2013) CRCWSN: Presenting a routing algorithm by using re-clustering to reduce energy consumption in WSN. Int J Comput Commun 8:61–69
10. Singh VK, Sharma V (2014) Elitist genetic algorithm based energy balanced routing strategy to prolong lifetime of wireless sensor networks. Chin J Eng
11. Tillett J, Rao R, Sahin F (2002) Cluster-head identification in ad hoc sensor networks using particle swarm optimization. In: Proceedings of the IEEE international conference on personal wireless communications, pp 201–205
12. Yadav A, Kumar S (2017) A teaching learning based optimization algorithm for cluster head selection in wireless sensor networks. Int J Future Gen Commun Netw 10(1):111–122. ijfgcn.2017
13. Rao RV, Savsani VJ, Vakharia DP (2011) Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. Comput Aided Des 43 (3):303–315
14. Kavita M, Dr. Kashyap RC (2010) Improved bat algorithm based clustering in WSN. Int J Eng Dev Res 4(4). ISSN 2321-9939
15. Yang X-S (2013) A new metaheuristic bat-inspired algorithm. Stud Comput Intell 284:65–74
16. Yang X-S, Algorithm Bat (2013) Literature review and applications. Int J Bio-Inspired Comput 5(3):141–149
17. Gandomi AH, Yun GJ, Yang XS, Talatahari S (2010) Chaos-enhanced accelerated particle swarm algorithm. Commun Nonlinear Sci Numer Simul 18(2):327–340
18. Alatas B (2012) Chaotic bee colony algorithms for global numerical optimization. Expert Syst Appl 37:5682–5687
19. Talatahari S, Sheikholeslami R, Farahmand Azar B, Gandomi AH (2013) Imperialist competitive algorithm combined with chaos for global optimization. Commun Nonlinear Sci Numer Simul 17(3):1312–1319
20. Gandomi AH, Yang XS, Talatahari S, Alavi AH (2000) Firefly algorithm with chaos. Commun Nonlinear Sci Numer Simul 18(1):89–98

The Impact of Spatial Resolution on Reconstruction of Simple Pattern Through Multi Layer Perceptron Artificial Neural Network

Pardis Jafari and Saeideh Sarmadi

Abstract This study investigates the effect of spatial resolution factor on prediction of symbol ‘I’, the first Roman number, from scatter pixel images of this symbol through multiple perceptron neural network with adjusted network training parameter. The corrupted images of this symbol with different spatial resolution and pixel size has applied to the feed forward neural network model. The results of modelling have revealed acceptable correlation coefficient values with low and high resolution desired images and low and high resolution predicted images. From other aspect, a very less difference between these correlation coefficients emphasize on insignificant consequence of pixel size variation over the symbol with simple geometry.

Keywords Artificial feed forward neural network · Multilayer perceptron
Pattern recognition · Spatial resolution

1 Introduction

Identification of a simple/complex character, one potential branch of pattern recognition, still found as a challenging issue in image processing technology [1]. Establishing intelligent and fast mechanism with an acceptable accuracy is the main goal of any pattern recognition method [2]. A general approach for recognition of a pattern by human brain is extracting any related feature from content of an image [3]. These features helps human brain for pattern recognition by analyzing huge amount of data obtained by eye to identify for instance size and color of stationary and mobility object details only within approximately 100–200 ms [4]. Observing a certain feature within an image greatly rely on spatial pixel resolution of that image

P. Jafari · S. Sarmadi (✉)
Department of Computer Engineering, Persian Gulf University, Boushehr, Iran
e-mail: s.sarmadi@pgu.ac.ir

P. Jafari
e-mail: pardisabh@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_71

[5]. In the other hand the capability to recognize the smallest object within an image can be obtained at higher spatial resolution [6]. Artificial neural network with different training algorithms known as a very strong tool helps for recognizing a pattern or a character [7]. The similar facts between neuron learning process on recognizing a pattern in the human brain and artificial neural network make it worthy to open up a review on the analogous to artificial and real neuron performance.

In human brain, regardless the nature of information, all data initially transform to form electrical signals and triggering the neurons for constructing a network and preparing the brain for learning an experience [8]. This experience develops by adopting the nervous system to the surrounding environment and further participation of neurons as information-processing units in the human brain [9]. Over a series of knowledge acquired by the network and deeper learning process, the synaptic weights (i.e. interneuron connection strength) will update through the time [10]. The inherent fact behind any physical mechanism which finally leads to generation of input signals applicable for neuronal system is its nonlinearity [11]. One of the usual paradigm to treat the nonlinear learning either in bio-neuron or artificial neuron systems is mapping the output and input with the use of supervisor for modification of the synaptic weights [12]. This method requires two different sets of data. The first set consists of training examples and known as input signals to the neural network and the second set is labeled corresponding to the desired response [13]. The network in non-bio-systems artificially starts with picking up an example randomly from the first set and altering the network's weights for further shrinking the difference between target response and input signals to the network upon on appropriate statistical criterion [14]. The training of the network will move on continuously and weights are modified repeatedly up until no further significant changes appears in the weights [15]. Adaptivity of neural network is important when the weights need to be updated in response to the changes in the surrounding environment [16]. Despite of minor or major changes in the operating environmental condition, the critical fact is real updating the neuron's weights over a statistical environment. Neural network architecture is adaptable corresponding to variety of applications including pattern classification, signal processing, and control applications. If the network intelligently adaptive to any type of these processing, and weights updating operates for error minimization, then the system is reliable to be used over a dynamic environment. It must be noticed, sometimes, the adaptivity of the network may not comply with desire values needed for the assigning to the weights and under this circumstance instead of minimization of network's output error, it starts to maximizing the errors or randomly fluctuates to an unknown trend [17]. The good example of robust errors may find in an adaptive system with short-time constants or overhasty variation. Thus, the system tends to respond to spurious disturbances and finally lead to an enormous deterioration of the performance [18]. Following the above matters, this study aims to evaluate the artificial neural network performance over a dynamic pixel variation of low and high spatial resolution of unique simple pattern which here is 'I' the first Roman number. The objective task here is seeking for analyzing the performance of the multi-layer perceptron (MLP) feed forward neural network when assigned inputs to the network comprises different pixel number and pixel aspect ratio. The next sections in this stream will offer a short briefing about artificial neural network design, network parameters and a comparison between achieved modelling results.

2 Multi Layer Perceptron Network

The layout presented in Fig. 1 is the schematic of multilayer perceptron model with only one hidden layer. The capability of this model for nonlinear prediction and classification task makes it prominent for variety of applications including image processing [19]. The transformed images labeled with image 1, image 2 and so on are the inputs to the first hidden layer neuron of the MLP model. The transformed images are modified in the form of numerical vectors suitable to be applied into the hidden neurons.

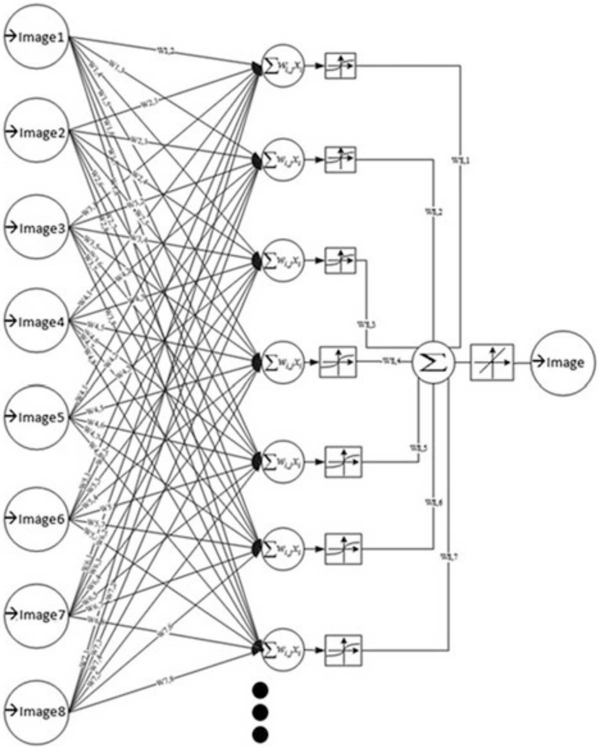


Fig. 1 Schematic of MLP model used to predict Roman number I in Roman numeric system from corrupted images with different pixel numbers and pixel aspect ratio

The number of hidden neuron (HN) is intentionally kept at only one layer in Fig. 1 to represent a simple model and minimize the features of the model. The real implementation in this stream have been conducted with several hidden neuron layers. The first set of arrows coming out from input layer and pointing to middle layer are called input-hidden neuron weights. The nodes in middle layer consists of many summation nodes responsible for treating the incoming weights and data from input layer viz., the symbolic sign is $\sum w_{i,j}x_i$. The output at each node then sends to neuron activation (transfer) function which here, as can be seen in Fig. 1, is selected as Tangen-Sigmoid transfer function at hidden neuron layers. This activation function offers a better compatibility between transformed input images and transformed output images [20, 21]. MLP derives the prediction potency from nonlinear processing in the hidden neurons. The cruciality imposed by Tangen-Sigmoid transfer function at two-stage (i.e. middle layer output and output layer) is evident from Fig. 1. The output of each node at hidden neuron layer can be written as:

$$ML_{node} = \frac{1}{1 + e^{-\left(\sum w_{i,j}x_i + b_j\right)}} \tag{1}$$

where, $w_{i,j}$ is the weight between input and the hidden neuron layers. x_i is the transformed images at input layer and b_j is the bias which assigned to each neuron at middle layer (*ML*). The final output at third (last) layer goes under another nonlinear transformation, and MPL model result which is a transformed image generally computes by [19, 22]:

$$OL_{node} = \frac{1}{1 + e^{-\left(\sum W_{i,j} \times \frac{1}{1 + e^{-\left(\sum w_{i,j}x_i + b_j\right)}} + B_j\right)}} \tag{2}$$

The capital $W_{i,j}$ and B_j in Eq. 2 convey the same concept similar to $w_{i,j}$ and $b_{j,}$ but this time with the emphasize on only one output node at last layer. Among several training methods applicable to use on the input data, here the Levenberg-Marquard (LM) algorithm which approved as a general algorithm for training the input data within the literature, has selected to use on MPL model [23–25]. The outstanding features of this algorithm are fast performance, time preserving with acceptable accuracy, as well as fast error convergence to global minimum error [26, 27]. It has been investigated by Sheela and Deepa et al. (2013) that no rule or rules were found in the literature on how to choose the number of hidden neurons [28]. It has been also found out by Islam and Murase et al. (2001) that the minimal errors are obtained by the increase of number of hidden neurons [29]. The investigations by other researchers only suggested a method rather than a confirmed solution for defining the value of hidden neurons in designing a neural network [28]. Among many approaches suggested by other

Table 1 Some well-known approaches to estimate HNs in MPL model

No.	Proposed by	Equation to estimate HNs values
1	Li (1995)	$\left(\sqrt{1+8n}-1\right)/2$
2	Tamura and Tateishi (1997)	$n-1$
3	Zhang (2003)	$2^n/n+1$
4	Hunter (2012)	2^n-1

researchers for computing HNs, the following methods has used within the structure of the MPL model in this study.

In Table 1, n is the number of inputs applied to the MPL model. The outcomes of investigation on HNs is out of this study's scope, and since the initial results after running the model proved that the less HNs affords a better efficiency by the means of reconstruction of desired image only method No. 2 in Table 1 suggested by Tamura and Tateishi et al. (1997) has selected to use in the model and further analytical treatment's. Next section is presenting the importance of achieved results from the proposed MPL model shown in Fig. 1.

3 Result and Discussion

The image illustrated in Fig. 2 is the desired image of Roman number one and the target image for MPL model in Fig. 1. This symbol with red color has aspect ratio of 1:7 (1 units wide and 7 units high) within an image ratio of 5:7. The number of pixel is assumed to be small for shortening the simulation time. This target image (Fig. 2) used as a part of MPL model assessment for two types of high and low resolution corrupted images applied to the model.

The following images in Fig. 3 are presenting low resolution of corrupted images of 'I' with the similar aspect ratio of image in Fig. 2. Whereas, the images in in Fig. 4 are the ones with aspect ratio of 15:17. The greater number of pixels within the image certainly increases the processing time of MPL model. The more pixels of an image mean the more matrix elements as the input to the MPL model and more complexity of weighting of the neurons in HNs layer. Further comparison between the images in Figs. 3 and 4 indicates the pattern similarity of each image with their counterparts in the other set. Thus, the only difference between two sets of images in Figs. 3 and 4 can be interpreted as different grade of resolution. In the other word, the main assessment of MPL model in this study is about reconstructing a low-resolution image of 'I' with the corrupted images of different resolution grades. Each image first transformed from matrix form to the vector form which is an acceptable format of input data enter to the MPL model. The five corrupted images provide five vectors as the inputs to the model. Hence, the number of HNs as it discussed in previous section has chosen as four. There are several parametric trainings also adjusted in addition of training type where here is LM algorithm.

Fig. 2 Image of desired Roman number 1 or ‘I’

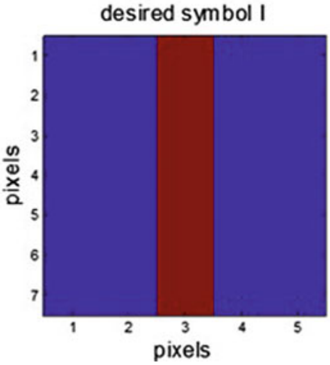
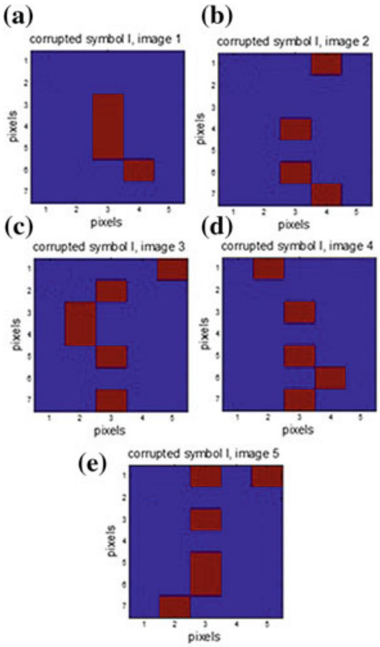


Fig. 3 Corrupted images of Roman number 1 or ‘I’ with low resolution



Some parts of these adjustments as can be seen from Fig. 5a, b are considered similar for low and high resolution input images. Among these adjustments, the most important ones, number of iteration and performance goal are set to 100 and 1e-12, respectively. The verification parameter ‘R’ or correlation coefficient in Fig. 5a, b has reached to a very acceptable value. The similarity between predicted images (Fig. 5c, d) by MPL model and desired image shown as Fig. 2. The literal coefficients 0.76 and 0.83 values with positive sign (Fig. 5c, d) and numerical values close to 1, emphasizing that the prediction fit line (blue line in Fig. 5c, d)

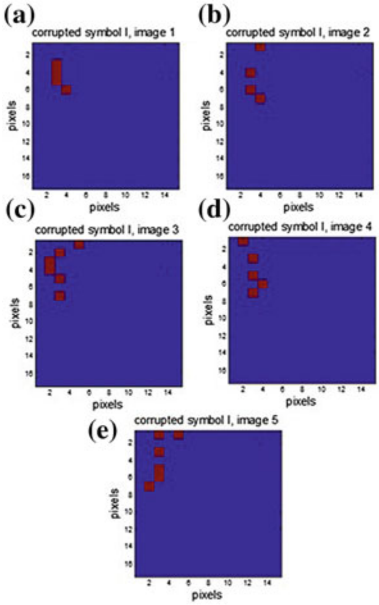


Fig. 4 Corrupted images of Roman number 1 or ‘I’ with high resolution

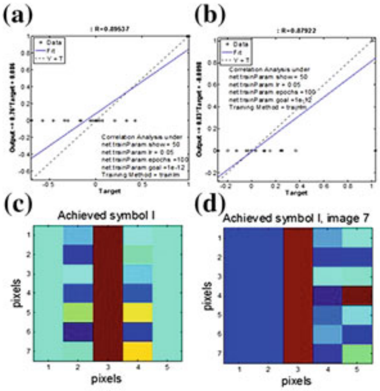


Fig. 5 Correlation analysis and prediction result of ‘I’ Roman number with high resolution (**a**, **c**) and low resolution (**b**, **d**) input images

slope at each low/high resolution case is ratherly close to the best fit line (dashed line in Fig. 5c, d).

Further analysis on verification of reconstructed images in Fig. 5 using mean square error (MSE) factor provides the value of 0.0445 and 0.0448 for high and low resolution inputs, correspondingly. Overall, the results are showing the good reconstruction of Roman number ‘I’, but still the neighboring pixels of symbol ‘I’ are showing several corruption and error propagation within the images, and further modifications on network structure and training parameters are needed for improving the network prediction.

4 Conclusion

This study reviews the effect of high and low spatial resolution corrupted images representing symbol ‘I’ as the inputs of MPL neural network on reconstruction and prediction of flawless image of first Roman number ‘I’. Response of designed MPL model after several time running the model has shown very insignificant role of spatial resolution impact on recognizing a certain simple pattern for instance ‘I’ as a target image.

References

1. Fink GA (2014) Markov models for pattern recognition: from theory to applications. Springer, Redmond
2. Saba T, Rehman A (2013) Effects of artificially intelligent tools on pattern recognition. Int J Mach Learn Cybernet 4(2):155–162
3. Nixon MS, Aguado AS (2012) Feature extraction and image processing for computer vision. Academic Press, Cambridge
4. Haykin SS, Haykin SS, Haykin SS, Haykin SS (2009) Neural networks and learning machines, vol 3. Pearson, Upper Saddle River, NJ, USA
5. Keren D, Peleg S, Brada R (1988) Image sequence enhancement using sub-pixel displacements. In: Proceedings CVPR’88, computer society conference on computer vision and pattern recognition, IEEE, pp 742–746
6. Pawley JB (2006) Fundamental limits in confocal microscopy. In: Handbook of biological confocal microscopy. Springer, US, pp 20–42
7. Ripley BD (2007) Pattern recognition and neural networks. Cambridge University Press, Cambridge
8. Oja Erkki (1982) Simplified neuron model as a principal component analyzer. J Math Biol 15 (3):267–273
9. De Ruyter van Steveninck RR, Zaagman WH, Mastebroek HAK (1986) Adaptation of transient responses of a movement-sensitive neuron in the visual system of the blowfly Calliphora erythrocephala. Biol Cybern 54(4):223–236
10. Sporns O, Tononi G, Kötter R (2005) The human connectome: a structural description of the human brain. PLoS Comput Biol 1(4):e42
11. Chen L, Narendra KS (2001) Nonlinear adaptive control using neural networks and multiple models. Automatica 37(8):1245–1255
12. Bengio Y, Lamblin P, Popovici D, Larochelle H (2007) Greedy layer-wise training of deep networks. In: Advances in neural information processing systems, pp 153–160
13. Hagan MT, Menhaj MB (1994) Training feedforward networks with the Marquardt algorithm. IEEE Trans Neural Netw 5(6):989–993
14. Xu L (1993) Least mean square error reconstruction principle for self-organizing neural-nets. Neural networks 6(5):627–648
15. Ghaboussi J, Pecknold DA, Zhang M, Haj-Ali RM (1998) Autoprogressive training of neural network constitutive models. Int J Numer Meth Eng 42(1):105–126
16. Geman S, Bienenstock E, Doursat R (1992) Neural networks and the bias/variance dilemma. Neural Comput 4(1):1–58
17. Grossberg S (1988) Nonlinear neural networks: principles, mechanisms, and architectures. Neural Netw 1(1):17–61
18. Grossberg S (ed) (1988) Neural networks and natural intelligence. The MIT Press, Cambridge

19. Samarasinghe S (2016) Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition. CRC Press, New York
20. Karatzas KD, Papadourakis G, Kyriakidis I (2008) Understanding and forecasting atmospheric quality parameters with the aid of ANNs. In: IEEE international joint conference on neural networks, 2008. IJCNN 2008 (IEEE world congress on computational intelligence), IEEE, pp 2580–2587
21. Panda SS, Mahapatra SS (2010) Online multi-response assessment using Taguchi and artificial neural network. Int J Manuf Res 5(3):305–326
22. Diaconescu E (2008) Prediction of chaotic time series with NARX recurrent dynamic neural networks. In: Proceedings of the 9th WSEAS international conference automation information. World Scientific and Engineering Academy and Society, Bucharest, Romania, pp 248–253
23. Chelani AB (2010) Prediction of daily maximum ground ozone concentration using support vector machine. Environ Monit Assess 162(1):169–176
24. Piotrowski AP, Napiorkowski JJ (2011) Optimizing neural networks for river flow forecasting—Evolutionary computation methods versus the Levenberg–Marquardt approach. J Hydrol 407(1):12–27
25. Jiang Q, Zhang ZG, Ding GL, Silver B, Zhang L, Meng H, Mei L et al (2006) MRI detects white matter reorganization after neural progenitor cell treatment of stroke. Neuroimage 32 (3):1080–1089
26. Demuth H, Beale M (1992) Neural network toolbox. For use with MATLAB. The MathWorks Inc.
27. Pires JCM, Martins FG (2011) Correction methods for statistical models in tropospheric ozone forecasting. Atmos Environ 45(14):2413–2417
28. Sheela KG, Deepa, SN (2013) Neural network based hybrid computing model for wind speed prediction. Neurocomputing 122:425–429
29. Monirul IM, Murase K (2001) A new algorithm to design compact two-hidden-layer artificial neural networks. Neural Netw 14(9):1265–1278

Analysis of the Role of Cadastre in Empowerment of Informal Settlements (Case Study: Ahvaz City)

Seyed Sajjad Ghoreyshi Madineh, Ramatullah Farhoudi and Hasan Roosta

Abstract Ahvaz is the capital of the province and the second largest city in terms of size after Tehran and the 7th Iranian city in terms of population. In Ahwaz, the issue of marginalization is a function of the process of marginalization of the whole country. The purpose of this study is to analyze the role of cadastre in the empowerment of informal settlements in Ahvaz city. This paper addresses the socio-economic dimensions of eight marginal neighborhoods and the effect of cadastre on empowerment of its inhabitants. Among the various types of cadastre, we emphasize the civil-legal cadastre in this research. In order to investigate the role of cadastre in decision making and decision making, library resources and questionnaires and questionnaires were used to survey the statistical population and face to face interviews with relevant managers and experts at different levels. In the quantitative analysis of the questionnaires, statistical techniques such as correlation and regression were used in the SPSS software package that resulted in the following results: Using the correlation test, it was determined that between the security of capture and the incentive to invest in the housing sector as well as the motivation for participation People have a direct relationship with improvement plans and using regression it has been found out that secretive and perceived security has a greater impact on the security of legal seizure on the incentive to invest in the housing sector as well as the incentive for people to participate in improvement projects. Therefore, it can be concluded that in informal settlements, in addition to enhancing the legal dimension of seizure, with tools such as the implementation of civil-law cadastre and

S. S. G. Madineh
Surveing Group, Faculty of Engineering, Larestan Unit,
Islamic Azad University, Larestan, Iran
e-mail: ir.apjco@gmail.com

R. Farhoudi (✉)
Tehran University, Tehran, Iran
e-mail: rfarhudi@ut.ac.ir

H. Roosta
Department of Surveing Engineering, Faculty of Civil Engineering,
Islamic Azad University Larestan Branch, Dubai, UAE
e-mail: Hroosta@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_72

the consolidation of ownership and documentation of seizure, it should strengthen and increase the conventional and perceptual dimensions of security of seizure by measures such as increasing Sense of belonging to the environment and promoting the quality of urban, educational and social services. The collection of library resources, as well as interviews with managers and experts, resulted in the ignorance of informal settlements and the prevention of the granting of property, the method of dealing with and the prevention of the spread of these gatherings, as the effects of informal neighborhoods such as poverty and corruption, and Other social disruptions have a direct impact on the city.

Keywords Cadastre · Empowerment · Informal settlements · Ahvaz

1 Introduction

The emergence and expansion of marginal neighborhoods and informal settlements is one of the most important challenges in the development of the city of Ahwaz. So that the problems of marginalization in all cities of Ahwaz are not observed in any cities of the country. Of all types of cadastre, The civil-law cadastre, we emphasize in this research. The proposed issue in this research is the empowerment of informal settlements. According to studies, one way of empowerment is regulating and proving security for the occupants of these settlements. And on the other hand can be a cadastre a good tool for securing security, Therefore, in this research, the role of cadastre in the empowerment of informal settlements will be analyzed.

1.1 Problem Statement

Informal settlement and marginalization are phenomena that are Following structural changes and the emergence of socio-economic problems, such as the rapid flow of urbanization And unbridled rural immigration has emerged in most countries of the world, especially Third World countries. As social marginalization and habitat are combined with a combination of low-income groups, often with informal jobs and unsustainable urbanization, many social damages are considered. The metropolis of Ahwaz, as one of the economic pillars of the country, has always attracted the population. Some of the amenities, culture, education and training have caused the aspirations to attack the city of Ahwaz. The city has been exposed to the phenomenon of marginalization as an official and informal part of a developing country. And a large part of the immigrant population has made construction inappropriate due to economic weakness in parts of the city. These settlements occupy a large part of Ahwaz city And informal settlements Have been created.

1.2 Research Goal

The main purpose of this research is to analyze the role of cadastre in the empowerment of informal settlements in Ahwaz. Since empowerment schemes in informal settlements are not apart from economic, social and physical analyzes. Therefore, the research objectives will be:

1. Explain the civil-law cadastre relationship with security capture
2. nvestigating Different Dimensions of Segregation Security and their Impact on empowerment of Informal Settlements.
3. Investigating Different Dimensions of Segregation Security and their Impact on empowerment of Informal Settlements.

1.3 Research Questions

1. Does the implementation of cadastre can affect the security of capturing residents of informal settlements?
2. Is implementation of civil-law cadastre influencing decision making and decision-making in order to motivate the investment of households living in informal settlements in the housing sector?
3. Does the implementation civil-law cadastre can affect the willingness of people to participate in plans to improve informal settlements?

1.4 Research Methodology

The present project is based on descriptive-analytic research methodology and is applicable to the target. In the descriptive section, documentary, library and electronic resources are used.

2 Theoretical Concepts and Fundamentals

2.1 Concepts Related to Cadastre

2.1.1 Parcel

A piece of land cadastre Is a continuous piece of land in which the rights and the unique exploitation of its individually and uniquely identified [1].

2.1.2 Information Systems

All computer systems that maintain, store, process, and possibly analyze data or information are known as information systems.

2.1.3 Geographic Information System (GIS)

Systems with computer management that store, store, process, analyze and disseminate geographic data And their output may be descriptive information or digital maps [2: 12]

2.1.4 Earth Information System (LIS)

The Land Information System in the International Federation of Surveying has been defined as follows: "The land information system is including legal, managerial, and economic decision-making and planning assistance, on the one hand, databases containing physical information Reference ground is defined for a given region and, on the other hand, it including techniques for collecting, updating, processing, and distributing data" [3: 25] (Chart 1).

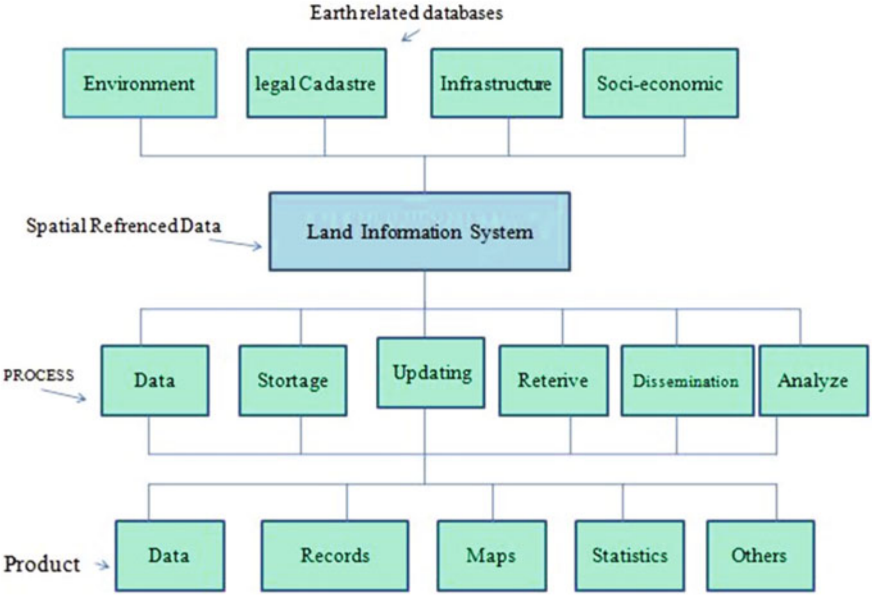


Chart 1 A landing information system [2: 12]

2.1.5 Cadastral Objectives

The purpose of the cadastre is to create a precise, simple, fluid, reliable, and time-consuming military system for the rule of the real estate and lands belonging to real, legal, state, and endowed persons; and to review information and maps throughout the program and Finally, the existing Registered system changes to cadastre.

Quantitative Objectives: Determine the Legal Range of Ownership of Types of Towers and Land Covered by Real, Legal, Governmental and Endowed Individuals.

Qualitative Objectives: To create a workflow in creating a reliable military system for the issuance of a property certificate; to ensure confidence in real estate transactions and property consolidation; to reduce the number of civil claims in courts; to reduce corruption in the real estate sector in the country; to assist in settling A fair tax system in the country and increasing the efficiency of development projects in the country [4: 15–16] (Chart 2).

2.1.6 Cadastral Types

Tax Cadastre

This cadastre is the oldest kind of cadastre, which is based on the measurement and land use of estates and products and the use of land by governments in the old days. The collection of taxes and burdens in the old governments of Iran, Babylonia and Egypt, and elsewhere, in order to confirm ownership and consolidate it, but more to provide the important chapter in the revenue of the state and government treasuries. In the history of drawings and cartography, it is also noted that, since taxes were required, surveys and cadastral maps also emerged. This has also been considered a sign of justice, and in some countries, these cadastral tax receipts are still important treasuries of that country [5: 126–127] civil-legal cadastre.

In some countries, even the beginning of the cadastre was a legal cadastre. And the role of maps in this kind of cadastre is more important and valuable day after day. To coordinate between the codes and the number plates in the documents and drawings (and now in the computers), each owner can use his assets and the interest from which the property belongs to him. Therefore, according to the nature of the real estate, the two categories are civil estate cadastre and civil land cadastre. therefore, according to the nature of the real estate, the two categories are civil estate cadastre and civil land cadastre [5: 128].

(1) Urban Civil Cadastre

In this type of cadastre, the purpose was to determine the location and registration of real estate, tents and streets in the cities. Due to the value, importance and density of urban land, precise methods are used to prepare plot maps.

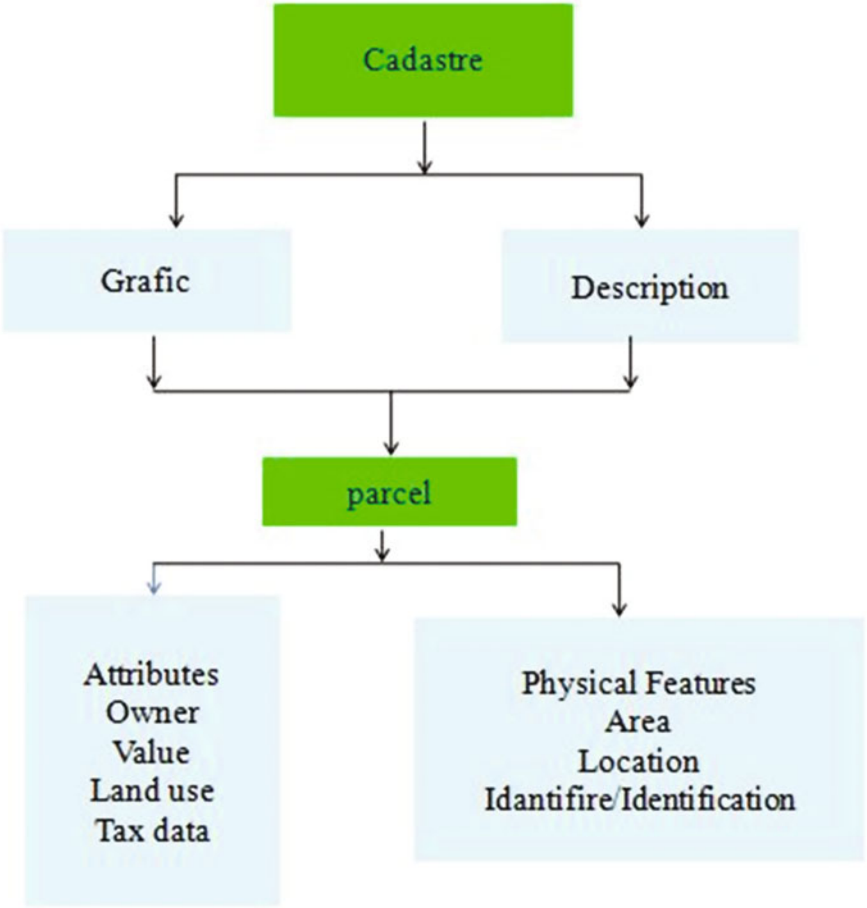


Chart 2 Cadastral system arms [2: 30]

(2) Non Urban Civil Cadastre

The purpose of this type of cadastre is to determine the limits and limits of non-state real estate and In addition, it provides descriptive information about the environment and its applications, and the accuracy of the preparation of the parts map in this type of cadastre is less than urban civilian cadastre [4: 11].

Linear Cadastral

If a cadastre is provided in linear maps with descriptive and descriptive information, the cadastre is linear. In the linear cadaster, the “map” with the “record” is two separate categories. That is why the records of such records refer to clerical documents or lead wire documents. In this kind of cadastre, the description of the dimensions and timing of the estate is with sentences. Like the north at the length of 8 m and 25 cm to the…. Today, globally, due to the development of information technology and the advancement of technology and computer science, the cadastral system is replaced by a three dimensional and four dimensional cadastral system [3: 42–43].

Digital Cadastral

Digital cadastre is a kind of new and electronic registration. This kind of cadastre is considered to be the most common cadastre in the world because of the advantages. In digital cadastre, geometric and descriptive dimensions of land and property are combined [3: 44–45].

Property Cadastre

This type of cadastre, although not widely used in cadastral countries, is another complementary cadastre and is a complementary cadastre. The use of this cadastre is the grading of private, public and public property. The necessity of realizing such a cadastre is the existence of accurate and accurate information from immovable assets on the one hand and updating information and clarifying them over time on the other [3: 46].

Political Cadastre

The political cadastre is related to regional divisions of the provinces, sections, cities and international borders. In this way, the information needed for planning regional and national divisions is collected. More precisely, the political cadastre is divided into national and transcendental cadastres And refers specifically to transboundary cadastres to ground, blue and air boundaries [1: 199].

Geographic Cadastre

Geographic cadastre is a general concept of cadastre which collectively covers different political, human and natural geographic areas, as well as the management of development strategies [4: 11].

Watery Cadastre

The purpose of this type of cadastre is to determine the borders of the countries and the extent of the influence of each country in adjacent waters and the management of coasts and ports [4: 12].

Online Cadastre

Quick and immediate access to a large amount of timely information from cadastre and geospatial databases is Online Cadastre. The first factor that has a profound effect on this type of cadastre is the Internet. The advancement of Internet technologies enables the use of geographic information over the Internet [1: 199].

Multipurpose Cadastre

Multi-purpose cadastre is an integrated system for managing cadastral data. As it combines geographic and non-geographic data and displays graphics in the form of a map. In fact, it includes real estate and property rights management, property rights management, landlord and tenant, view layers of maps and subject maps and supervision of land use [6]. This kind of cadastre in addition to the objectives of civil and physical cadastre, is responsive to the needs and programs of development and economics and the provision of social services in relation to land plots [1: 200].

2.2 Informal Settlement

2.2.1 Definitions and Common Terminology for Informal Settlements

Marginalization

Marginalization consists of a settlement or form of residence that is based on the spontaneous and spontaneous movement of people (residents), based on the desire, motivation and based on their potential and facilities. Marginalization consists of a settlement or form of residence that is based on the spontaneous and spontaneous movement of people (residents), based on the desire, motivation and based on their potential and facilities. Due to the characteristics of the car and the unplanned turnaround, these settlements are weak in providing services and infrastructure, and they live in minimal services and facilities. As well as land ownership and the separation of land and the use of land outside the legal framework and without obtaining the necessary permits of official devices [7: 107–108].

2.3 The Relationship Between Cadastre and the Security of Land Occupation in Informal Settlements

According to the definition of the World Bank, the empowerment of the informal settlement of land regulation is one of the ways of empowerment to improve the physical environment of informal settlements. And it can be said that one of the aspects of land regulation is the stabilization of ownership of individuals and the determination of rights to land. The existence of a cadastral system with the storage of records of property, causes the property to be protected and defended in civil litigation and disputes. Or after crises such as earthquakes, fires, floods, etc. It is possible to prove ownership. Also, in the cadastre of registration issues such as the limitation of the limits of occupancy of land and land rights to consolidate and consolidate ownership, and this cadastre can provide for the acquisition of legal grounds for the registration of land and thereby secure the seizure of land law Upgrade. It is necessary to note that the cadastre of registration is not the grantor of property rights to the possession of individuals. But merely as a geospatial reference database that contains geometric information and records of property and property rights And it can both legal and geometrically enhance the security of legal possession of the land and provide a place for consolidation and consolidation of ownership and land tenure. for example: If the property has a registered record in the records and offices of the registration offices of the documents and real estate, the property occurs. And these records have been stored and stored at the Cadastre database, although the property is not officially owned, but, despite the existence of the records and the hierarchy in the Cadastre database, the legal seizure of that property can be guaranteed and defended [8].

3 Case Study

Ahwaz is the second largest city in terms of size after Tehran and the seventh Iranian city in terms of population. In Ahwaz, the issue of marginalization is a function of the process of marginalization of the whole country. The emergence and expansion of marginalized neighborhoods is one of the most important challenges in the development of Ahwaz. Therefore, due to the high volume of marginalization and its problems and problems, the extent of these areas in the city, in fact, indicates the quality of life of some residents of Ahvaz. As of 2011, Ahwaz has a population of 1,080,955 people, of which more than 400,000 people have been marginally resident in informal settlements (Fig. 1).

In this study, eight informal settlements neighborhoods of Ahwaz are described in the Table 1.

Fig. 1 The location of the Ahwaz informal settlements

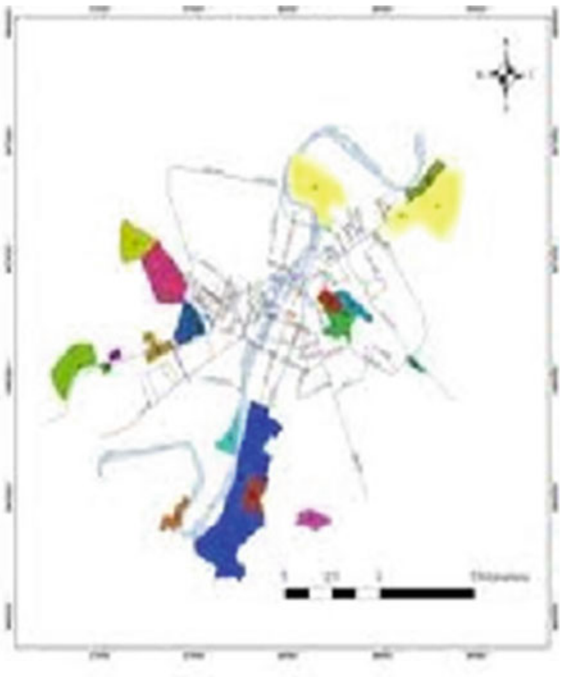


Table 1 Population dimensions and area of the 8 marginal areas studies

Area code	Neighborhood name	Total population	Household dimension	Number of households	Area (sq. km)
11	Shilan G Abad and Sayyahi	58,621	7.5	10,313	11.29
52	Eyn	13,288	7.5	2451	2.22
63	Molashieh	26,094	5.1	5060	3.03
13	Al Safi	7717	6	1320	0.96
14	Water source	18,944	4.82	3925	1.58
15	Hasir Abad	26,795	4.8	5724	0.75
16	Zargan	10,784	5.2	2084	1.27
17	Zovieh	13,924	4.6	2043	92.10

4 Discussions and Findings

4.1 Reviewing the Implementation of the Cadastral Plan on the Security of Capturing Residents of Informal Settlements and in Deciding and Deciding to Motivate Investment

The data needed for this study was collected through field studies and field visits, completion of a citizens' questionnaire, and the determination of the sample size from the 8 marginal areas of the study population of 322 citizens by simple random sampling. The validity of the questionnaire was assessed through Cronbach's alpha. The validity of the questionnaire was 0.76, which indicates the acceptable validity of the questionnaire. To investigate the implementation of the cadastre plan on the security of capturing residents of informal settlements. In connection with the above hypothesis, the indicators (land ownership status and residential unit, certificate of construction and completion of work, major materials used in residential units, sub area and The age of the residential building, the number of rooms in the residential unit, facilities and facilities (water, electricity and gas) in the residential unit and how to access these facilities has been studied. The results of descriptive statistics showed that there is a significant relationship between the implementation of cadastral plan and security capture And the security index is 0.799 and the cadastral plan is 0.475. The results indicate that there is a strong relationship between the implementation of the cadastral plan and the occupancy security of the residents of the neighborhood that will lead to the security of capture by citizens through the implementation of the cadastre project in the region, which will contribute to investment and optimal quality of life (Table 2).

In order to investigate the effect of the implementation of civil and legal cadastre, in the decision-making process to motivate the investment of the households living in informal settlements in the housing sector, out of the 265 respondents mentioned, they participate in the projects, based on data Table 3, 40% of respondents stated that they are willing to participate in the implementation of the cadastre project in the implementation of the projects of participation participation (building, physical, asphalt, etc.), 27.92% of the respondents (74 people) to participate Thoughts (studying and presenting the plan), 14.71% (39 people) believed in financial participation and 17.35% participation in all cases.

To measure the significance of the above hypothesis and considering the level of measurement of the variable, T-test of a single sample has been used (Table 4).

According to the results of Table 3, and based on the significant level of T-test (0.00), it can be said that at a level less than 0.05, there is a meaningful relationship between the implementation of the cadastre plan and the motivation of citizens to invest in the physical section of the city.

Using a question at the scoring level, two categories of respondents were asked to indicate whether they were willing to contribute to the cadastre plan in community development and improvement projects (plans that would improve living

Table 2 Implementation of the cadastral plan on the security of capturing residents of informal settlements

Components	Number	Pearson correlation statistics	Meaningful (sig)
Capture security	322	0/719	0.000
Run the cadastre plan	322	0/475	0.000

Table 3 Distribution of responses according to the type of participation and investment

Participation and investment motivation	Number	Percentage relative
Financial participation	39	14.71
Intellectual participation (studying and presenting the plan)	74	27.92
Executive participation (building, physical, asphalting, etc.)	106	40
All items	46	17.35
Total	265	100

Table 4 T-test sample (economic indicators)

Single sample T test				
Difference in averages	The significance level	Degrees of freedom	T Statistics	
14.11491	0.000	265	55.215	Score

Table 5 Distribution of responses by participation in community improvement and organization projects

Participation	Number	Percentage relative
Yes	265	82.3
No	57	17.7

conditions in the residential neighborhood). And the answer options in the order are yes = 1 and no = 2. Table 5 provides descriptive statistics.

4.2 The Opinion of the Relevant Experts

4.2.1 Identification of Effective Factors in the Formation of Informal Settlements from the Perspective of Experts

Understanding the fact that the factors that are more active in the formation of informal settlements are highlighted in the first place can be helpful in helping the planning authorities to solve the problem. For this purpose, in the present study, we tried to use a five-point Likert scale (very high = 5, high = 4, mean = 3, low = 2 and very low = 1) in a questionnaire. The variable discussed will be discussed. The Table 6 describes the status of the dispersion of responses in terms of individual

Table 6 Frequency distribution of respondents in terms of the effects of formal recognition on land ownership in informal settlements

Positive and negative impact on the ownership of land in informal settlements	Frequency	Valid percentage
Increase the residence time of residents in the neighborhood	0	–
Encouraging more populations of low-income families to settle in the neighborhood and increasing congestion	4	30.8
Supervision of the security of the capture of residents in the occupied arena	6	46.2
Prevent the spread of physical, economic and social problems of the neighborhood	1	7.7
Encouraging residents to participate in community development and improvement projects	2	15.3
Total	13	100

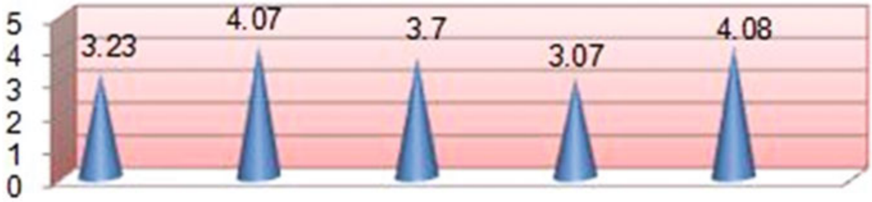


Chart 3 Distribution of the distribution of responses in terms of effective factors in the formation of informal settlements

items. From the perspective of urban experts, the game of land and housing is another important factor in the formation of informal settlements. This, in turn, causes an uncontrollable increase in the price of land and housing in parts of the city with suitable facilities, as well as a factor for directing low-income groups to a part of the city that is in poor condition and It is not environmentally friendly (Chart 3).

4.2.2 Investigating the Positive and Negative Impacts of Land Ownership on Informal Settlements

In order to measure the positive and negative effects of formal recognition on land ownership in informal settlements, 5 options were presented, which is explained in the table below and its frequency and percentage.

According to the information in the table above, the majority of experts, in other words, 46.2% of them considered better urban management in the neighborhood to capture residents in occupied areas as a positive effect on the ownership of land in informal settlements. Approximately 30.8% encouraged and inhabited by more and more middle-class and low-income people to choose their neighborhood to live and

Table 7 Frequency distribution of respondents in terms of the effect of cadastre plan on informal settlements in the face and economic, social, spatial and environmental structure of Ahwaz city

Influence of informal settlements on the face and economic, social, spatial and environmental structure of Ahwaz city	Frequency	Valid percentage
Too much	1	7.7
Alot	6	46.2
Somewhat	4	3.8
Little	2	15.4
Very little	0	–
Total	13	100

increase their density, 15.4% of the officials encouraged residents to participate in community improvement and organization projects and the rest (7.7%) to prevent the spread of problems The physical, economic and social aspects of the neighborhood are among the positive effects of the recognition of land ownership.

4.2.3 The Impact of Informal Outreach on the Faces and Economic, Social, Spatial and Environmental Structure of Ahvaz City

Among other variables questioned by the experts of the city, the impact of informal settlements on the face and economic, social, spatial and environmental structure of Ahwaz was studied, of which 53.9% believed that the issues Economic, social, spatial and environmental informal settlements have a great influence on the situation in Ahvaz city and about 30.8% of officials are of the opinion that the problems of the region to some extent this city is difficult and 4/15% of them expressed low neighborhood impact on increasing problems in Ahvaz city (Table 7).

5 Conclusion

The structure of cities consists of various social, economic and physical factors, and it will be impossible to examine the space of a city without considering these factors. The structure of cities changes in the course of its physical expansion under the influence of various factors, such as marginalized and informal neighborhoods. The role of informal settlements is unclear in creating undesirable spatial, economic, social and environmental impacts on urban structures. And as these settlements expand, the volume of undesirable effects on the city's image will be more pronounced And ultimately leads to a structural fragmentation of residential areas from the whole city, leading to a reduction in the level of social service and housing quality, as well as the marginalization of the inhabitants of these neighborhoods. Civil and legal studies of land based on field studies show a clear gap between informal settlements and quality of life indicators and the official district of Ahwaz.

The main problems of informal settlements in the social, economic, physical and environmental areas are as follows: In interviews with managers and experts, in order to investigate the effect of cadastre in decision making and decision making, different views were expressed that it can be concluded that a proper and efficient legal and legal cadastre is available due to the availability of geometric information and The legal rights of individuals' possessions can be effective in realizing better and faster urban design plans for informal settlements. The ignorance of informal settlements and the prevention of ownership is not the way to deal with and prevent the spread of these gatherings, as the effects of informal neighborhoods, such as poverty and corruption, and other social disruptions, have a direct impact on the city. Also, the existence of a cadastre is not built on land, and plots of land can hinder the expansion and penetration of informal settlements. This could be an effective step in preventing the spread of these settlements and an appropriate tool for protecting public land. On the other hand, the ownership of these settlements should not be such as to encourage residents to sell land and, as a result, transfer these settlements to another location.

References

1. Henssen J (1995) Basic principles of the nain cadastral systems in the world. In: Proceedings of the one day seminar held during the annual meeting of commission 7, cadastre and rural land management, of the international federation of surveyors (FIG), May 16, Delft, The Netherlands
2. Yousefi R (1380) Cadastre dignami, 1st Printing. National Surveying Organization Publishing House, Tehran
3. Behnam HP, Hanifi M (1391) Cadastral Law (New Registration), First Printing, Tehran, Publishing House 14, Ganj Danesh
4. Assi MR (1391) Recording surveying (Cadastre), First printing. Sima Danesh Publications, Tehran
5. Pourkmal M (2002) Introduction to cadastre recognition and its applications, 1998, 1st edn. Gorgan Geographical Information Center, Tehran
6. Larsen G (1997) Land registry and cadastre systems: tools for information and land management (trans Pourkmal M). GHG Publications of Tehran, Tehran
7. Karegar B, Sarvar R (2011) City marginal and social security, 1st edn. Gorgan Army Geographic Publications, Tehran
8. Mujahed M (1393) Analysis of the role of cadastre in the empowerment of informal settlements (case study: Ghaemshahr). Master's dissertation, Tehran Central Azad University

Threats of Social Engineering Attacks Against Security of Internet of Things (IoT)

Mohsen Ghasemi, Mohammad Saadaat and Omid Ghollasi

Abstract Internet of Things is a fast growing technology. Home appliances, clothes, traffic lights, cars, cameras, and more things used by human are prone to be connected to the IoT. Any new technology has its own challenges. Security is one of the greatest challenges of IoT. Security has been a challenge about the Internet, but in IoT domain, security is wider problem. For instance, manipulating the system of traffic lights will disrupt the security and public discipline. The daily life of human being is involved in the IoT, always and anytime is carried or controlled by users, therefore a vital role is going to be played in human interactions by IoT. Influencing the social interactions of individuals and their everyday lives can mean the penetration in the IoT and challenging the security. Social engineering is not a new concept; SE is an old category that is growing steadily with no end in its vision. SE is effective because human prefer to trust naturally. Social engineers target human factors and use the trust to thieve the purposed information. Generally, SE influences social interactions. Using SE against IoT and lack of knowledge among society will be able to cause a tragedy. SE is introduced in this paper and then security challenges of IoT will be discussed. People's behaviors to keep their security will be discussed in this paper.

Keywords Internet of things · Social engineering · Security of IoT
Social engineering attacks · IoT challenges

M. Ghasemi (✉)
Department of Computer Science, Tehran University, Tehran, Iran
e-mail: moghasemi@ut.ac.ir

M. Saadaat
Department of Computer Science, Khayyam University, Mashhad, Iran
e-mail: msaadaat@live.com

O. Ghollasi
Department of Computer Science, Shiraz Industrial University, Shiraz, Iran
e-mail: Omid.ghollasi@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_73

1 Introduction

It's not the time to know who has created the world, you have to see who are going to destroy it; Chomsky said [1]. The definition of social welfare has been changed, since 1980s. Today, an important feature of social welfare is feeling of security [2]. Cybercriminals who are targeting people security in cyberspace, in fact they are targeting social welfare of society. Social welfare is one of the main tasks of governments, and failure to establish it will cause chaos. Therefore, cyber security is connected directly with national security. The main challenge in today's digital and informal world is the cybercriminals; people who are ahead of technological advancements. Cybercriminals costs governments millions of financial losses.

IoT is an emerging and fast-growing technology in the world today. Nowadays, stored data is being increased rapidly. This data can be accessed and processed by devices such as sensors, cameras, and mobile phones; these devices are connected by internet and share data with each other. This sharing of data and connection of various devices is called Internet of Things (IoT) [3]. IoT is a new technology in which it provides the ability to send data over the Internet-based communication networks to any entity. All real and legal things in IoT space are things with ID and share data over a integrated network. IoT provides the ability to communicate all things to each other and human with a specific ID under an integrated network. In general, the IoT is the set of Standards, Protocols and technologies for communicating and transferring data over smart devices globally [4]. Cisco experts have chosen years 2008–2009 for appearance of IoT; because the number of connected devices to the Internet surpassed the global population during this period. The Internet which was known as "Internet of people" was renamed to "The Internet of things" [5]. According to the Gartner Institute, by 2020, there will be more than 26 million internet-connected devices which are subset of IoT [6]. This mass of devices shows the mass of data which are producing and being increased.

The IoT has its own echo system. Various aspects must be considered in advancements of IoT; such as investment, storage, integration of devices, hardware, software developing, analysis for using in industry and most important that others, the security of data shared by IoT devices. Security has also various aspects; such as Antiviruses, Secured designing, customer service for security and user's conducts. Unsecured conducts of users are discussed in this paper. Using the interactions on users for unauthorized access to sensitive information is known as "Social Engineering". SE is recognized as the most serious threat in cyberspace and it is an effective tool to attack to information systems. Employees provide a purposed environment for Attackers by services they are using today [7]. The main issue which is discussed in this paper is "weakness of IoT users against SE threats". SE and its tricks will be introduced in past "Social Engineering". The research method in this paper is based on a questionnaire distributed among 1570 students in a field study and its results are discussed in part "challenges and weakness of users conducts". Simulation in SE is done by scenario. In this paper simulation is done by real scenarios. The title of this paper has no literature review.

2 Social Engineering

SE is knowledge of using social interactions as a tool to persuade a person or organization in order to agree with the request submitted by attacker [8]. SE is “duping” in slang. “duping” is not a new category but it has always been used by human in their interactions. The globalization has made “duping” a knowledge. Various definitions of SE have been presented but the theoretical form of SE has been discussed since 1980s. Comprehensive definition of SE was presented by authors in [9]. Social engineering is the art of deceiving human and hacking their social conducts in order to gather sensitive information [9].

SE attacks are different from other cyber-attacks. SE attacks can be implemented both in form of cybercrime and also as form of face-to-face crime. Cybercrime form of SE attack is known as computer-based SE attack and the face-to-face attack is called human-based. The basic form of SE attacks is happening in human-based and according to human-based form of SE attacks, a new form is adopted as computer-based form in cyberspace. The model of SE attacks consist of seven items or classes showed in Fig. 1. Various parts of the model are not explained to avoid prolonging the paper.

SE attacks are very diverse and can be newer and innovative, depending on creativity and agility of attacker. In next part a SE attack is simulated by a real scenario and is adopted with SE attack model above.

3 Simulation of a Social Engineering Attack

Victor Lustig is one of the biggest social engineers in the history, a man who sold the Eiffel Tower. The idea of selling the Eiffel Tower came Lustic mind after reading an article in the newspaper. Lustig immediately got started. He first

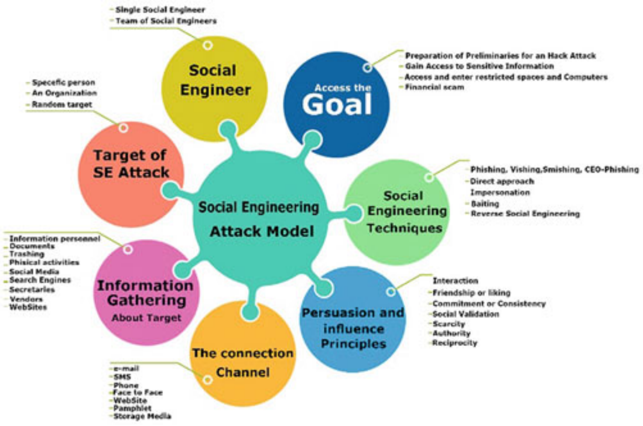


Fig. 1 Social engineering attack model [9]

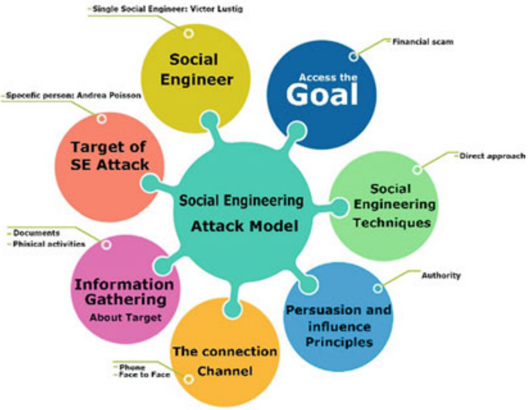

-Single Social Engineer: Victor Lustig
Social Engineer
-Financial scam
Access the Goal
-Specefic person: Andrea Poisson
Target of SE Attack
Social Engineering Attack Model
-Direct approach
Social Engineering Techniques
-Documents
-Phisical activities
Information Gathering About Target
-Authority
Persuasion and influence Principles
-Phone
-Face to Face
The connection Channel

Fig. 2 Model of Victor Lustig Attach adopted on SE attack model

collected documents in which he introduced himself as Deputy Chief of the Ministry of Post and Telegraph. Then Fake letterheads was prepared by Lustig. Lustig invited six famous businessmen to a governmental and secret meeting in Creon hotel; A hotel known for diplomatic appointments. Six invited businessmen showed up on time. Victor stated that the government is in bad financial condition and the maintenance costs of the Eiffel Tower are out of government ability. I am on the government's mission to sell the Eiffel Tower despite the savage and regrettable. The best customers are honest businessmen and you six persons are the best, Lustig said. Victor emphasized that it should remain hidden and secret until it is finalized because of the possibility of a general opposition to this issue. Andre Poisson was one of six businessmen who was beginner and wanted to walk through the centuries. Victor as a clever social engineer realized Poison. The businessmen offered their prices and then Victor announced to Poisson that he had won the tender; and the documents for signing and delivery of the tower are ready in the hotel. Lustig announced that he was a simple employee in order to Persuade Poisson to pay bribes. After the bribes were paid, the transaction documents were signed, and Andrea Poisson considered himself owner of the Eiffel Tower. The next day, when Andrea Poisson and his workers were arrested for the destruction of the Eiffel Tower, Victor was miles away from Paris [10]. The simulation of Viktor Lustig's attack based on the model of social engineering attacks is showed in Fig. 2.

4 SE Attack Is Unpredictable Technically

Nowadays, while the use of the Internet is increasing rapidly, the security of information is deteriorating. Anti-virus's software is one of the best solutions to improve the security of systems. Antivirus software has the pattern of a virus, worm and malware and if user run programs according to the patterns that the antivirus already has, detects it and informs the user [11]. Social engineers focus on human's conducts; and people at any moment show unpredictable behaviors and are directly influenced by the surroundings and individual interactions. therefore, it's possible to create patterns of human momentum behaviors, so it's impossible to develop a software such as antivirus based and pattern of human conducts. The weaknesses of human behavior that social engineers use as a vulnerability are unpredictability, under the influence of external factors, think and action instantly and more important than others desire of individual for trust. This is the reason why SE attacks are unpredictable and needs more attention than Hack attacks. In some attacks such as phishing, it can be somewhat less dangerous technically; but in many SE attacks, there are no other means of preventing except training and modifying user behavior.

5 Introducing Some Social Engineering Attacks

- **Phishing, Vishing and Smishing**

Phishing means hunting the user's password through the prey (fake website) [12]. Different definitions for "phishing" has been purposed. The PhishTank forum is a global and public project that helps people identify, detect, and share phishing sites' addresses. The PhishTank has provided the following definition for phishing attacks: Phishing is a fraudulent attempt, usually made through email, to steal your personal information [13].

The attackers copy the website of the organization that provide online services for users and users have account in these websites such as banking websites. In the next step, the fake website will be shared by email, pop-ups and online advertising. If users trust this website and use their confedtials to enter their accounts, First, they have given their personal information to Phishers, then they are guided by a trick to the main website to not doubt the scenario. The address of the fake websites is also similar to the original site. For instance, the address of Mellat Bank's website is "Mellatbank.ir", the Phishers will choose the address "Melllatbank.ir" for their fake website. The art of SE in Phishing attacks is persuasion and deception of users to trust the fake website. "Vishing" is the same phishing attack that is carried out using a phone and in a telephone service such as "Telephone Banking". In Smishing attacks, Fake links are provided for users by SMS. There are other isotopes of phishing that are not discussed in this paper, such as CEO Phishing, Spear Phishing.

- **Baiting**

The Baiting is another trick that social engineers use. Prey like a flash memory that contains malware is released somewhere like office and coffee shops. Victims of this prey take it by the perception of being lost by the owner and used it with curiosity. They install malware on their own computer with satisfaction. Malware is not a part of social engineering; That the purpose of the baiting is to use flash memory and run it by victim [13].

- **Reverse Social Engineering**

The Reverse SE is that the attacker first creates a technical problem in victim's computer system. When the user failed to resolve the problem, the attacker indirectly identifies himself an expert in that area. The victim is in hurry and wants to keep up his work, that's why the victim asks the attacker to help him and provides the excellent situation for the attacker [14].

- **Direct approach**

A typical example for the direct approach is Lustig's attack. In this attack, the user has a direct and face-to-face relationship with the target and these attacks require strong social and communication abilities. Identity theft, piggybacking, Reverse Social Engineering are some kinds of direct approach.

- **Nigerians attacks**

Nigerians attacks are emails that target passions of victims and then they ask them to pay for various excuses. Sense of greed is raised in victims. There were funny things about Nigerian attacks, like a winning claim, huge amounts in lottery, helping people in Indonesia's tsunami who need aim, helping the Syrians, inheriting the high amount from where it does not exist.

6 Challenges and Weakness of Users' Conducts

Security challenges of IoT are more important in comparison with old computer which used The Internet. Old and current computers in case of security threat are prone to leak private information and industrial data, but hackers are not able to perform an action by these hacked computers. The story for the IoT is very different. For example, if the control of the heater and security sensors of a home which uses the IoT is hacked, it will be possible to fire a house for hackers. The IoT is fast growing technology, In the near future, all parts of social life will be involved by IoT. Control of various devices connected to the IoT must be done by the owners of the things. Controls of things that are connected to the Internet have different layers and many security features are involved in their security. If controllers of sensors and things that use the IoT fall into criminal's hands, there will be tremendous consequences for personal and family life, and in large scale for national security of the community. Security challenges for IoT are shown in Fig. 3. Its items will be discussed further and social engineering will be discussed in detail.

- **Authentication**

IoT provides Internet communication between different things that belong to different people. Authentication is an issue that matters in all Internet-based technologies. Whether the person who is using the device is the person who is allowed to use it, is called "authentication". Technologies like fingerprints, Two-step verification, Face Recognition and Password are used for authentication.

- **Confidentiality**

Information is only available to authorized users who request purposed information, this method of access to information is called "Confidentiality". The principle of confidentiality means that data is only available to authorized people at the required time. For example, on the disclosure of US military intelligence information by Edward Snowden, Snowden was not allowed to access to information Because Snowden was a contract worker, that's why the principle of confidentiality has been violated in this process.

- **Privacy**

The right of individuals to be protected from any interference and search in their private life, is called privacy; Privacy, both physical and virtual, and personal information [15]. In IoT the privacy principle is seriously threatened, because device producers cooperate with governments to disclose information. All efforts to protect privacy are aimed at protecting the privacy from each other violation and it's not protected from government. Privacy settings on the devices of IoT are set by Companies and it is necessary for users to adjust it according to their instructions after purchase.

- **Secure middleware**

Middleware is actually the operating systems that devices use. Smart TVs, headphones, and other devices of IoT use specific middleware belong their own device. Security of the middleware of the IoT is a challenge that must be taken seriously.

- **Mobile security**

The mobile device is considered as a tool for controlling the IoT. Things connected to the IoT must be controlled by a mother device owned by the owner. The security of this mother device is more important than other devices. The new WikiLeaks publishes show that The US intelligence agency has control of Android and iOS devices.

- **Trust**

The concept of "trust" is used in different fields and different meanings. Trust is a complex theory, and there is no definitive agreement on its meaning in scientific literature, but its importance has been discussed extensively. Devices connected to the IoT are things that are always where the owner is, and are fully humanized. Therefore, they are often present in public places and is connected to wireless networks. Social interactions, friendships, and communities affect people to use each other's devices, and this makes it possible for someone to let their friend use their device [16]. Managing trust in these relationships is one of the challenges for security of the IoT. SE exactly focuses on social interactions, so IoT is a good domain for Social engineers.

- **Policy enforcement**

Organizations have their own security policies, and these policies are notified to employees. The IoT in addition to individual life is also widely used by organizations, therefore, organizations need to take the precautionary measures to ensure that their policies are enforced by employees.

- **Access control**

Organizations should control the access of their employees to IoT devices. Access permission to information for people with different positions should not be the same.

- **Social Engineering Tricks**

Social engineering focuses on the behavior of IoT users; behaviors whose weaknesses provide the opportunity for social engineers to take advantage of. In this paper, users' behavior is discussed in four sections: Choosing Password, Importance of privacy policies, pay attention to the App permissions and terms of use, use personal information to browse through cyberspace. Each of these cases is discussed separately. The statistics of the four questions are shown in the Chart 1.

- **Choosing Password**

Password is one of the vital tools of today's electronic world. The password is required to use the IoT, because the first method of authentication is using the password. Imagine someone using same password for all devices connected to the IoT and his password is also very simple, and social engineers can guess it with little effort; disaster will occur and criminals are able to control his life.

The behaviors that individuals show in in case of choosing password can directly affect their security. Social engineers collect identifying information of a person who uses this information as password, and they can guess the password with trial and error. Sequential passwords like "123456" is one of the worst ways to pick a password. In this paper, 1570 undergraduate students have been asked which policies they follow to choose their password. The question and choices are as below:

Question: What policy do you have for choosing a password?
First choice: Personal information such as ID number?
Second choice: Sequentially like "123456" and "abcdef"?
Third choice: Exclusive password depend on your own?
Fourth choice: Using password management software?

37% of the respondents have secured behaviors, while other respondents do risky behaviors. The idea of some people who use a simple password is that because they do not have a valuable virtual asset, they do not have to worry; but the most valuable asset of users is their identity at first. Today, user credentials that has been

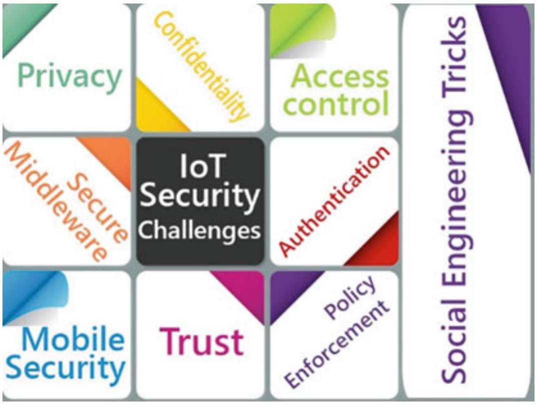
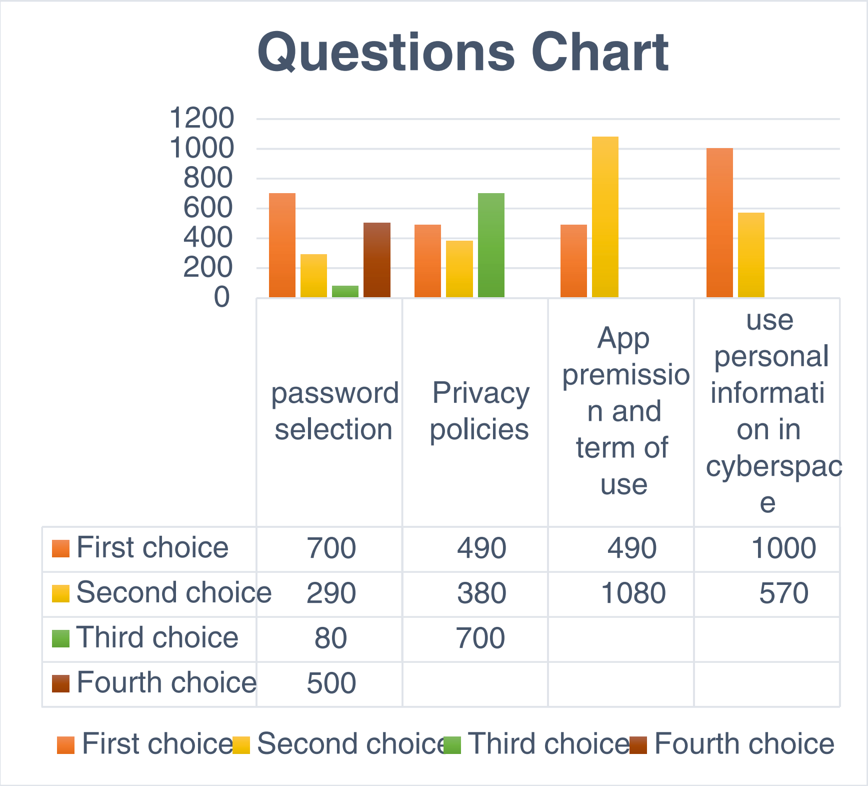


Fig. 3 Security challenges for IoT



	password selection	Privacy policies	App premission and term of use	use personal information in cyberspace
🟧 First choice	700	490	490	1000
🟨 Second choice	290	380	1080	570
🟩 Third choice	80	700		
🟫 Fourth choice	500			

Chart. 1 The statistics of the four questions asked 1570 students

hacked from companies such as Yahoo, has been released in a Dark web and used by criminals for identity theft.

- **Importance of privacy policies**

Application developers and producers of devices will consider policies for their products and, with that in mind, offer options in products that users with their settings can decide if their privacy protections are strict or not. Privacy settings are not set as default by the manufacturer strictly, and the user himself must set them up to his liking. For example, there are two-step verification in many famous applications, but low percentage of users use it. The sensitivity and awareness of users about privacy policies and their related settings has been evaluated using the question and the statistic are shown in Chart 1 and the questions with choices are as below:

Question: How do you deal with privacy settings?
First choice: I'm not familiar with Privacy settings.
Second choice: It does not matter for me.
Third choice: I'm sensitive and set it according to my liking.

The statistics are not promising; 44.5% of respondents said they were sensitive and the rest were either unaware or not reckless.

- **Pay attention to the App permissions and terms of use**

The applications are becoming a big threat since there is no monitoring of their distribution and security standards. Developers can easily reach out to their massive audience without filtering them to assess their security. Applications that are used on devices can have access to different parts of the device. Criminals use this feature and provide applications with specific accesses, and the user installs it, and then the device is virtually controlled by criminals. It has been seen that the calculator application with the has access to the memory card and the user's camera, while it does not need any access to it. Users' attention to these accesses is very important. The behavior of users in this area is also evaluated with the question. The statistics are shown in Chart 1 and the question with choices are as below.

Question: When you install an application, do you review the terms of service and accesses of it?
First choice: Yes!
Second choice: No!

Unfortunately, in this behavior, the results are not promising and only 31% notice this option.

- **Use personal information to browse through cyberspace**

Creating an account on a site like Gmail or Facebook requires the registration of individual information; this information includes identity information and photo. This information is asked again when the user forgets his password to recover it. Social engineers can guess password, If the social engineers have the purpose information of target and the individual has recorded his real personal information. The service provider's sites have somehow thought about the risks, but the risk is still not resolved. The behavior of users in this case has been evaluated using a questionnaire. The question and choices are as below.

Question: Will you enter your personal information accurately when you open an account on a site like Gmail?
First choice: I enter correct!
Second choice: I enter wrong information!

1000 of respondents said they were entering correct information and showing that the situation is very worrying.

7 Conclusion

There is no patch for human stupidity. Human behavior will never be 100% secured as the technology is not secured. The technology's speed is very high; however, the criminals do not stay away from this speed. Therefore, organizations and individuals must be trained and ready to maintain their own security. Technology culture is slower than its development, and needs to be reviewed seriously. Social engineers are a potential threat to the technological world today because security systems have made significant advances in protecting data. Hacking is an essential element of knowledge and technical excellence; while social engineering does not require technical knowledge and has a better return. Social Engineers instead of attacking security systems, goes to the authorities of these systems and hacks their minds. Difference of hacking and SE is shown in Fig. 4.

The Internet of Things can be a great option for social engineers because it has entered into human interactions. The strategy to prevent and mitigate the risks of social engineering attacks is increasing user's awareness. Training should be continuous and sustainable to be useful; because the types of attacks are constantly changing, and users are forgetful.

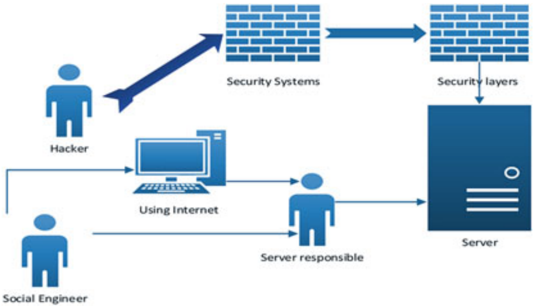


Fig. 4 Difference between Hack and Social Engineering

References

1. Chomsky N (2017) Noam Chomsky website, 17th March 2017. Retrieved from https://chomsky.info
2. Hezar Jeribi J, Safari Shali R (2011) Social welfare and factors affecting it. J Res Plann 1–22
3. Kholod I, Petuhov I, Efimova M (2016) Data mining for the internet of things. Springer, St. Petersburg, Russia, 26–28 Sept 2016
4. Nejad Kamali M, Ali Abadi S, Kahani M (2017) Analyzing the Mirai Botnet based on IoT. In: 2nd conference on cyber space security incidents and vulnerabilities (CSIV 2017), Mashhad, Iran
5. Evans D (2011) The internet of things. how the next evolution of the internet is changing. Cisco White Paper, Cisco System
6. Gartner (2014) Gartner says the internet of things installed base will grow to 26 billion units by 2020
7. Krombholz K, Hobel H, Huber M, Weippl E (2015) Advanced social engineering attacks. J Inf Secur Appl 22:113–122 Elsevier
8. Mouton F, Leenen L, Venter H (2016) Social engineering attack examples. Comput Secur 59:186–209 Elsevier
9. Ghasemi M, Saadaat M (2017) Toward introduction of Social engineering as a threats against the security of personal and professional information. In 2nd conference on cyberspace security incidents and vulnerabilities (CSIV 2017), Fersousi University, Mashhad, Iran
10. Maysh J (2017) Smithsonian, 9th March 2016. Retrieved from http://www.smithsonianmag.com/history/man-who-sold-eiffel-tower-twice-180958370/. Accessed on 24 Mar 2017
11. Dien NK, Hieu TT, Thinh TN (2014) Memory-based multi-pattern signature scanning. In: First international conference on future data and security engineering, Ho Chi Minh City, Vietnam
12. Hadnagy C (2009) Social engineering. The art of human hacking. Wiley
13. PhishTank Retrieved from http://www.phishtank.com/what_is_phishing.php. Accessed on Mar 30 2017
14. Ghasemi M, Saadaat M (2016) Social engineering. Naghoos Publication, Tehran
15. Aslani H 04 05 1391. Retrieved from http://www.urmialawyer.ir/articles/7755. Accessed on 1396 1 18
16. Sicari S, Rizzardi A, Grieco L, Coen-Porisini A (2014) Security, privacy and trust in internet of things: the road ahead. Comput Netw

Assessment and Modeling of Decision-Making Process for e-Commerce Trust Based on Machine Learning Algorithms

Issa Najafi

Abstract Decision-making on trust in an e-commerce environment is associated with several other elements such as security, risk, satisfaction, loyalty and reputation. The life cycle of e-trust in online Business to Consumer (B2C) transactions takes multiple stages from beginning to end. The first stage involves the complete unawareness of online shoppers about online sellers. Then, individual trust begins to develop, endure, recover and ultimately sustain or diminish. Since it is a complicated task to gain trust, there have been numerous solutions, methods and models proposed so far to create, maintain, measure, enhance and prevent loss of trust. One of the solutions that has long been adopted to determine trust level in B2C is a concentration on the history or background of online shoppers (customers) and online sellers (companies) so as to obtain reliable data or identify trust level. This paper attempted to adopt the machine learning algorithms to analyze decisions about the past and history of individuals/companies and trust in e-business/ e-commerce (EC). Moreover, efforts were made to identify and assess the key contributing factors to decision-making. The results demonstrated that corporate factors and business models left the greatest impacts on customer decision-making in e-business trust or distrust during electronic transactions.

Keywords e-trust · e-commerce · Decision-making · Assessment
Modeling · Machine learning algorithm

1 Introduction

Despite improvements in today's smart practices within the e-commerce environment to resolve two major challenges in gaining trust (virtual meeting of the parties to a contract and intangible presentation of goods/services during an online transaction), there are grounds for a great deal of risk against the parties to an online

I. Najafi (✉)
Computer Engineering Department, Quchan University of Technology, Quchan, Iran
e-mail: najafy@qiet.ac.ir; issa.najafi@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_74

transaction. Each party is particularly concerned about online implications of such interaction. In this regard, numerous studies have been conducted, a few of which will be explored in the next section. Nonetheless, the multidimensional structure of trust in e-commerce as well as the difficulty of gaining trust make it inevitable to carry out further extensive research. This study intended to discuss the issue from a new perspective.

Reference [1] explored trust in Internet banking across India. Considering that Internet banking in this country is newly emerging, this paper attempted to explore various aspects of trust. The main objective was to construct e-banking user profiles through intelligent algorithms as well as machine learning algorithms. In the first stage, the variables potentially contributing customer trust in e-banking were identified. Then, the analysis focused on user profiles through different machine learning algorithms such as classification, regression tree, support vector machines (SVMs) and neural networks (NNs). At the end, the results of machine learning predictive models were compared against the results of statistical methods, which indicated an improvement through the newly proposed method.

In [2], effort was made to analyze the client-client, client-device and client-domain transactions through machine learning techniques and propose a solution to enhancing security. In fact, this method served to demonstrate the effect of machine learning algorithms on accelerating the security of e-transactions. In fact, higher security of e-transactions facilitates the customer trust in online trading.

In [3], it was pointed out that users have to make financial transactions with unknown persons in many parts of very large distributed financial systems. Any decisions made about the security of this type of trading significantly influences the trust in the entire system. The traditional solutions view this issue through client history. If positive, the customer will be trustworthy; otherwise, the customer will be untrustworthy. According to that very traditional method, this paper intended to outline a framework for specifying the trust level in unknown customers in very large systems through machine learning algorithms. For this purpose, a large number of features were extracted from transactions and user behaviors. Then, a trust predictive model was developed based on machine learning algorithms.

In [4] explored several methods of determining trust level in e-transactions. In other words, this paper explored the techniques specifying which variables are most important in measuring the trust level in e-business. The methods were assessed through the opinions of 5 subject-matter experts. Finally, the optimum method was introduced based on machine learning algorithms, i.e. k-nearest neighbor.

In [5], the machine learning algorithms were used to analyze the individual behavior within a loop, thereby to assess the trust level in e-business. The main challenging question to which an answer was provided in this study was: *Is the training set for machine learning algorithms selected randomly or objectively without any random process involved*? The results demonstrated that if the training set for machine learning algorithms involves a search for keywords and insertion of results, the accuracy of learning algorithms will improve.

Reference [6] first discussed the importance of trust in e-business, especially in cases where sellers/buyers are anonymous. Then, a model was proposed to predict the trust level of anonymous sellers/buyers through several features such as product type, transaction amount, transaction time etc. The new model involved two tree algorithms namely K-D-B and CMK. The results indicated that CMK was more accurate.

In [7], a new model was offered and investigated to predict the trust in counter party through Bayesian networks. For this purpose, a hierarchical model based on Bayesian networks was proposed to derive information about trust level in e-business. The newly proposed model is resistant to any potential noise or unwanted interference in different environments. The proposed model is based on statistical analyses derived from behavioral characteristics of counter party. Bayesian network is actually a machine learning model based on the theory of probability and Bayes' Theorem.

2 Research Structure

This paper was composed to first explore the theoretical foundations of machine learning algorithms, and then describe the new algorithm. Finally, the model was appropriately implemented and evaluated based on customer behavior data.

3 Proposed Method

This section initially defines and discusses the machine learning algorithms adopted for developing the predictive model of customer trust in e-business. The newly designed model will then be proposed through those machine learning algorithms. The next section will provide basic and fundamental elaborations on classification.

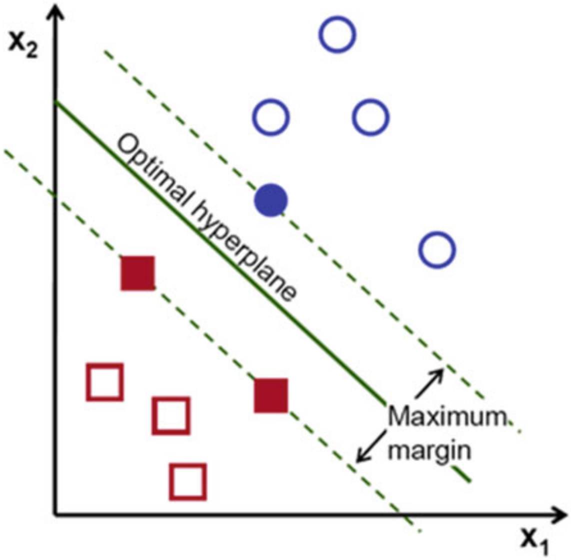
3.1 Main Issue of Classification

There are a number of predetermined classes and a number of input data. For instance, suppose the input data are fruit, and the classes are apples, peaches and pomegranates. The main objective in classification is to assign each input fruit to the relevant class with respect to its elements. Generally, the classification methods are divided in two categories: supervised and unsupervised. In the supervised method, the classifier is initially trained on a number of data whose class has already been specified. This dataset is called training dataset. Having finished the learning stage, the classifier is applied on the main dataset, i.e. testing data, for evaluation. In the unsupervised method, there are no training datasets, since the classifier classifies the input data according to its parameters. This section will elaborate on classification techniques of SVM, C4.5 Decision Tree, Bayesian Classifier, and K-Nearest Neighbors.

3.2 Support Vector Machine (SVM)

SVM is a supervised learning method used for classification and regression. It is among the relatively new techniques proving desirable performance in recent years compared to traditional classification techniques such as perceptron neural networks. The SVM classifier focuses on linear data classification by selecting the line with the greatest safety margin. The equation to find the optimal line for data is solved through combinatorial methods well-known in solving limited problems. Before line division, the machine transfers the data through the function ∅ into a far higher dimension space so as to be able to classify data with immense complexity. The high-dimension problems can be solved through these methods by adopting the Lagrangian dual theorem to transform the minimization problem into its dual form where the complex function ∅ involving a high-dimension space is replaced by a simpler function called *core*, which is the vector product of ∅. There are several core functions to be employed such as exponential, polynomial and Sigmoid [8]. Figure 1 shows the classification of support vector machine.

Fig. 1 Support vector machine algorithm [9]



3.3 C4.5 Decision Tree

A decision tree is a tool for supporting decisions adopting the trees for modeling. Decision trees are typically used in operational research, specifically in decision analysis, to identify a strategy most likely realizing an objective. Another application of decision trees is to describe the conditional probability calculations. In a decision analysis, a decision tree is used to depict the expected values of competitions alternately calculated. A decision tree has three types of nodes [10]: (a)—Decision node typically represented by a square. (b)—Chance node represented by a circle. (c)—End node represented by a triangle. The Fig. 2 shows how the algorithm functions as an example.

3.4 Bayesian Classifier

Also known as *Bayesian belief network*, a Bayesian network is a directed acyclic graph representing a set of independent random variables and their independent links. For instance, a Bayesian network can represent the relationship between cause of a disease and the disease itself. With the factors detected, the risk of a particular disease can be estimated in a patient.

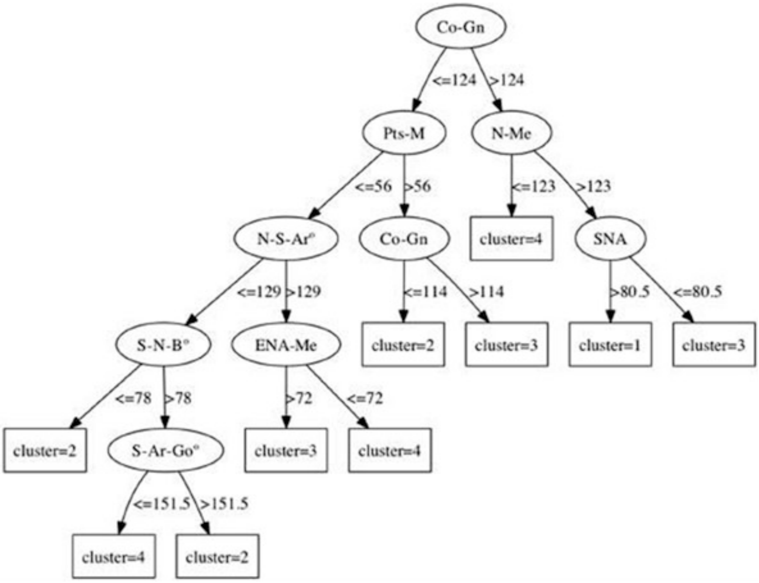


Fig. 2 Decision tree algorithm [11]

Bayesian network is a relatively new tool for identification of possible relationships in order to predict or assess the membership class. In short, it can be argued that a Bayesian network is a meaningful reflection of unknown relationships between parameters within a given field. The Bayesian network of a directed acyclic graph uses nodes and arcs to represent random variables and possible relationships between variables, respectively. Bayesian networks are widely used in probabilistic reasoning, where they transform into the tree connected to the reasoned probabilities. Bayesian networks transform into the main subgraph of the connected maximum tree and are used more frequently than connected trees. Bayesian networks are generally openly distributed with acceptable initial values and interrelationships between variables broadly used in real-world problems. Over the last few years, many scholars have explored the Bayesian networks and biology teams have adopted it in gene networks. Bayesian network is a graphical model to display the possibilities between variables. Moreover, Bayesian networks are a great way to depict the Continuous Probability Distributions exponentially and compressively, allowing for effective probabilistic calculations. They adopt the graphical model structure for the independent rules between chance variables. Bayesian networks are often used for probabilistic model scenarios, making it easier to reason under uncertainty. A Bayesian network includes a qualitative part (structural model) providing a visual representation of the interactions among variables and a quantitative part (a set of local probability details) allowed to deduce probabilities and make numerical measurements, affecting variables or a set of variables. The qualitative part involves unique Continuous Probability Distributions defined on all variables [12]. Figure 3 explains how this algorithm works as an example.

3.5 K-nearest Neighbor

Searching for the K-nearest neighbor will indicate the K neighbors closer to the query point. This method is usually employed in analysis/forecasting to estimate or classify a point based on the consensus of its neighbors. The K-nearest neighbor is a graph where each point on the graph is connected to its K-nearest neighbor. The near neighborhoods at a constant radius is a problem where all the members in a set of points within the Euclidean space should be effectively detected at a fixed distance from a given point. In this case, the data of structures should function on a fixed distance. The Fig. 4 is an example of running this classification [14].

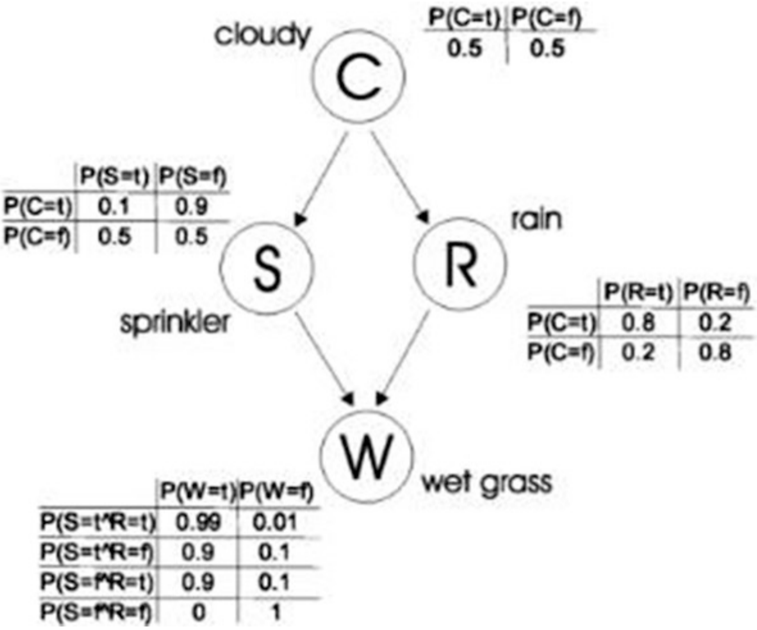


Fig. 3 Bayesian network algorithm [13]

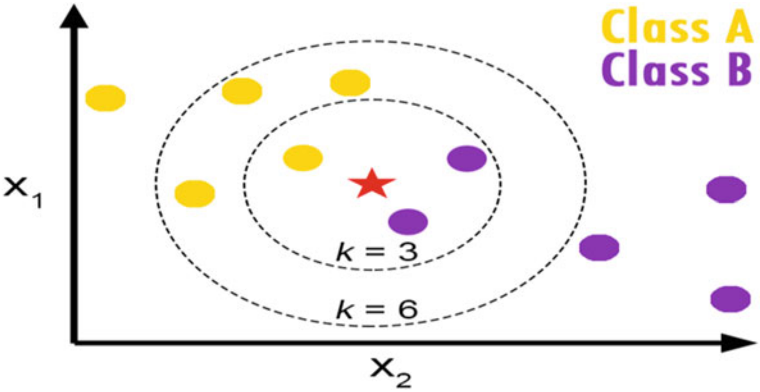


Fig. 4 Nearest neighbor algorithm

3.6 Distance Metric Learning

Assume that *S* is a set of similar data (i.e. behaviors of customers trusting in e-business) and D is a set of dissimilar data (i.e. customers distrustful of e-business model). Distance metric learning between the two datasets is obtained as follows:

$$\text{Min}\sum_{i,j\in S}\left\|x_i-x_j\right\|_A^2=\sum_{i,j\in S}\left(x_{ii}-x_j\right)^T A\left(x_i-x_j\right)\qquad \text{s.t}\quad \sum_{i,j\in D}\left\|x_i-x_j\right\|_A^2\geq\alpha \quad (1)$$

S: Similarity dataset; D: Dissimilarity dataset; α: Threshold for determining the degree of similarity or dissimilarity.
A: Distance Metric Matrix.
$A \geq 0$: A is a positive definite matrix, where all elements are greater than or equal to zero. That is because the distance value is positive.

Matrix A in the above equation is unknown and obtained by solving the above equation. Note that after solving the above equation and learning A from the training dataset, each new data is calculated through the distance metric or S data representing the distrustful customers. If the distance is short, it will be identified as a distrustful customer.

The above equation can be solved as follows: The logarithm of the condition expression is taken and then subtracted from the equation itself. The outcome will be:

$$g(A)=\sum_{i,j\in S}\left\|x_i-x_j\right\|_A^2-\log\left(\sum_{i,j\in D}\left\|x_i-x_j\right\|_A^2\right) \quad (2)$$

Then, this equation is solved through random initialization of A and using the gradient method.

$$A_{n+1}=A_n-\lambda H^{-1}\nabla \quad (3)$$

where H represents Hessian matrix and ∇ is the error gradient function g(A) [15].

The distance metric is used when the problem is extremely complex with multiple variables and when it is impossible to train the classifier accurately according to standard predefined criteria. In this case, instead of using the standard predefined distance metric, distance metric is trained by the training dataset, which is adopted in solving the classification problem.

4 Modeling of Decision-Making About Business Trust Using Machine Learning Algorithms

As mentioned in the introduction, an e-business procedure entails several features and factors below [16]:

(1) Business model: This feature represents an effective business model in e-trust, covering different values depending on whether the user is buyer/customer.

(2) Personal characteristics: These features reflect the individual characteristics (demographic) effective in e-business procedures such as age, gender, occupation and living place.
(3) Corporate characteristics: These features indicate the e-business trust of companies in e-business procedures such as e-trust certificates, warranties, etc.
(4) Infrastructure features: They reflect the infrastructure of e-business procedure such as natural and legal rules and technological infrastructure.

Table 1 displays these features along with their value ranges.

According to the current features in the system and their diversities, several machine learning algorithms were adopted to model the e-business customer trust system. Finally, the efficiency and accuracy of the algorithms will be compared to each other. In fact, this paper intended to model the customer decision-making process on trust/distrust in e-business through mathematical functions of machine learning algorithms. This could apply various factors and features within the decision-making, and ultimately extract the key parameters of customer trust.

The newly proposed system comprises the following overall 6 steps (Fig. 5).

The next section will explore the above mentioned machine learning algorithms as pseudocode for how the customer decision-making model on e-business trust learns. Cross-validation was used to determine the training dataset and testing dataset [17]. In this procedure, the entire dataset was first divided into *K* equal parts and the following steps were taken for *K* times: (1)—Selecting one of K sections not selected in the previous iterations as the testing dataset and selecting the K − 1 section left as the training dataset. (2)—This procedure is performed for all K sections.

The accuracy and efficiency of the proposed algorithm will be the mean accuracy and efficiency of K training and testing datasets. This method was used due to accidental elimination of training and testing datasets. In this procedure, all sections of the datasets are considered training and testing, so as to bring the efficiency and accuracy of the system closer to reality. Figure 6 demonstrates pseudocode of learning in customer decision model for e-business trust. As can be seen in the figure, this algorithm is based on support vector machine composed of two subroutines. In the first part, the training dataset and testing dataset are determined

Table 1 Variables and their definitions

Feature	Description	Values range
Business model	Buyer or seller	Company, wholesaler, retailer, industrial buyer, citizen, government, employee
Individual characteristics	Age, gender, occupation, living place	Male or female, young, middle-aged, old, student, self-employed, employee, provincial, capital city
Corporate features	Letters of guarantees	E-certificate, insurance, trust marks, trust seals
Infrastructure features	Technology and natural/legal rules	Website security, website ranking, security certificate

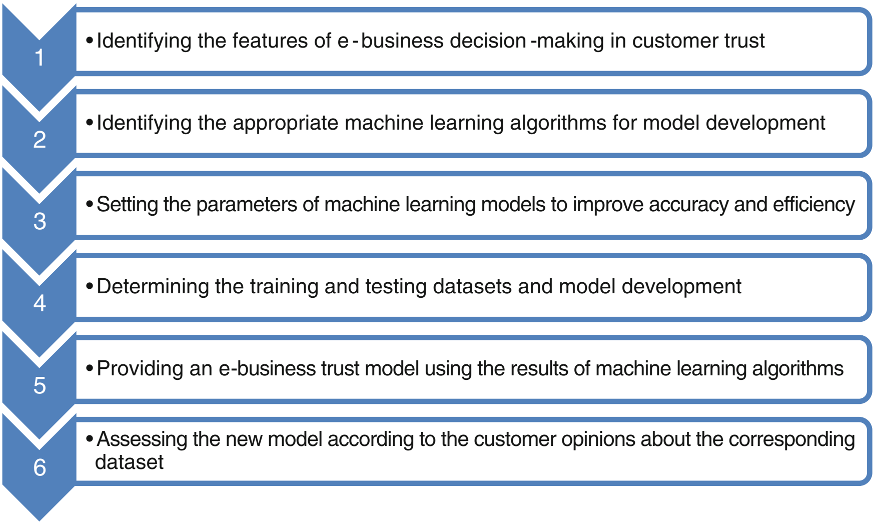
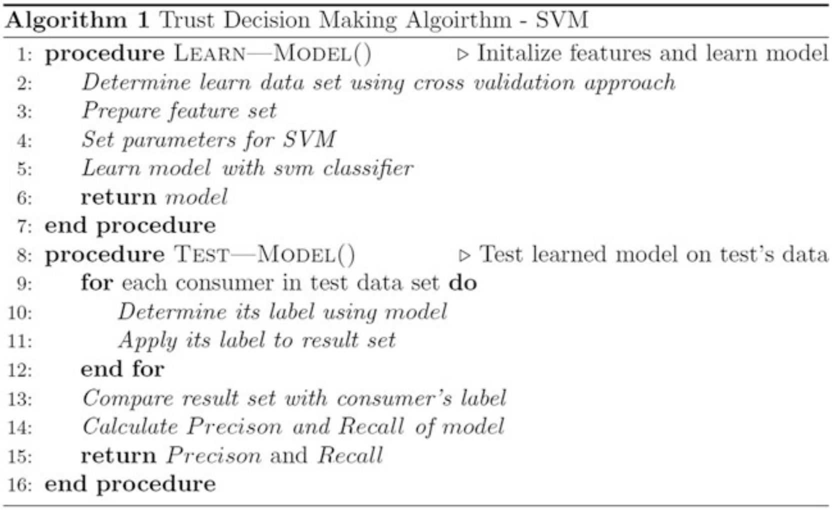

1
• Identifying the features of e - business decision -making in customer trust
2
• Identifying the appropriate machine learning algorithms for model development
3
• Setting the parameters of machine learning models to improve accuracy and efficiency
4
• Determining the training and testing datasets and model development
5
• Providing an e-business trust model using the results of machine learning algorithms
6
• Assessing the new model according to the customer opinions about the corresponding dataset

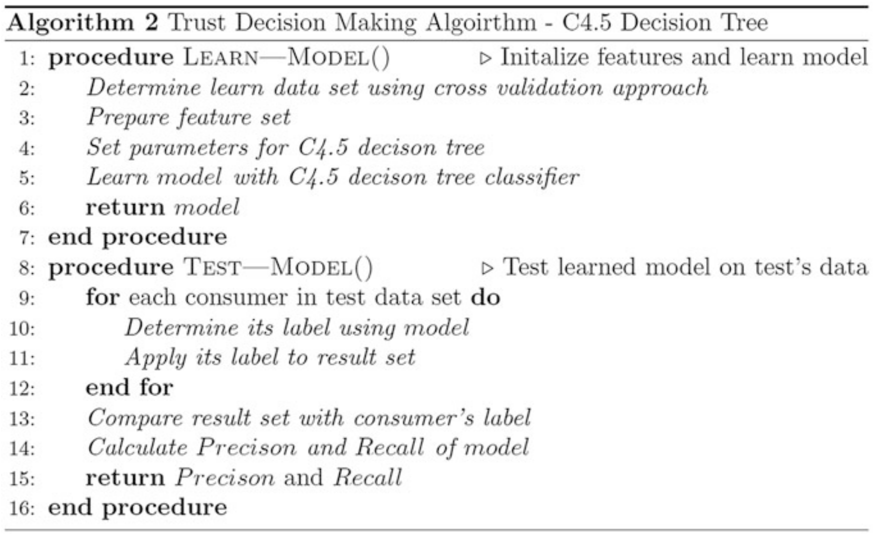
Fig. 5 Methodology of the newly proposed system



```
Algorithm 1 Trust Decision Making Algoirthm - SVM
procedure LEARN—MODEL()                         ▷ Initalize features and learn model
    Determine learn data set using cross validation approach
    Prepare feature set
    Set parameters for SVM
    Learn model with svm classifier
    return model
end procedure
procedure TEST—MODEL()                          ▷ Test learned model on test's data
    for each consumer in test data set do
        Determine its label using model
        Apply its label to result set
    end for
    Compare result set with consumer's label
    Calculate Precison and Recall of model
    return Precison and Recall
end procedure
```

Fig. 6 Pseudocode of support vector machine algorithm for learning of customer decision model in e-business trust

based on cross-validation. After specifying the parameters of support vector machine, the new model was trained through the algorithm. In the second subroutine, the model trained in the previous stage predicts the type of trust/distrust in e-business for each customer depending on their characteristics mentioned in the previous section.



```
Algorithm 2 Trust Decision Making Algoirthm - C4.5 Decision Tree
procedure LEARN—MODEL()                ▷ Initalize features and learn model
    Determine learn data set using cross validation approach
    Prepare feature set
    Set parameters for C4.5 decison tree
    Learn model with C4.5 decison tree classifier
    return model
end procedure
procedure TEST—MODEL()                 ▷ Test learned model on test's data
    for each consumer in test data set do
        Determine its label using model
        Apply its label to result set
    end for
    Compare result set with consumer's label
    Calculate Precison and Recall of model
    return Precison and Recall
end procedure
```

Fig. 7 Pseudocode of decision tree algorithm for learning of customer decision model in e-business trust

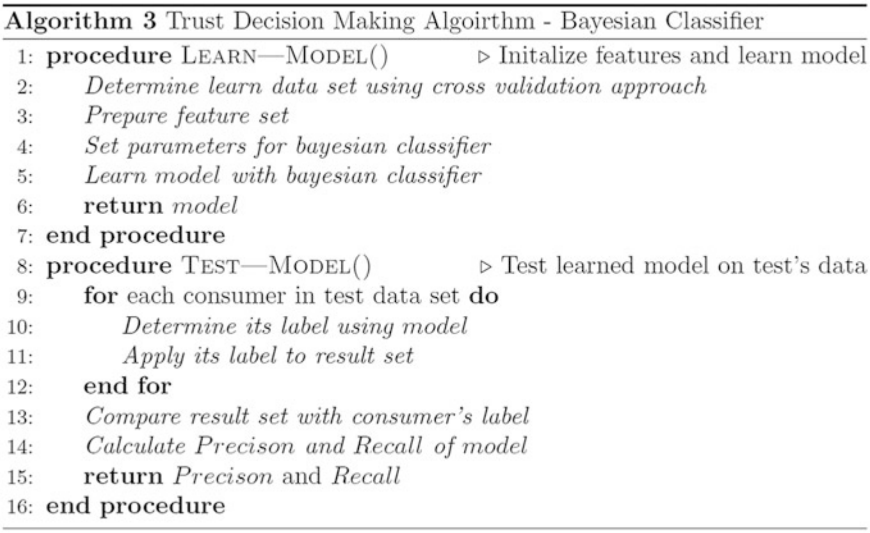
Figure 7 illustrates the pseudo-code for learning of customer decision model about trust in e-business using C4.5 decision tree algorithm. Similar to support vector machine, C4.5 decision tree is composed of two main parts, the first of which specifies the testing and training datasets and the values of parameters in the decision tree algorithm. The second part uses the model trained in the first part to predict trust/distrust in the e-business model for each customer.

Figures 8, 9 and 10 display the pseudocode of learning algorithm for the customer decision model about trust in e-business using the Bayesian classifier algorithms, k-nearest neighbor and distance metric, respectively. This algorithm resembles the support vector machine in terms of overall procedure.

5 Model Implementation and Presentation

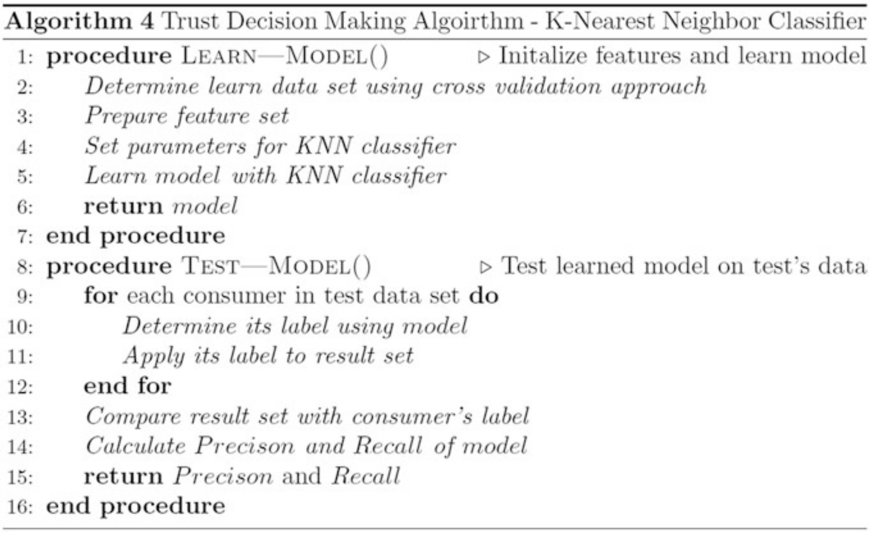
This section introduces the implementation of new algorithms, compares them against the results of users' behavior about the e-business trust model and evaluates the results.

The dataset consisted of comments from 3440 individuals about e-business trust obtained through 6 questionnaires. Each questionnaire focused on a key issue concerning the e-business model, where the collected data were used through cross-validation to specify the training and testing datasets. After implementing the newly designed machine learning algorithm, the most important factors in



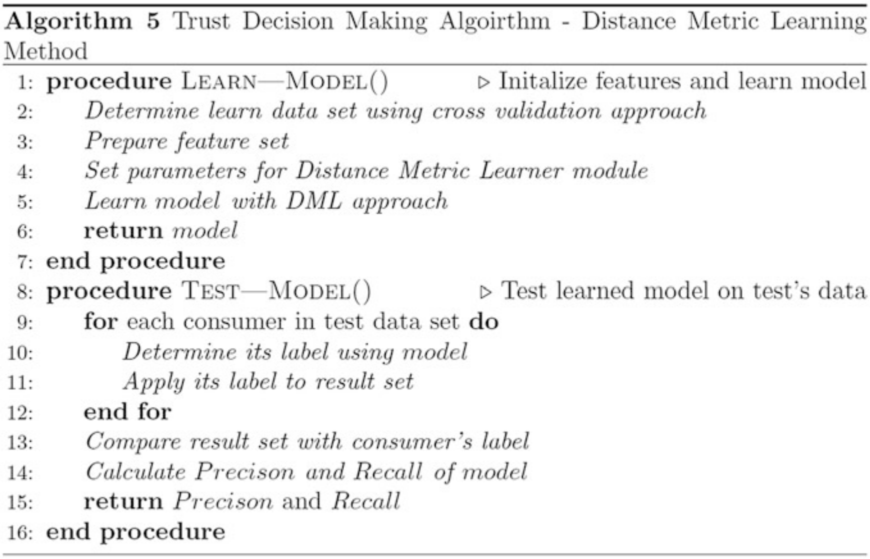
```
Algorithm 3 Trust Decision Making Algoirthm - Bayesian Classifier
procedure LEARN—MODEL()            ▷ Initalize features and learn model
    Determine learn data set using cross validation approach
    Prepare feature set
    Set parameters for bayesian classifier
    Learn model with bayesian classifier
    return model
end procedure
procedure TEST—MODEL()             ▷ Test learned model on test's data
    for each consumer in test data set do
        Determine its label using model
        Apply its label to result set
    end for
    Compare result set with consumer's label
    Calculate Precison and Recall of model
    return Precison and Recall
end procedure
```

Fig. 8 Pseudocode of Bayesian classifier algorithm for learning of customer decision model in e-business trust



```
Algorithm 4 Trust Decision Making Algoirthm - K-Nearest Neighbor Classifier
procedure LEARN—MODEL()            ▷ Initalize features and learn model
    Determine learn data set using cross validation approach
    Prepare feature set
    Set parameters for KNN classifier
    Learn model with KNN classifier
    return model
end procedure
procedure TEST—MODEL()             ▷ Test learned model on test's data
    for each consumer in test data set do
        Determine its label using model
        Apply its label to result set
    end for
    Compare result set with consumer's label
    Calculate Precison and Recall of model
    return Precison and Recall
end procedure
```

Fig. 9 Pseudocode of k-nearest-neighbor algorithm for learning of customer decision model in e-business trust



```
Algorithm 5 Trust Decision Making Algoirthm - Distance Metric Learning Method
procedure LEARN—MODEL()                          ▷ Initalize features and learn model
    Determine learn data set using cross validation approach
    Prepare feature set
    Set parameters for Distance Metric Learner module
    Learn model with DML approach
    return model
end procedure
procedure TEST—MODEL()                           ▷ Test learned model on test's data
    for each consumer in test data set do
        Determine its label using model
        Apply its label to result set
    end for
    Compare result set with consumer's label
    Calculate Precison and Recall of model
    return Precison and Recall
end procedure
```

Fig. 10 Pseudocode of distance metric algorithm for learning of customer decision model in e-business trust

decision-making on e-business and key points of opinions from 3440 individuals were extracted and then a conceptual model was developed accordingly.

RapidMiner was employed to implement the machine learning algorithm used to develop the predictive and decision model about customer trust in e-business [18].

Table 2 lists the features defined for customer decision model learning.

This section provides the results of each machine learning algorithm.

Figure 11 shows the output of support vector machine. This figure displays the value of cost function used in support vector machine as an optimization factor and correct assignment of classes to data.

Table 2 List of dataset characteristics

Feature	Description	Feature	
F1	Buyer or seller	F6	E-trust certificate
F2	Company, wholesaler, retailer, industrial buyer, citizen, government, employee	F7	LoG, insurance
F3	Gender	F8	Website ranking
F4	Age	F9	Website security
F5	Occupation		

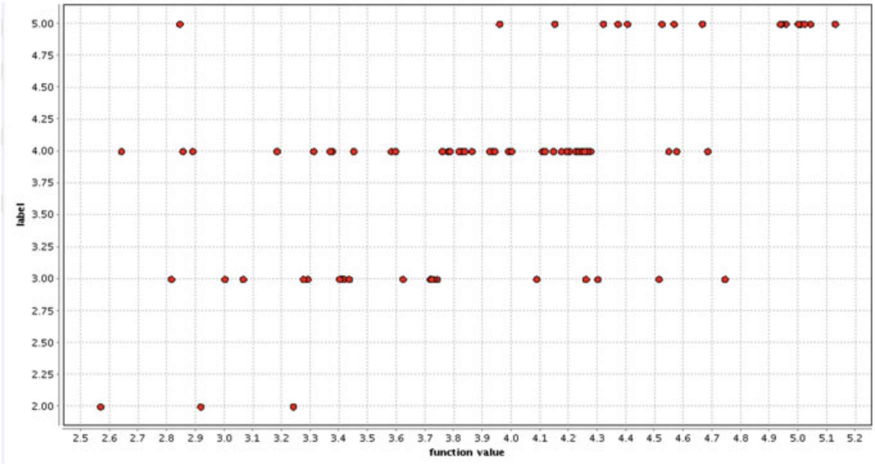


Fig. 11 The output of support vector machine algorithm (values of the cost function for labels)

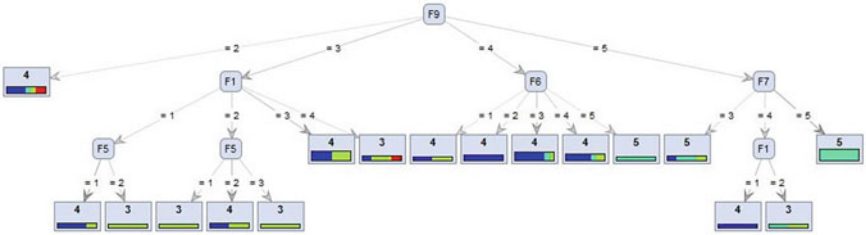


Fig. 12 Overview of the decision tree algorithm

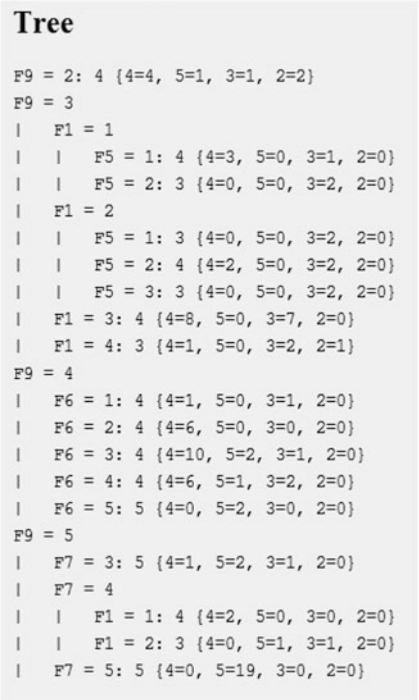
As indicated in the figure, the labels consist of five values. There are 5 different classes as follows:

1. Trust: Customers with complete trust in the business model.
2. Tendency toward trust: Customers tending to trust in the business model.
3. Neutral: Customers who neither trust nor distrust in the business model.
4. Tendency toward distrust: Customers tending to distrust in the business model.
5. Distrust: Customers with complete distrust in the business model.

Figure 12 illustrates the output of decision tree algorithm along with the various features given their importance at different levels. Figure 13 shows the same algorithm in script.

Figure 14 shows the density of a feature sample such as F1. The values of different densities are used in Bayesian classifier.

Fig. 13 Script output of the decision tree algorithm



Tree

```
F9 = 2: 4 {4=4, 5=1, 3=1, 2=2}
F9 = 3
|   F1 = 1
|   |   F5 = 1: 4 {4=3, 5=0, 3=1, 2=0}
|   |   F5 = 2: 3 {4=0, 5=0, 3=2, 2=0}
|   F1 = 2
|   |   F5 = 1: 3 {4=0, 5=0, 3=2, 2=0}
|   |   F5 = 2: 4 {4=2, 5=0, 3=2, 2=0}
|   |   F5 = 3: 3 {4=0, 5=0, 3=2, 2=0}
|   F1 = 3: 4 {4=8, 5=0, 3=7, 2=0}
|   F1 = 4: 3 {4=1, 5=0, 3=2, 2=1}
F9 = 4
|   F6 = 1: 4 {4=1, 5=0, 3=1, 2=0}
|   F6 = 2: 4 {4=6, 5=0, 3=0, 2=0}
|   F6 = 3: 4 {4=10, 5=2, 3=1, 2=0}
|   F6 = 4: 4 {4=6, 5=1, 3=2, 2=0}
|   F6 = 5: 5 {4=0, 5=2, 3=0, 2=0}
F9 = 5
|   F7 = 3: 5 {4=1, 5=2, 3=1, 2=0}
|   F7 = 4
|   |   F1 = 1: 4 {4=2, 5=0, 3=0, 2=0}
|   |   F1 = 2: 3 {4=0, 5=1, 3=1, 2=0}
|   F7 = 5: 5 {4=0, 5=19, 3=0, 2=0}
```

As explained in the previous section, the two parameters of Recall and Precision are used for comparison and evaluation of new algorithms. This section discusses the parameters.

It should be noted that TP stands for True Positive, TN for True Negative, FP for False Positive and FN for False Negative. Each term is clarified below:

TP: The correct values also detected correctly by the algorithm.
TN: The incorrect values detected correctly by the algorithm.
FP: The correct values detected incorrectly by the algorithm.
FN: The incorrect values detected incorrectly by the algorithm.

Calculation of precision:

$$Precision = \frac{number\ of\ true\ positives}{number\ of\ true\ positives + number\ of\ false\ positives} \tag{4}$$

Calculation of recall:

$$Recall = \frac{number\ of\ true\ positives}{number\ of\ true\ positives + number\ of\ false\ negatives} \tag{5}$$

Fig. 14 Density of characteristic F1 (Bayesian classification algorithm)

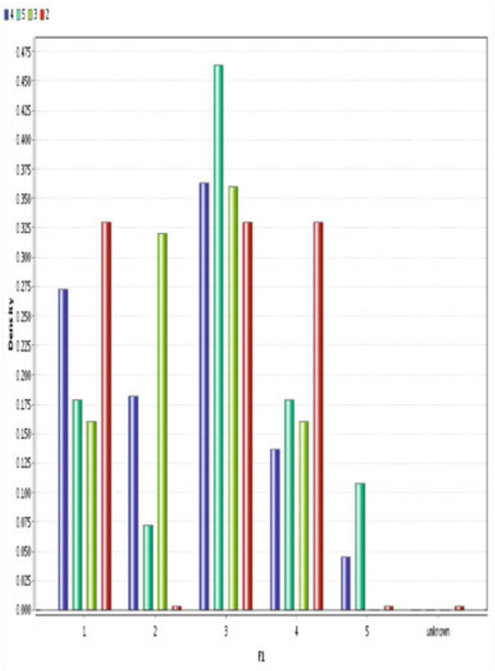


Table 3 Results of different algorithms

Name of algorithm	Precision (%)	Recall (%)
Support vector machine (SVM)	84.97	83.9
Decision tree (C4.5)	72.64	69.79
Nave Bayesian	75.21	54.21
K-NN	72.19	60.73
Distance metric learning (DML)	81.67	79.94

Table 3 lists the values of Precision and Recall for support vector machine, decision tree, Bayesian classifier, K-nearest neighbor and distance metric.

Figure 15 displays the values graphically.

As can be seen in the figure, the SVM outperformed the other machine learning algorithms. It can be used as the main algorithm for developing the customer decision model about trust in e-business.

Therefore, the support vector machine algorithm can be employed to train the customer feedback in the decision model as illustrated in Fig. 16. Evidently, the trained model rather concentrates on the corporate and infrastructure variables. In other words, variables F1, F2, F6, F7, F8, F9 play a decisive role in identification of the customer decision model.

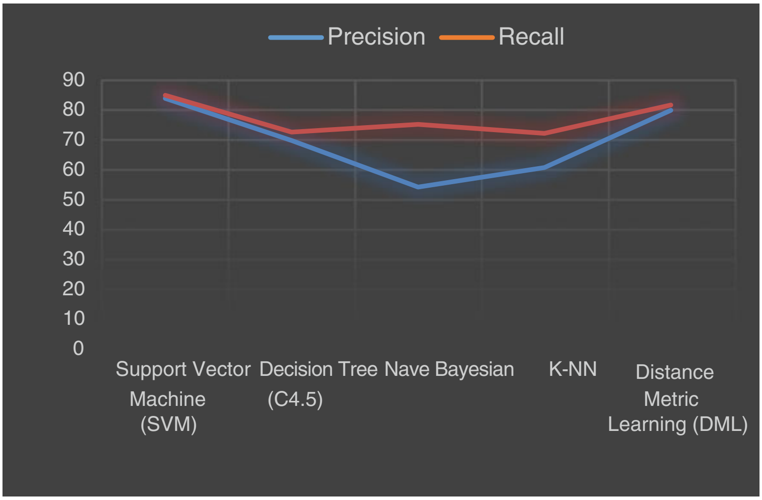


Fig. 15 The results of different machine learning algorithms

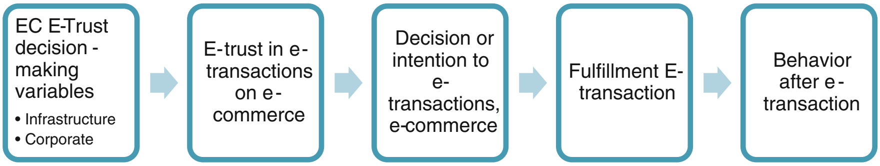

EC E-Trust decision - making variables
• Infrastructure
• Corporate
E-trust in e-transactions on e-commerce
Decision or intention to e-transactions, e-commerce
Fulfillment E-transaction
Behavior after e-transaction

Fig. 16 Conceptual model of decision-making about e-business trust using machine learning algorithms (SVM)

6 Conclusions

According to the results obtained in this study, the trust of public customers in e-business can be enhanced and promoted throughout the society by monitoring agencies who provide the mutual security and trust in a transaction under an accurate scientific framework. As is clear in results, security certificates, e-trust symbols and website security play utmost roles in realizing the customer trust in e-transactions. In addition to the symbols and certificates, efforts should be made to design a comprehensive monitoring and evaluation mechanism where every factor such as banking transaction security, web security, site ranking, etc. are covered, so that customers make online banking transactions more confidently.

References

1. Ravi V, Carr M, Sagar NV (2007) Profiling of internet banking users in India using intelligent techniques. J Serv Res 7(1):61
2. Pitroda SG, Desai M (2015) Facilitating establishing trust for a conducting direct secure electronic transactions between a user and a financial service providers. Google Patents
3. Liu X, Tredan G, Datta A (2014) A generic trust framework for large-scale open systems using machine learning. Comput Intell 30(4):700–721
4. Liébana-Cabanillas F et al (2013) Analysing user trust in electronic banking using data mining methods. Expert Syst Appl 40(14):5439–5447
5. Cormack GV, Grossman MR (2014) Evaluation of machine-learning protocols for technology-assisted review in electronic discovery. In: Proceedings of the 37th international ACM SIGIR conference on research and development in information retrieval, ACM
6. Zhang H et al (2015) ReputationPro: The efficient approaches to contextual transaction trust computation in E-commerce environments. ACM Trans Web (TWEB) 9(1):2
7. Teacy WL et al (2012) An efficient and versatile approach to trust and reputation using hierarchical Bayesian modelling. Artif Intell 193:149–185
8. Cortes C, Vapnik V (1995) Support vector machine. Mach Learn 20(3):273–297
9. Meyer D, Wien FT (2015) Support vector machines. The interface to LIBSVM in package e1071
10. Quinlan JR (1996) Bagging, boosting, and C4.5. In: AAAI/IAAI, vol 1
11. Rokach L, Maimon O (2014) Data mining with decision trees: theory and applications. World Scientific
12. Cheng J, Greiner R (1999) Comparing Bayesian network classifiers. In Proceedings of the fifteenth conference on uncertainty in artificial intelligence, Morgan Kaufmann Publishers Inc
13. Friedman N, Geiger D, Goldszmidt M (1997) Bayesian network classifiers. Mach Learn 29(2–3):131–163
14. Peterson LE (2009) K-nearest neighbor. Scholarpedia 4(2):1883
15. Xing EP et al (2003) Distance metric learning with application to clustering with side-information. Adv Neural Inf Process Syst 15:505–512
16. Khodadad HSH, Shirkhodayee M, Keronaich A (2009) The factors affecting customer trust in ecommerce (B2C model). Q Teach Humanit 13(2)
17. Moreno-Torres JG, Sáez JA, Herrera F (2012) Study on the impact of partition-induced dataset shift on-fold cross-validation. IEEE Trans Neural Netw Learn Syst 23(8):1304–1312
18. Hofmann M, Klinkenberg R (2013) RapidMiner: data mining use cases and business analytics applications. CRC Press

Three-Band, Flexible, Wearable Antenna with Circular Polarization

Milad Najjariani and Pejman Rezaei

Abstract In this chapter, we show a circular polarization (CP), wearable antenna, with a flexible substrate for GPS, DCS, and PCS applications. The CP was created using different methods, for example, we truncated two opposite edges of a square patch, as well as using other methods. We used a coplanar waveguide for feeding and the substrate material antenna is Rogers RO4003. The dimensions of the proposed antenna is 76 × 76 mm, with a thickness of 0.8 mm, a relative permittivity of 3.55, and dielectric loss tangent of 0.0027. The chapter goes on to show that the CP, wearable antenna has a 3-dB axial ratio (AR) bandwidth of 520 MHz (26%), a return loss (RL) of 10 dB, and impedance bandwidth of 2.9 GHz (97%). The antenna was simulated by HFSS (High Frequency Electromagnetic Field) software.

Keywords Coplanar waveguide (CPW) · Wearable antenna · Circular polarization (CP) · Axial ratio (AR)

1 Introduction

A microstrip antenna consist of patch, Ground and substrate, that patch putting on the substrate and Ground is under that [1]. usually used different method for feeding microstrip antenna, microstrip feedline, coaxial feed and coplanar waveguide (CPW) are important of them. but in this paper used cpw for antenna feeding.

The radiation characteristic is one of the most important features of an antenna. In recent years, the wearable antenna has become an interesting topic for

M. Najjariani (✉)
Department of Electrical and Computer Engineering,
Adiban Institute of Higher Education, Garmsar, Iran
e-mail: milad.najjariani@gmail.com

P. Rezaei
Electrical and Computer Engineering Faculty, Semnan University, Semnan, Iran
e-mail: prezaei@semnan.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_75

researchers. One of the best ideas associated with this is the integration of cloth and the antenna to make smart clothes [2–4].

The wearable antenna has applications for the military, navigation, global positioning systems, firefighting, emergency response, medical treatments, and intelligent systems, along with other applications. However, use of this antenna has specifically gained increased interest in terms of its potential military applications [5, 6]. Textiles and flexible antennas are good materials for making of intelligent cloth, because they are light, cheap, have a low profile, and are able to have within them microwave circuits [7–9]. Recently, there has been a lot of research and development in wearable antennas installed on the body [10, 11].

In this chapter, an L-shaped wearable antenna is presented, with CPW feed and wide bandwidth, for wearable applications. Circular polarization (CP) is preferred to linear polarization because it has multiple benefits. we should making difference in radiation patch till have circular polarozation. we should making difference in radiation patch until the microstip patch has a circular polarization. In fact should be 90° Phase difference in current distributions for have cp in patch antenna [12–14].

The proposed antenna is designed to cover standard GPS, PCS, and DCS bands while its frequency range for each is 1.575, 1.710–1.880, and 1.850–1.990 GHz, respectively. The structure of a wearable antenna with CP is planar and uses a flexible substrate and conductor for patch and GND plate. The material used for the designed antenna was provided by the Rogers Corporation, a material which has multiple features [15, 16]. The antenna has a thickness of 0.8, a relative permittivity of 3.55, and a dielectric loss tangent of 0.0027, representing a thin substrate.

2 Introduction to the Structure of the Designed Antenna

The geometry of the proposed antenna is shown in Fig. 1. The antenna is made from a Rogers RT4003 substrate with dimensions of 67 × 67 mm, a thickness of 0.5 mm, and a relative permeability constant of 3.55. The structure of this antenna

Fig. 1 Dimensions of the proposed antenna

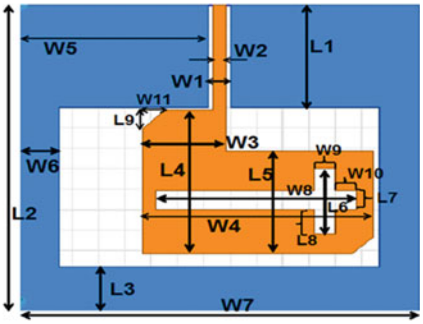


Table 1 Dimensions of the proposed antenna shown graphically in Fig. 1 (all measurements are given in millimeters)

W_1	W_2	W_3	W_4
6	2.6	16	44
W_5	W_6	W_7	L_1
36	7.5	76	25
L_2	L_3	L_4	L_5
76	12	35	25
W_8	W_9	W_{10}	W_{11}
38	4	4	4.5
L_6	L_7	L_8	L_9
16	4.6	6	4.5

is a surface-waveguide-type with planar feeding which matching of that is 50 ohms. The structure of the antenna is composed of a ground plan having square rings and a radiation patch which is L shaped while it has two truncate triangle in opposite of them and cross slot In the middle of patch which causes CP, improvements in the impedance and polarization bandwidth and passing standard GPS, DCS, and PCS bands the current distribution Rotated in the patch antenna and 90° phase deference in the phase and eventually Circular polarization is caused in Table 1. Table 1 presents the dimensions and sizes of the slots for the designed antenna.

3 Results of Antenna Simulation

In the analysis of the antenna, the simulation results of the return loss diagrams, axial ratios, radiation patterns, right and left gain patterns, vector current distributions, and 3-D patterns are investigated.

The return loss, the return wave caused by discontinuity in the microstrip antenna port, has an optimal value of less than −10 dB. The axial ratio diagram indicates the presence of CP which has an optimal value of less than −3 dB.

In Fig. 2a the geometry of the proposed antenna is shown. The proposed antenna structure is a CPW so that the ground plane and radiation patch appear on one plate. As seen in the base design, this antenna consists of a ground plane, which has a rectangular slot in the middle of it, and a radiation patch that includes a rectangle with a thin rectangular coplanar waveguide feed line. In the first step, as shown in Fig. 2b the radiation patch structure of the designed antenna is L shaped. In the second step (Fig. 2c), a cross-shaped slot has been created in the middle of the radiation patch. Finally, in the third step (Fig. 2d), two opposite corners of the radiation patch have triangles cut from them. Test results for the antenna are given in the next section.

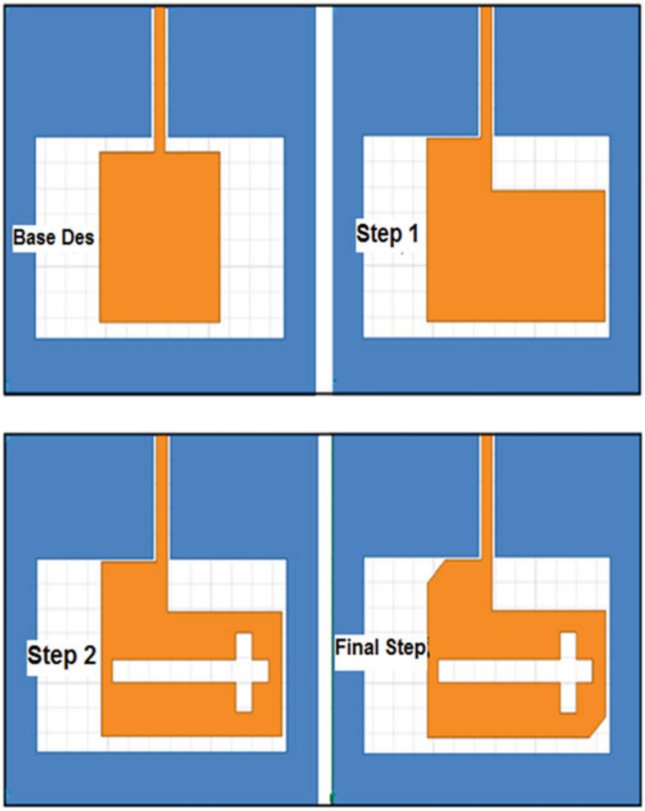


Fig. 2 Step-by-step design of the proposed antenna

3.1 Return Loss

In microstrip antenna one can reduce their size and cost by using special techniques using CP with only one feed line. In actual fact, there is no need to use two feeds with 90° positional differences. This can be achieved, with a little change in the microstrip antenna patch. The following figures are suitable for creating CP with a feed because it causes a rotation of current in the patch surface. Figure 3 is a return loss diagram. As can be seen in the return loss diagram for the base design, within the range 2.65–3 GHz, and higher than 4 GHz, the diagram drops below −10 dB (pass band). In the proposed antenna, because the range greater than 3 GHz was not important to us it has not shown in the figure. The impedance bandwidth of the basic design for this range is around 350 MHz. However, in the first step, by changing the shape of the radiation patch from rectangular to L shaped, the return loss bandwidth drops under −10 dB between 1.33–1.66 GHz and 1.94–2.15 GHz. Its bandwidth is around 535 MHz. In the second step, by creating a cross slot in the

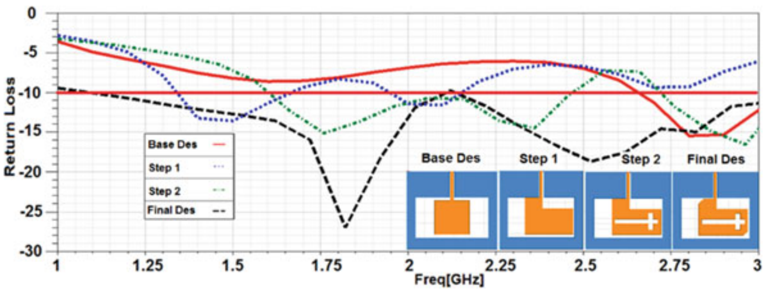


Fig. 3 Step-by-step return loss for the proposed antenna

middle of the radiation patch, that diagram drops under −10 dB in range 1.6–2.46 GHz and also 2.7 GHz. Which has bandwidth around 1.16 GHz which band width achieve to 1.6 GHz. In the third step, by creating two opposite slits in the corners of the radiation patch, the range of drops under −10 dB from 1.1–3 GHz. Only at 2.2 GHz does it reach the edge of the −10 dB range. which it is negligible. which band width achieve to 2.9 GHz. In this case, the standard GPS, DCS, and PCS bands are covered. And also significant that analysis range is than 1–3 GHz.

3.2 Axial Ratio

Figure 4 shows the axial ratio diagram for the designed antenna. As seen in the axial ratio diagram for the base design, within the frequency range 1.84–1.95 GHz the axial ratio drops below −3 dB, which has a bandwidth around 110 MHz which

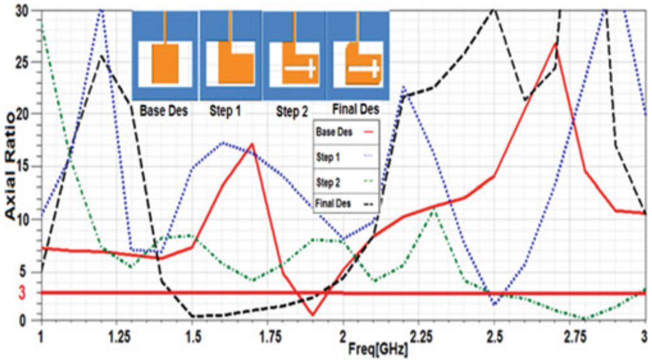


Fig. 4 Step-by-step diagram of the axial ratio for the proposed antenna

covers the PCS band. However, in the GPS range the diagram is about 7 dB and it only covers a part of the DCS band. On the other hand, to have an acceptable CP, the diagram should go under −3 dB in the Within range. According to this fact, in the base design the antenna is not desirable for CP. However, in the first step, by changing the shape of the radiation patch from rectangular to L shaped, the axial ratio drops under −3 dB within the range 2.46–2.53 GHz, which has a bandwidth of around 50 MHz and covers the Bluetooth band. Additionally, the diagram shows that for the desired bands is higher than −10 dB. In this step, changing the L-shaped patch causes a 90-degree phase difference in the current distribution phase. Due to its geometry the current distribution in the patch has a circular rotation, which suggests CP. In the second step, a cross shape in the radiation patch has been created. The range of this goes to −3 dB in the range 2.49–3 GHz. With a bandwidth of around 510 MHz, the effect of changes made at this step provides evidence for the creation of a polarization bandwidth. But the diagram in the limit of the bands are below −10 dB and the frequency shift to lower frequencies is required. Finally, in the third step, the opposite two corners of the radiation patch are truncating, in the form of triangles. The range of this diagram is between 1.42 and 1.94 GHz when it drops below −3 dB, having a bandwidth of around 520 MHz. As seen in the final result, by creating a slit at the corner of the patch the optimal status of the diagram has not changed, but it has shifted to a lower frequency. The result is coverage of GPS bands, DCS bands, and the lower, and the majority of the upper, PCS bands.

Finally, it should be said that the creation of a triangular slit in the corner of the patch, as described earlier, has a significant effect on the final result. Also, it should be mentioned that the axial ratio diagram is more sensitive to variation of the radiation patch. The analytical range is from 1 to 3 GHz.

3.3 Current Distribution on the Antenna Surface

Figure 5 shows the current distribution in the designed antenna in vector form. As shown earlier, the rotation of the current distribution is due to CP. This is displayed in the figure [17]. The current distribution is shown for 0, 90, 180, and 270 phases at a frequency of 1.6 GHz.

It can be seen that in the 0-degree phase, the direction of the current (direction of the vectors) is to the right; in the 90-degree phase it is upward; in the 180-degree phase it is to the left; and finally in the 270-degree phase it is downward—representing right-handed CP.

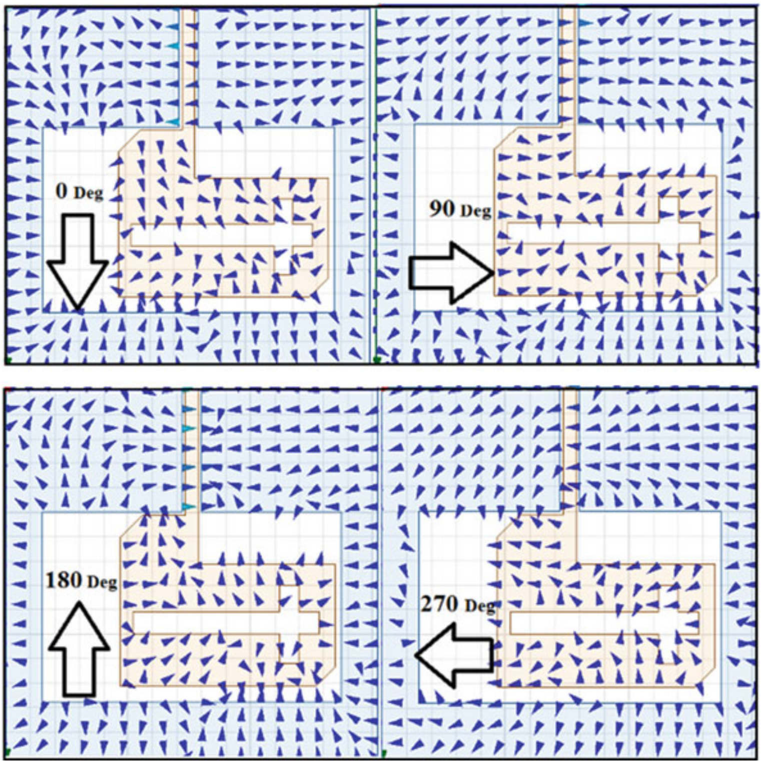


Fig. 5 Vector current distribution for the designed antenna

3.4 *Right-Handed and Left-Handed Radiation Patterns*

Figure 6 shows the right-handed and left-handed radiation patterns from the designed antenna at 0°, 90°, 180°, and 270°, at a frequency of 1.6 GHz. As seen in the figure, in the 0-degree phase, the greatest gain follows a right-handed pattern with a value of 2.5 dB; in the 90-degree phase, the greatest gain follows a right-handed pattern with a value of 7.5 dB; in the 180-degree phase, the greatest gain follows a right-handed pattern with a value of 2.5 dB; and finally, in the 270-degree phase, a left-handed pattern has the highest gain of 5.7 dB.

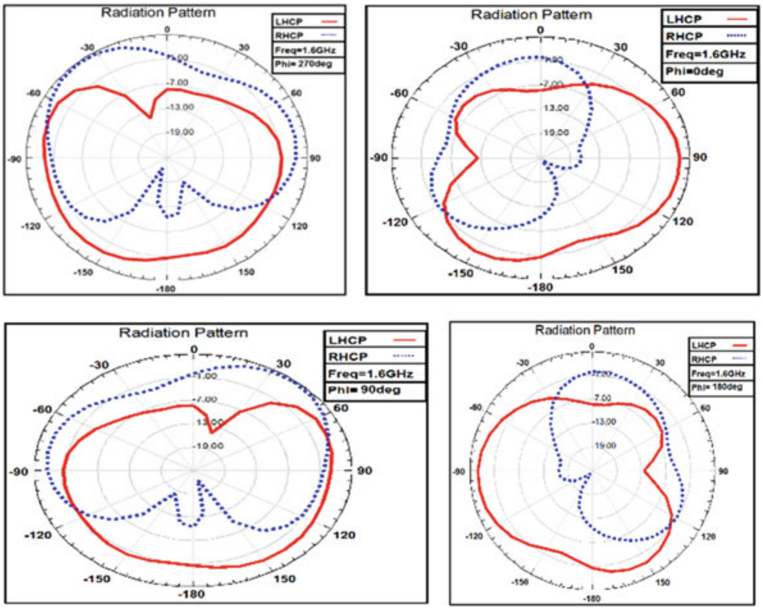
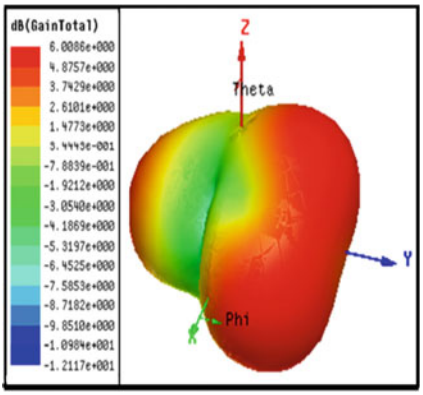


Fig. 6 The right-handed and left-handed radiation patterns of the designed antenna. *Notes* RHCP, right-handed circular polarization; LHCP, left-handed circular polarization

3.5 3-D Radiation Pattern

Figure 7 shows the 3-D radiation pattern of the designed antenna.

Fig. 7 3-D radiation pattern



4 Conclusion

This chapter presents a CP wearable antenna of size 67 × 67 mm. It is designed on Rogers Corporation RT4003 substrate with a CPW structure. This antenna is designed for use in the GPS, DCS, and PCS bands. The proposed antenna consists of an L-shaped radiation patch. In the middle of it there is a cross-shaped slot, and it has two opposite truncate triangular slots at the edge of the patch. The ground also has a slot and has CP in GPS, DCS, and PCS. The designed antenna has an impedance bandwidth of around 2.9 GHz and a polarization bandwidth of about 520 MHz.

References

1. Locher I et al (2006) Design and characterization of purely textile patch antenna. Adv Packaging, IEEE Trans on 777–788
2. Jais MI, Jamlos MF, Malek MF, Jusoh M (2012) Conductive E-textile analysis of 1.575 GHz rectangular antenna with H-slot for GPS application. In: Antenna and propagation conference, Loughborough, UK
3. Lui KW, Murphy OH, Toumazou C (2013) A wearable wideband circularly polarized textile antenna for effective power transmission on a wirelessly-powered sensor platform. IEEE Trans Antennas Propag 61(7):3873–3876
4. Salam A, Khan AA, Hussain MS (2014) Dual band microstrip antenna for wearable applications. Microwave Opt Technol Lett 56(4):916–918
5. Hertleer C, Rogier H, Vallozzi L, Langenhove LV (2009) A textile antenna for off-body communication integrated into protective clothing for fire-fighters. IEEE Trans Antennas Propag 57(4):919–925
6. Kaivanto EK, Berg M, Salonen E, Maagt PD (2011) Wearable circularly polarized antenna for personal satellite communication and navigation. IEEE Trans Antennas Propag 59(12):4490–4496
7. Massey PE (2001) GSM fabric antenna for mobile phones integrated within clothings. IEEE, vol 3, pp 8–13
8. Sankarlingam S, Gupta B (2010) Development of textile antennas for body wearable application investigations on their performance under bent conditions. Progr Electromagnetic ResB 22:53–71
9. Klenm M, Troester G (2006) Textiles transaction on antennas and propagation 54(11):414–422
10. Zhang L, Wang Z, Volakis JL (2013) Textile antennas and sensors for body-worn applications. IEEE Antennas Wirel Propag Lett 11:1690–1693
11. Wang H, Zhang Z, Li Y, Feng Z (2013) A dual-resonant shorted patch antenna for wearable application in 430 MHz band. IEEE Trans Antennas Propag 61(12):6195–6200
12. Vallozzi L, Van Torre P, Hertleer C, Rogier H (2009) A textile antenna for off-body communication integrated into protective clothing for firefighters. IEEE Trans Antennas Propag 57(4):919–925
13. Kaivanto E, Lilja J, Berg M, Salonen E, Salonen P (2010) Circularly polarized textile antenna for personal satellite communication. In: European conference *on* antennas and propagation, pp 1–4

14. Klemm M, Locher I, Troster G (2004) A novel circularly polarized textile antenna for wearable applications. In: Proceeding 7th European Conference Wireless Technologies pp 285–288
15. Locher I, et al (2006) Design and characterization of purely textile patch antenna. Adv Packaging, IEEE Trans 777–788
16. Rais NHM, Son PJ, Malek F, Ahmad S, Hashim NBM, Hall PS (2009) A Review of wearable antenna. In: Antennas and propagation conference, pp 225–228
17. Najjariani M, Rezae P (2016) The slot wearable microstrip antenna with circular polarization for Use in GPS, bluetooth bands. In: International conference in electrical and mechanic engineering

A Multi-objective Distribution Network Reconfiguration and Optimal Use of Distributed Generation Unites by Harmony Search Algorithm

Mojtaba Mohammadpoor, Reza Ranjkeshan and Abbas Mehdizadeh

Abstract In this paper, the method of network reconfiguration and simultaneous use of distributed generation resources (DG) in optimal location and capacity is analyzed in order to minimize losses and reach the optimum level of voltage stability and voltage profile. A new method for this purpose is proposed by use of the Harmonic Search Algorithm (HSA). To investigate the effectiveness of the proposed method, the capabilities of the MATLAB software and the DPL language linked to the DIGSILENT application is used. The 33-node distribution network of IEEE standard was selected for investigation. The simulation results shows that by using the proposed method, network losses were minimized and the voltage level and voltage profiles were improved correctly.

Keywords Network reconfiguration · Distributed generation · Harmonic Search Algorithm (HSA) · Network losses · Voltage profile · Voltage stability

1 Introduction

Volatility, high voltage degradation and high losses are as the most important problems in distribution networks. Network restructuring, which is a process of changing the topological structure of feeders by changing the open/closed state of the partitions and switches, can have a beneficial effect on solving these problems. On the one hand, network reconstruction and installation of distributed generation resources have several advantages such as: improving power quality, reducing losses, improving the voltage profile, and improving reliability of the network. In order to achieve the required level of load demand, improve voltage stability, high

M. Mohammadpoor · R. Ranjkeshan (✉)
Electrical and Computer Department, University of Gonabad, Gonabad, Iran
e-mail: ranjkeshanreza@gmail.com

A. Mehdizadeh
Department of Computing, Faculty of Science and Technology,
Nilai University, Putra Nilai, Negeri Sembilan, Malaysia

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_76

reliability, as well as gain economic benefits such as minimizing losses, energy efficiency, DG units are connecting to the distribution networks.

Grid rebuilding is a complicated technique due to the fact that the limitation of optimization problem is not recognizable as well. Several algorithms have been presented in this way the past where some of them are being reviewed.

Merlin and Beck [1] has analyzed network restructuring and used the branch-and-bound-type optimization technique. The disadvantage of this technique is that the solution is time-consuming because the possible structures of the system are as 2^n, where n is the number of lines with the key. Based on Merlin-Beck method, an exploratory or mental algorithm was proposed by Shirmohammadi and Hong in [2]. The disadvantage of this algorithm is that simultaneous keying is not considered in the feeder restructure. Civanlar et al. in [3] have proposed a method of exploration algorithm in which a simple formula was developed to determine changes in power losses caused by the transformation (or changing) of the branch.

The disadvantage of this method is that the operation of a key pair is considered at any one time, and the network restructure depends on the key state. Das in [4] has presented a subjective-based algorithm and a fuzzy multi-objective approach to optimize network structure. The disadvantage of this method is that there is no criterion for choosing membership functions for targets.

Nara et al. in [5] have proposed a method for solving this problem using the Genetic Algorithm (GA) to find out the least structure losses in the distribution system. In [6], an improved genetic algorithm (RGA) was proposed to reduce the losses in the distribution system. Rao et al. in [7] have proposed Harmony Search algorithm (HSA) to solve the network redistribution problem, which resulted in optimized simultaneous key switching in the network to minimize the real power losses of the distribution network.

The reorganization of the electricity market in many countries has created a new perspective on the distributed generation of electrical energy using renewable low capacity energy sources. DG units with a capacity of 5–10 KW are placed close to the consumer in order to provide electrical power. Since choosing the best position and size of DG units is still a complex optimization problem, many methods have been presented for solving it in recent years.

Rosehart and Nowicki in [8] presented a Lagrangian approach to determining the optimal position of DG placement in distribution systems, taking into account the economic and sustainability constraints. Celli et al. in [9] have presented a multi-objective algorithm using GA to locate and determine DG's size in the distribution system.

Wang and Nehrir in [10] have developed an analytical method for determining the optimum position of DG in the distribution system in order to minimize power losses. Agalgaonkar et al. in [11] has discussed the positioning approach and effect of DG inputs within the SMD framework. In [12] HAS algorithm is presented to solve the problem of network redistribution in the presence of distributed generation in order to reduce losses.

In this paper, a method to reconstructing and simultaneously installing DG distributed resources to minimize losses, and improve the voltage stability level and voltage profile simultaneously, is proposed, which benefits from the Harmonic Search Algorithm (HSA) algorithm for optimal optimization.

2 Distributed Generation Resources (DG)

Distributed generation units (DGs) are defined as any type of small-scale electrical power generation technology that can be installed and adapted to the distribution system. In some cases, due to poor location or capacity selection, it may not be economical to use the DG, but in addition to economical economics, there are other issues involved in the use of these generators, which cause an increasing use of this technology. Some of them are application of consumption management programs, reduction of lines losses and reduction of environmental pollution, pinching, correction of voltage profile and control, improving the quality of power and reliability in order to meet the needs of different customers and, in general, increasing the efficiency of energy production process.

3 Formulation of the Problem

In this paper, there are three objective functions are considered: increasing voltage stability, reducing network losses, and improving the voltage profile, which is described below.

3.1 Voltage Stability

Voltage stability is the ability of the power system to maintain a durable, acceptable voltage across all system shafts under normal and disturbing conditions. Voltage stability is divided into two categories: high voltage turbulence stability and small voltage disturbance stability. The main factors of voltage collapse are: reactive power control or generator voltage limitations, load characteristics, reactive compensators characteristics, and the performance of voltage control devices such as ULTCs. Several methods have been proposed for the purpose of voltage stability, usually made up of four general methods. These methods include using P-V curve, using Q-V curve, sensitivity analysis method, CPF continuous load distribution method. In this paper, sensitivity analysis was used to analyze voltage stability.

The sensitivity analysis of power systems can be written linearly as:

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} J_{P\theta} & J_{PV} \\ J_{Q\theta} & J_{QV} \end{bmatrix} \begin{bmatrix} \Delta\theta \\ \Delta V \end{bmatrix} \tag{1}$$

Assuming that the reactive power (Q) is constant, the reactive power increment (ΔQ) in the bus is zero. Using the inverse of the data (2) we have:

$$\Delta P = \left(J_{PV} - J_{P\theta}J_{Q\theta}^{-1}J_{QV}\right)\Delta V \tag{2}$$

$$\Delta V = \left(J_{RPV}\right)^{-1}\Delta P \tag{3}$$

where, the JRPV has a reduced Jacobin matrix that determines the variation of the voltage value based on the DG power-injected variations. If we model the bass as PQ bass, $J_{Q\theta}$ is a will be reversible matrix and $\mathrm{J}_{\mathrm{Q}\theta}^{-1}$ will be possible. Therefore, this situation can occur normally in distribution systems, in the case of Slack bus, it is the only one that maintains the size of the voltage. Accordingly, a bus with the lowest voltage sensitivity to power changes will have the highest voltage level. Therefore, the coefficient ΔV in relation (2) is a desirable criterion.

F1 is defined as the first objective function, namely, increasing the voltage stability level based on statistical learning and the above theory, is presented in Eq. (4). As stated, for voltage stability analysis, the voltage sensitivity index to the active power variation ratio is selected ($\frac{dP}{dv}$), which expresses the average of all bus values. The maximum index indicates that the bus has the lowest voltage stability that should be applied in this formula. Meanwhile, w_1 and w_2 are the weight coefficients of the relationship.

$$F_1 = \mathrm{w}_1\left(\frac{dP}{dv}\right)_{average} + \mathrm{w}_2\left(\frac{dP}{dv}\right)_{max} \tag{4}$$

As the function of Eq. 4 is less, the network status will be more favorable for the stability of the voltage.

3.2 Voltage Profile

The system voltage profile is defined as a numerical value of the voltage of each network bus whose analysis indicates the network voltage range, as well as low points, which are important in terms of the electrical energy quality of the system. The network voltage profile (F_2) is defined in Eq. (5).

$$F_2 = \sum_{\mathrm{i}=1}^{\mathrm{n}} \left|\mathrm{V}_{\mathrm{ref}} - \mathrm{V}_{\mathrm{i}}\right| \tag{5}$$

where, $\mathrm{V}_{\mathrm{ref}}$ is the base voltage of the system, which is usually calculated based on the specified range, and V_{i} is the voltage of each bus and n is the total number of busbars. The lower the F_2 value, the more favorable the profile of the voltage.

3.3 Power Losses Using Network Restructuring

The purpose of network redesigning is to find an appropriate structure for a radial network that has the lowest power losses, while the necessary enforcement constraints have been imposed all parameter of the distribution system including system's voltage profile, feeder flow range and radial structure are satisfactory. The loss of the line power between k and k + 1 shins, after the network reconstruction, can be calculated by Eq. (6) [13].

$$P'_{Loss}(k, k+1) = R_k \frac{\left(P'^2_k + Q'^2_k\right)}{\left|V'_k\right|^2} \tag{6}$$

Total power losses in all parts of the feeder $P'_{T,Loss}$, which is the total loss in all sections of the network, can be written as Eq. (7).

$$P'_{T,Loss} = \sum_{k=1}^{n} P'_{Loss}(k, +1) \tag{7}$$

As the third objective function (F_3), the net reduction of power losses ΔP^R_{Loss}, in the system is the difference between power losses before and after the restructuring, which is obtained by Eq. (8).

$$\Delta P^R_{Loss} = \sum_{k=1}^{n} P_{Loss}(k, k+1) - \sum_{k=1}^{n} P'_{Loss}(k, k+1) \tag{8}$$

When the DG unit is installed in the desired location on the network, we will have the following:

$$P_{DG,Loss} = R_k \frac{\left(P^2_k + Q^2_k\right)}{|V_k|^2} + \frac{R_k}{|V_k|^2}\left\{P^2_G + Q^2_G - 2P_kP_G - 2Q_GQ_k\right\}\left(\frac{G}{L}\right) \tag{9}$$

The net power loss in the system ΔP^{DG}_{Loss}, is the loss of powers before and after the unit DG installation, which is defined as Eq. (10).

$$\Delta P^{DG}_{Loss} = \frac{R_k}{|V_k|^2}\left\{P^2_G + Q^2_G - 2P_kP_G - 2Q_GQ_k\right\}\left(\frac{G}{L}\right) \tag{10}$$

The positive value of ΔP^{DG}_{Loss} indicates a reduction in system losses by installing the DG, while a negative value indicates that the DG has increased the losses in the system. The third goal in this paper is to reduce the total network losses.

4 Harmony Search Algorithm (HSA) [14–16]

The search algorithm based on the principles of music production is a successful meta-exploration algorithm for routing in wireless sensor networks and in order to increase the lifetime of these types of networks. The HS algorithm has received more attention in recent years due to the applicability of discrete and continuous optimization problems, low mathematical calculations, simple concepts, low parameters.

This algorithm has lower mathematical requirements than other meta-heuristic methods, and can be adapted to different engineering issues with changes in parameters and operators. Another advantage of this method than the genetic method is that it uses all the solutions in its memory to create a new solution, unlike the genetic method that uses two vector-based solutions in the generation. This feature increases the flexibility of the algorithm in search of better resolution spaces. Another feature of the synchronization search algorithm is that it detects better resolution spaces over a reasonable period of time.

This feature is problematic if the studied parameter is of local optimality and is stopped at local optimality and cannot be globalized. The reason for this problem is the inadequacy of the algorithm in implementing local search for discrete optimization issues. This algorithm consist of five steps:

Initialize the optimization problem and initial parameters
Set up the harmonic memory
Create a new improved harmony
Update the memory of the harmony.

Repeat steps 3 and 4 until the final condition is met or repetitions are ended.

5 The Proposed Method

In Sect. 3, the objective functions of the paper were introduced, which included three objectives: 1—Improvement of voltage stability level 2—Improvement of voltage profile indicator. 3—Reduction of network losses. In order to achieve the stated goals, the reconditioning method and the installation of distributed generation resources with optimum location and capacity are used with the help of Harmonic Search Algorithm (HSA).

In order to optimize the optimization, the three-objective functions are applied to the corresponding algorithm by applying weight coefficients (w_1, w_2, w_3 and w_4)

$$\begin{aligned} F = \mathrm{w}_1 \left(\frac{dP}{dv}\right)_{average} + \mathrm{w}_2 \left(\frac{dP}{dv}\right)_{max} + \mathrm{w}_3 \sum_{\mathrm{i}=1}^{\mathrm{n}} \left|\mathrm{V}_{\mathrm{ref}} - \mathrm{V}_{\mathrm{i}}\right| \\ + \mathrm{w}_4 \left(\Delta P_{Loss}^{R} + \Delta P_{Loss}^{DG}\right) \end{aligned} \tag{11}$$

The constraints applied to the algorithm are expressed as follows:

- Voltage range

$$V_{min} \leq |V_k| \leq V_{Max}$$

- Line flow range

$$|I_{k,k+1}| \leq |I_{k,k+1,Max}|$$

- After reconstruction, all network loads should be fed well.
- The grid is radial.

6 Simulation Steps

To simulate the proposed theorem, using the MATLAB software, the HS algorithm is implemented and fed the selected data into DIGSILENT software. The interface between these two software is a text file called in DPL language from DIGSILENT software. By executing these data on the network, the target function, which includes loss reduction, improvement of the voltage profile and increase of the voltage stability level, is calculated. In order to prove the effect of the proposed method (network redundancy and installation of DG units simultaneously), the HSA method is applied to a standard 33 node distribution network.

6.1 Sample Network

The study network has a 33-node radial distribution system with five keys and 32 switches. On the network, the switches (normal closed) are numbered from 1 to 32 and the normal open keys are numbered from 33 to 37. The total active and reactive power of the entire system is 3715 kW and 2300 KVAR. System DG units are only selected in three excellent locations [5] (Fig. 1).

7 Software Simulation

Simulation of the proposed method is done using MATLAB and DIGSILENT software. The network is shown in Fig. 2 before optimization.

The capacity of the DG unit has been selected from 0 to +5 MW and −5, to +5 MVar. To simulate the proposed method, the parameters of the algorithm are as follows:

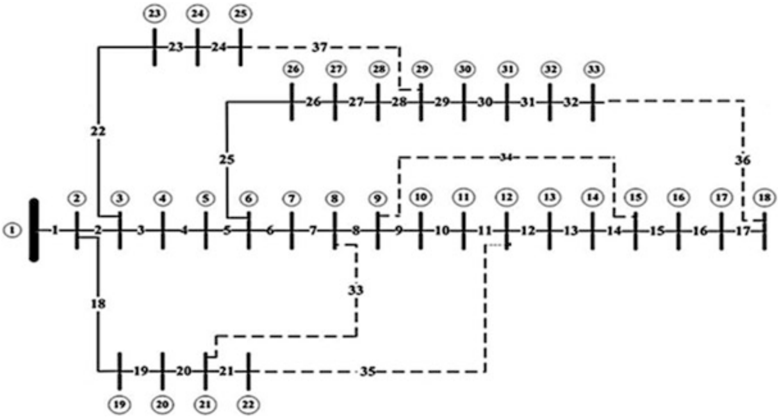


Fig. 1 Sample network 33 IEEE standard [17]

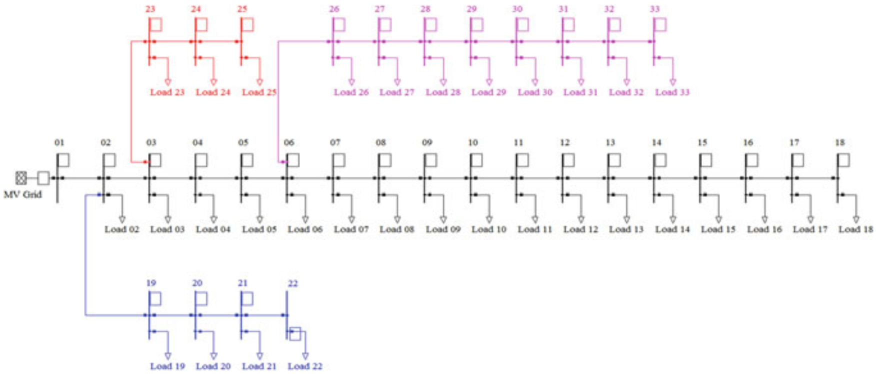


Fig. 2 Study network—before optimizing DIGSILENT software

Iterations	200
HMS	10
NHMS	10
HMCR	0.75
PAR	0.05

To ensure that the answer is optimal, the variables are executed 20 times with the same number of replicates, where more than 15 times the same answers are obtained. The variation of the target function during the implementation of the algorithm is shown in Fig. 3.

In Fig. 4, the simulated network is observed after the installation of distributed generation and reconstruction the network

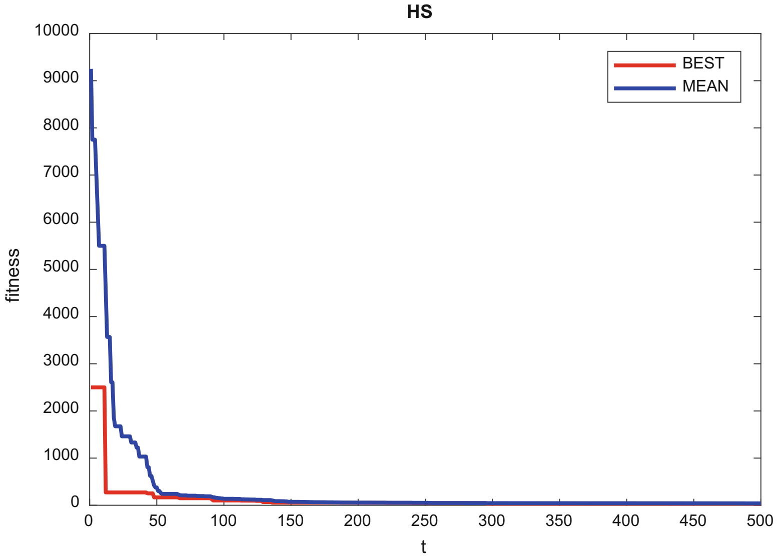


Fig. 3 The fitting curve of the target function

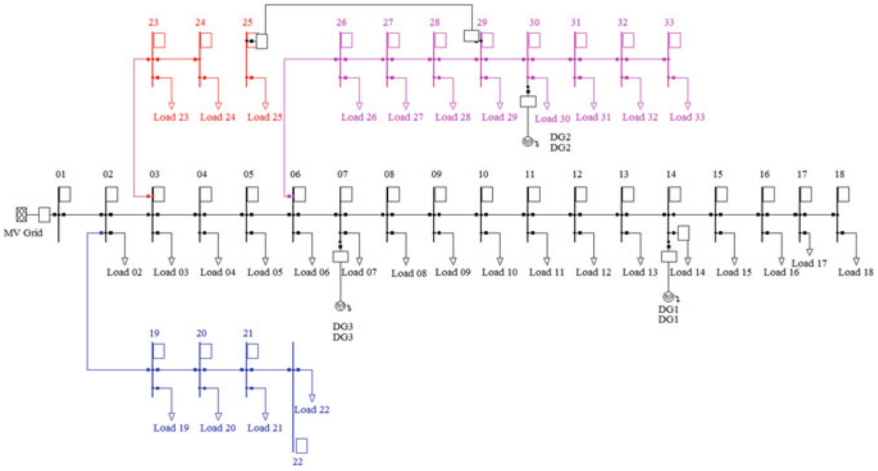


Fig. 4 Study network—after reconstruction and disposing of distributed generation sources in DIGSILENT software

8 Simulation Results

After implementing the proposed method in this paper, by selecting the optimal location and capacity of distributed generation sources according to Table 1 and reconstructing according to Table 2, after 20 times and 200 repetitions, the optimum response will be achieved by observing all relevant constraints.

By simulating and applying the results of the above optimization, Tables 3 and 4 are presented the variations of the three objective functions before and after optimization.

Changes in the network voltage profile before and after the proposed method, and voltage stability changes in two steps are presented in Fig. 5.

As shown in Figs. 5 and 6, the voltage stability index and the voltage profile after optimization (rearrangement and installation of distributed generation sources) have been improved.

Table 1 Choosing the location and optimal capacity of distributed generation sources with this proposed method

	Location	P (MW)	Q (MVar)
DG1	14	0.63357	0.03416
DG2	30	0.83892	1.0904
DG3	7	1.6285	0.93701

Table 2 The result of optimal reconstructing in the proposed method

Line		Line	
13–14	Close	9–15	Open
19–20	Close	8–21	Open
21–22	Close	12–22	Open
24–25	Open	25–29	Close
30–31	Close	18–33	Open

Table 3 Results of changes in target functions (voltage and loss stability)

	Average (VSI)	Max (VSI)	P_{loss} (kw)	Q_{loss} (kvar)
Before optimization	0.0341	0.081	189.34	126.24
After optimization	0.024974	0.067	16.98	17.24

Table 4 Results of voltage profile changes

	V_{min}	V_{max}	Voltage profile
Before optimization	0.94	1.03	1.0023
After optimzation	1.024	1.039	1.04678

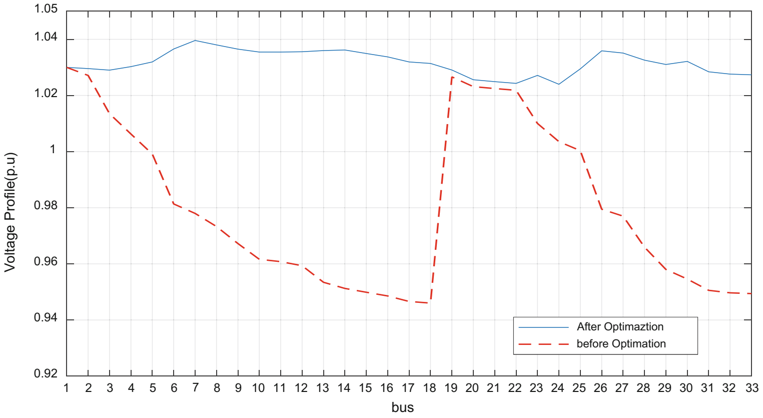


Fig. 5 Voltage profile status before and after optimization

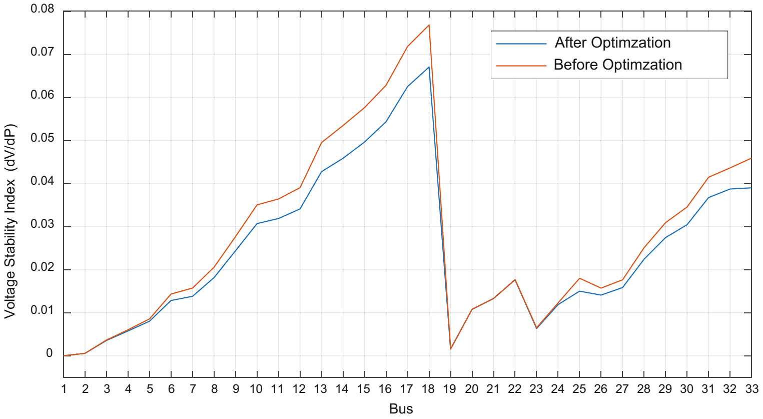


Fig. 6 Voltage stability status before and after optimization

9 Conclusion

In this paper, a new method is proposed for rearrangement and installation of DG units simultaneously in the distribution system. In addition, various methods for reducing losses, improving the voltage profile and increasing the voltage stability level have been simulated to demonstrate the superiority of the proposed method. An effective meta-method called HSA was used in the process of optimization for grid rebuilding and DG installation. The proposed method and other methods were tested on the system. The results showed that grid reorganization and DG installation simultaneously, compared with other methods, have a greater effect on reducing power losses and improving the voltage profile. The impact of the number of DG positions on reducing power losses at different load levels was studied, and, is shown the system voltage stability increased.

References

1. Merlin A, Back H (1975) Search for a minimal-loss operating spanning tree configuration in an urban power distribution system. In: Proceedings: 5th power system computation conference (PSCC), Cambridge, UK, pp 1–18
2. Shirmohammadi D, Hong HW (1989) Reconfiguration of electric distribution networks for resistive line losses reduction. IEEE Trans Power Deliv 4(2):1492–1498
3. Civanlar S, Grainger J, Yin H, Lee S (1988) Distribution feeder reconfiguration for loss reduction. IEEE Trans Power Deliv 3(3):1217–1223
4. Das D (2006) A fuzzy multi-objective approach for network reconfiguration of distribution systems. IEEE Trans Power Deliv 21(1):202–209
5. Nara K, Shiose A, Kitagawoa M, Ishihara T (1992) Implementation of genetic algorithm for distribution systems loss minimum reconfiguration. IEEE Trans Power Syst 7(3):1044–1051
6. Zhu JZ (2002) Optimal reconfiguration of electrical distribution network using the refined genetic algorithm. Elect Power Syst Res 62:37–42
7. Srinivasa Rao R, Narasimham SVL, Raju MR, SrinivasaRao A (2011) Optimal network reconfiguration of large-scale distribution system using harmony search algorithm. IEEE Trans Power Syst 26(3):1080–1088
8. Rosehart W, Nowicki E (2002) Optimal placement of distributed generation. In: Proceedings of 14th power systems computation conference, Sevillla, Section 11, Paper 2, pp 1–5
9. Celli G, Ghiani E, Mocci S, Pilo F (2005) A multi-objective evolutionary algorithm for the sizing and the sitting of distributed generation. IEEE Trans Power Syst 20(2):750–757
10. Wang C, Nehrir MH (2004) Analytical approaches for optimal placement of distributed generation sources in power systems. IEEE Trans Power Syst 19(4):2068–2076
11. Agalgaonkar P, Kulkarni SV, Khaparde SA, Soman SA (2004) Placement and penetration of distributed generation under standard market design. Int J Emerg Elect Power Syst 1:1
12. Geem ZW (2008) Novel derivative of harmony search algorithm for discrete design variables. Appl Math Comput 199(1):223–230
13. Prakash K, Sydulu M (2007) Particle swarm optimization based capacitor placement on radial distribution systems. In Proceedings of IEEE power engineering society general meeting, pp 1–5
14. Geem ZW, Kim JH, Loganathan GV (2001) A new heuristic optimization algorithm: harmony search. Simulation 76(2):60–68
15. Das S, Mukhopadhyay A, Roy A, Abraham A, Panigrahi BK (2011) Exploratory power of the harmony search algorithm: analysis and improvements for global numerical optimization. IEEE Trans Syst Man Cybern B Cybern 41(1):89–106
16. Geem ZW, Tseng C, Park Y (2005) Harmony search for generalized orienteering problem: best touring in China. In: Proceedings of ICNC, vol 3612, Springer, Heidelberg, pp 741–750
17. A Multi-objective Network Reconfiguration of Distribution Network with Solar and Wind Distributed Generation using NSPSO, Subas Ratna Tuladhar, Jai Govind Singh and Weerakorn Ongsakul, International Conference and Utility Exhibition 2014 on Green Energy for Sustainable Development (ICUE 2014) Jomtien Palm Beach Hotel and Resort, Pattaya City, Thailand, 19–21 March 2014.
18. Ghosh S, Sherpa KS (2008) An efficient method for load-flow solution of radial distribution networks. Int J Elect Power Energy Syst Eng 1(2):108–115

Multi-band Rectangular Monopole Microstrip Antenna with Modified Feed Junction for Microwave Wireless Applications

Mohammad Faridani and Ramezan Ali Sadeghzadeh

Abstract A multi-band microstrip antenna, based on a cheap FR-4 substrate and with thickness of 1.6 mm is proposed in this paper. A 50 O female SMA connector is welded to antenna and the antenna is fed by coaxial cable. The antenna has 8 bandwidths between 2.2 and 11 GHz frequency region. −68.04 dB return loss has been measured at 3.099 GHz in the second bandwidth. Finally, the antenna is used as array elements in 6 different linear and planar geometric arrangements. 17.2 dBi directivity is achieved at 7.9 GHz by 4 elements linear array antenna. The antenna has a simple structure and is suitable for microwave wireless applications such as energy harvesting. The voltage standing wave ratio (VSWR), return loss, radiation pattern and directivity are provided.

Keywords Multi-band antenna · Planar antenna · Array antenna

1 Introduction

In recent decades, wireless communication systems have become an inseparable part of modern life. With the swift progress and development of wireless systems, multi-band wireless systems which operate in different frequencies, are growing rapidly. Using one multi-band antenna instead of several antennas will simplify and miniaturize the wireless system structures, which are two important factors in designing communication systems.

M. Faridani
Department of Electrical and Computer Engineering,
Science and Research Branch, Islamic Azad University, Tehran, Iran
e-mail: mohammadfaridani@yahoo.com

R. A. Sadeghzadeh (✉)
Faculty of Electrical and Computer Engineering,
K. N. Toosi University of Technology, Tehran, Iran
e-mail: sadeghz@eetd.kntu.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_77

In literature reviews, there are various antennas which operate in more than one bandwidth; such as dual band [1, 2], triple band [3, 4] and quad band [5].

There are Also different antennas that are multi-band, for instance, quasi-yagi-type antenna [6], compact open-ended slot [7], antenna with a coupling feed and parasitic elements [8], narrow size inverted-F antenna [9], arc-shaped [10], low-profile ferrite [11], and other types [12–14].

Due to its appealing features such as low cost, low profile, compact size and light weight, the microstrip antenna is a well-known kind of antenna [15]. Different kinds of feeding techniques such as coaxial probe and microstrip line have been used for exciting the antenna. Because of its advantages, the microstrip antenna has a lot of applications in various fields like scientific research and medicine.

The multi-band monopole microstrip antenna is presented in this paper. A partial circle is added between the feed line and the radiation patch to modify this junction and decrease its bad effect. Some antenna parameters such as return loss and radiation pattern are developed. The proposed multi-band antenna is set in different array arrangements and the array antennas directivities are shown and compared.

2 Antenna Design

The geometry and configuration of the proposed multi-band rectangular microstrip antenna are shown in Fig. 1. The rectangular radiation patch has $L_P = 26$ mm length and $W_P = 47$ mm width. The patch is excited by the microstrip feed line, the length of the feed line is $L_F = 29$ mm and the width is $W_F = 2.5$ mm. To decrease the bad effect of the microstrip feed line to patch junction, a section of a circle with $R = 13$ mm radius and $a = 5$ mm, shown in Fig. 1, has modified the junction.

The antenna is printed on a FR-4 substrate with relative permittivity of 4.4, loss tangent of 0.02 and $h = 1.6$ mm thickness. The ground and substrate have the same width and length of $W = 55$ mm and $L = 65$ mm, respectively.

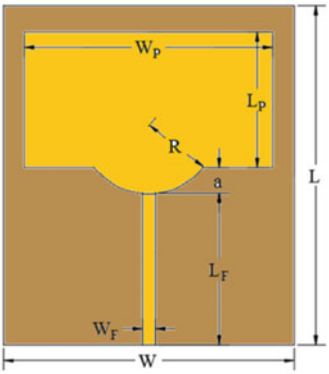


Fig. 1 Structure configuration of the proposed multi-band antenna

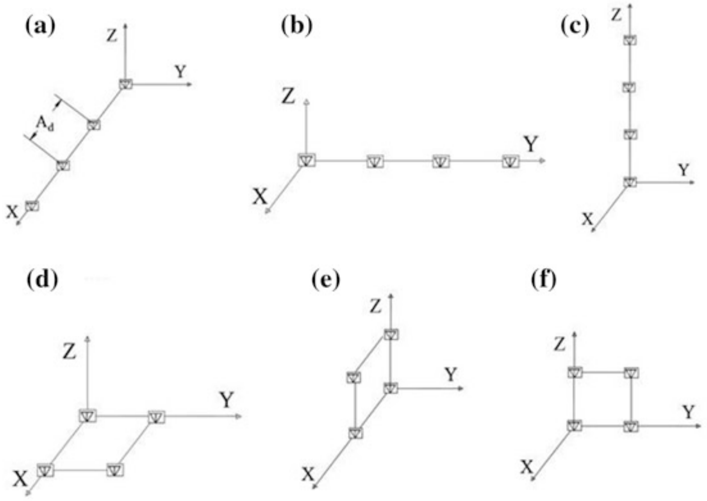


Fig. 2 Various array antennas arrangements, linear **a** X direction, **b** Y direction and **c** Z direction, planar in **d** X-Y plane, **e** X-Z plane and **f** Y-Z plane

Next, the antenna is considered as array elements for designing directive array antenna. As shown in Fig. 2, 6 array arrangements contain of 3 linear ones in X, Y and Z directions and 3 planar ones in X-Y, X-Z and Y-Z planes are presented. The distances between array elements are $A_d = 80$ mm and for simplify all elements are excited by same input signals with equal values and phases ($\text{Ø} = 0$).

3 Simulation, Measurement and Discussion

The simulated proposed antenna voltage standing wave ratio (VSWR) is shown in Fig. 3. The antenna resonant frequencies are at 2.5, 2.98, 3.92, 4.84, 5.9, 6.7, 7.12, 8.08, 9.34 and 10.26 GHz.

Next, the antenna is fabricated and measured in Tehran University's laboratory. The lab's temperature is 25 ± 5 CO and has relative humidity less than 60%. Agilent E8361C PNA series network analyzer with 10 MHz–67 GHz is utilized for measurement. The measurement devices are inside a shielded pyramidal anechoic chamber.

As shown in Fig. 4, the antenna has eight bandwidths which cover 2.446–2.58 GHz, 2.979–3.262 GHz, 3.998–4.418 GHz, 4.873–5.634 GHz, 6.212–6.811 GHz, 7.593–8.161 GHz, 8.957–9.352 GHz and 10.245–10.714 GHz with eight resonant frequencies at 2.512, 3.099 GHz ($S_{11} = -68.04$ dB), 4.19, 5.258, 6.443, 7.8, 9.145 and 10.483 GHz.

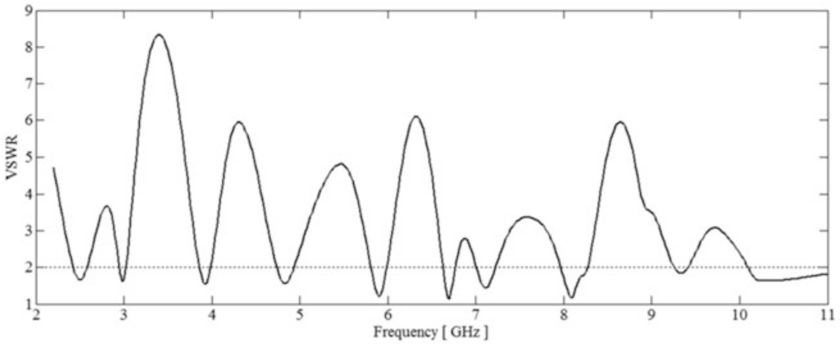


Fig. 3 Simulated VSWR of antenna

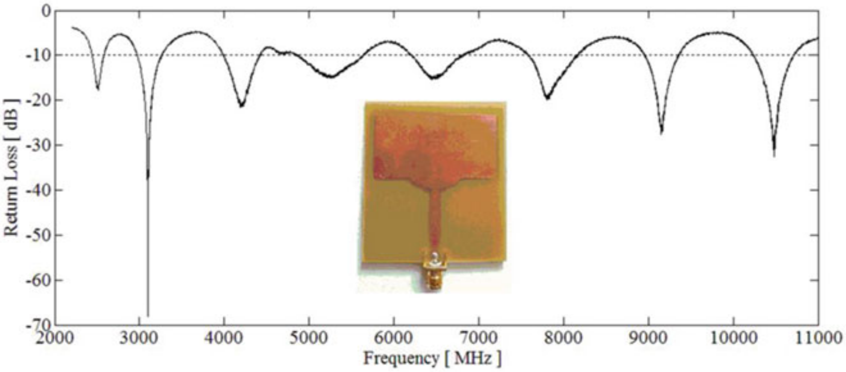


Fig. 4 Measured return loss of proposed antenna with photograph of the prototype antenna

The difference between simulation and measurement is due to the tolerance in the manufacturing process and the 50 Ω SMA connector, which is not presumed in simulation.

The radiation patterns at 5.4 GHz (fourth bandwidth) and at 7.8 GHz (sixth bandwidth) are illustrated in Fig. 5. At 5.4 and 7.8 GHz frequencies, the main lobe directions are at 90°.

The directivity plot which is shown in Fig. 6 indicates that the antenna is directive and has a maximum 11.3 dBi directivity at 7.9 GHz.

The linear array antennas have maximum directivities of 17.2, 14.8 and 14.7 dBi at 7.9, 10.3 and 7.8 GHz in X, Y and Z directions and the planar array antennas have maximum directivities of 14.8, 14.8 and 16.1 dBi at 10.4, 10.3 and 7.9 GHz in X-Y, X-Z and Y-Z planes respectively, as shown in Fig. 7.

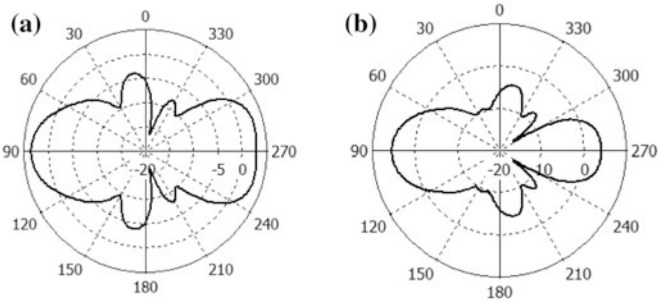


Fig. 5 Proposed antenna radiation pattern, **a** at 5.4 GHz and **b** at 7.8 GHz

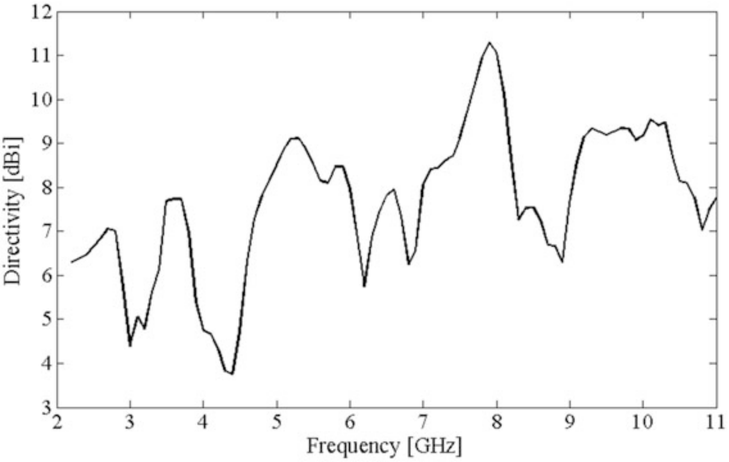


Fig. 6 Proposed antenna maximum directivity

4 Conclusion

This paper presents the multi-band rectangular microstrip antenna based on a single-layer substrate. The proposed planar antenna is simulated. Subsequently, it is fabricated and measured in Tehran University's laboratory. The antenna has 8 bandwidths, 8 resonant frequencies, a minimum return loss of −68.04 dB and a maximum directivity of 11.3 dBi. At the end, between the array antennas the 4 elements linear array antenna in X direction has the maximum directivity of 17.2 dBi. The proposed antenna is suitable for wireless applications like microwave energy harvesting.

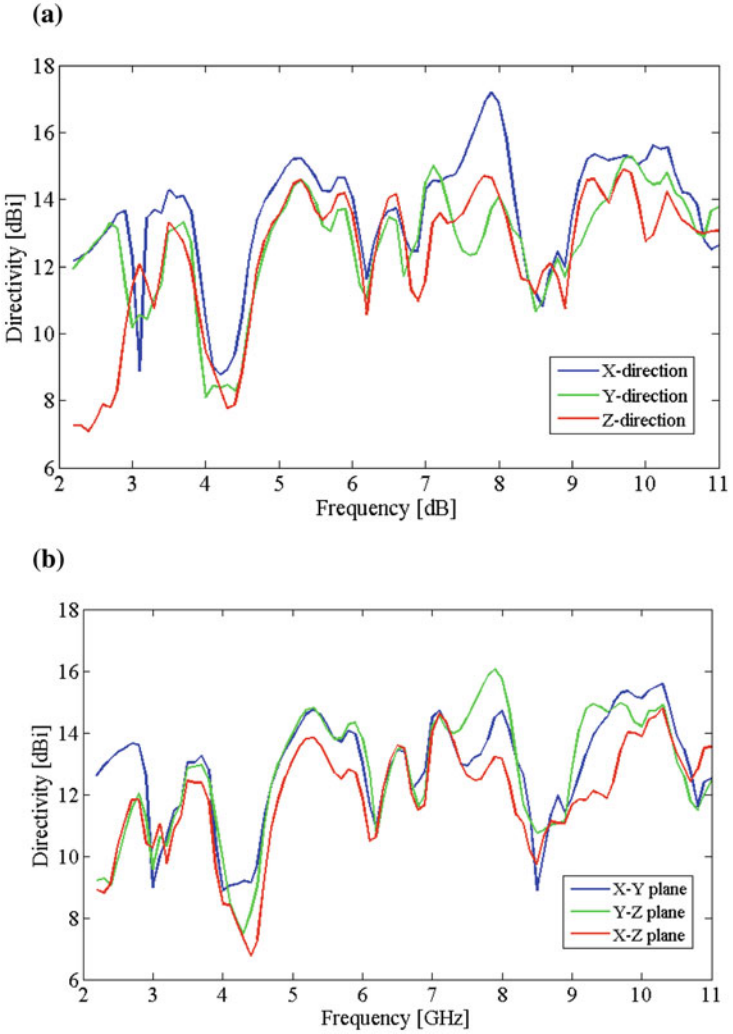


Fig. 7 Array antennas directivities **a** linear arrangements, **b** planar arrangements

References

1. Honari MM, Mirzavand R, Saghlatoon H, Mousavi P (2016) A dual-band low-profile aperture antenna with substrate-integrated waveguide grooves. IEEE Trans Antennas Propag 64 (4):1561–1566
2. Başaran SC, Erdemli YE (2008) Dual-band split-ring antenna design for WLAN applications. Turk J Elect Eng Comput Sci 16(1):79–86
3. Li L, Huang Y, Zhou L, Wang F (2016) Triple-band antenna with shorted annular ring for high-precision GNSS applications. IEEE Antennas Wirel Propag Lett 15:942–945

4. Wu RZ, Wang P, Zheng Q, Li RP (2015) Compact CPW-fed triple-band antenna for diversity applications. Electron Lett 51(10):735–736
5. Azaro R, Viani F, Lizzi L, Zeni E, Massa A (2009) A monopolar quad-band antenna based on a hilbert self-affine prefractal geometry. IEEE Antennas Wirel Propag Lett 8:177–180
6. Ding Y, Jiao YC, Fei P, Li B, Zhang QT (2011) Design of a multiband quasi-yagi-type antenna with CPW-to-CPS transition. IEEE Antennas Wirel Propag Lett 10:1120–1123
7. Cao Y, Yuan B, Wang G (2011) A Compact multiband open-ended slot antenna for mobile handsets. IEEE Antennas Wirel Propag Lett 10:911–914
8. Kim KJ, Lee S, Kim BN, Jung JH, Yoon YJ (2011) Small antenna with a coupling feed and parasitic elements for multiband mobile applications. IEEE Antennas Wirel Propag Lett 10:290–293
9. Pazin L, Leviatan Y (2011) Narrow-size multiband inverted-F antenna. IEEE Antennas Wirel Propag Lett 10:139–142
10. Verma S, Kumar P (2015) Compact arc-shaped antenna with binomial curved conductor-backed plane for multiband wireless applications. IET Microwaves Antennas Propag 9(4):351–359
11. Lee W, Hong YK, Park J, Choi M, Lee J, Baek IS et al (2016) Low-profile multiband ferrite antenna for telematics applications. IEEE Trans Magn 52(7):1–4
12. Liu Y, Shi D, Zhang S, Gao Y (2016) Multiband antenna for satellite navigation system. IEEE Antennas Wirel Propag Lett 15:1329–1332
13. Ahmed S, Tahir FA, Shamim A, Cheema HM (2015) A Compact kapton-based inkjet-printed multiband antenna for flexible wireless devices. IEEE Antennas Wirel Propag Lett 14:1802–1805
14. Hsu CK, Chung SJ (2015) Compact multiband antenna for handsets with a conducting edge. IEEE Trans Antennas Propag 63(11):5102–5107
15. Balanies CA (1982) Antenna theory analysis and design. Wiley, New York

Electrostatic MEMS Switch with Vertical Beams and Body Biasing

Armin Bahmanyaran and Kian Jafari

Abstract One of the easiest and the most appropriate solution for fabricating MEMS switch is Electrostatic actuation. However, it requires large parallel plates which can take a large surface on the substrate wafer and thus an expensive fabrication cost. In this paper, an improved electrostatic MEMS switch is presented. The proposed design is carried out so that it minimizes the dimensions of the MEMS device. Simulation results are also done by COMSOL based on the proposed design.

Keywords Switch · Electrostatic actuator · Micro electro mechanical systems Bulk terminal · Vertical beams · Pull-in voltage

1 Introduction

Nowadays, MEMS switches are under investigation for use in digital circuits with very low power consumption or in high temperature environments [1]. They can be used to complete or replace MOSFETs, but due to problems such as high latency, short life time and the large area occupied, they are still not comparable for mass production to semiconductor transistors [1]. When we want to make a full mechanical chip, which thousands (or millions) of MEMS switches are used inside, we need to design them as simple as possible and with the smallest possible area, which less complications, such as dielectrics, insulators and … will be used. In this report, the solution to reduce the area and compression of digital mechanical circuits without drawbacks was expressed. Other solutions create new problems. For example, reducing the air gap can reduce the area of parallel plates to have the same supply voltage, but it also reduces the spring constant [2], which increases the Possibility of adhesion between the plates and reduces their life time [2].

A. Bahmanyaran (✉) · K. Jafari
Electrical Engineering Department, University of Shahid Beheshti, Tehran, Iran
e-mail: arminiufb@yahoo.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_78

Step design

The simplest MEMS switch consists of three or four pieces of flexible metal (drain gate source and bulk) for the implementation of digital circuit. According to Fig. 1a, the potential difference between the gate and source leads to the electrostatic force and bends source plate onto the drain [3]. With The smaller area of designed switch on substrate the more compact digital circuits can be designed, but the use of actuation electrostatic force needs a large area of parallel plates. One solution to reduce the area is that plates place vertically on the substrate like Fig. 1b and then, by putting a number of them together, a compact digital circuit can be made like Fig. 1c. In this method, all three parallel beams are perpendicular to the substrate with a distinct air gap [4], but the electrostatic force only bends the source beam because the gate and drain beams are thicker and have less flexibility. The air gap between the drain and source is less than the air gap between gate and source to prevent the gate and source from connecting each other. This air gap ratio is better less than 1/3 so that the movement is stable with low velocity and the lifetime of the mechanical parts will increase [3]. In addition to the small area and compact circuit design, the other advantage of beams perpendicular to substrate is simplicity in fabrication since there is no need to use the sacrificial layer and the bulk fabrication technique can be used.

MEMS switches same as the MOSFETs, can have a fourth terminal (bulk) for tuning the pull-in/threshold voltage [5]. In this method, the electrostatic force is created by potential difference between the gate and bulk terminal to bend the plate of bulk and source onto the drain. Figure 2a is the final shape of the proposed design. Insulator is placed between the bulk and source to be electrically separated and mechanically connected [6]. The bulk should be completely isolated from the other terminals even when the switch is on [5]. If the bulk is connected to the ground, switch acts like a normal three-terminal N-type transistor, and as the bulk

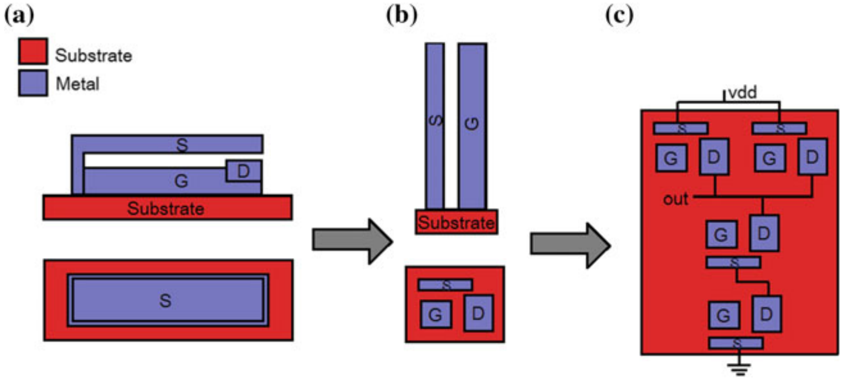


Fig. 1 Design steps. **a** Side and top view of a MEMS switch whose plates are horizontally located on a substrate and occupy a large area. **b** Side and top view of Improved design whose plates are vertically located on a substrate and occupy small area. **c** A compact digital circuit example using switches with vertical beams

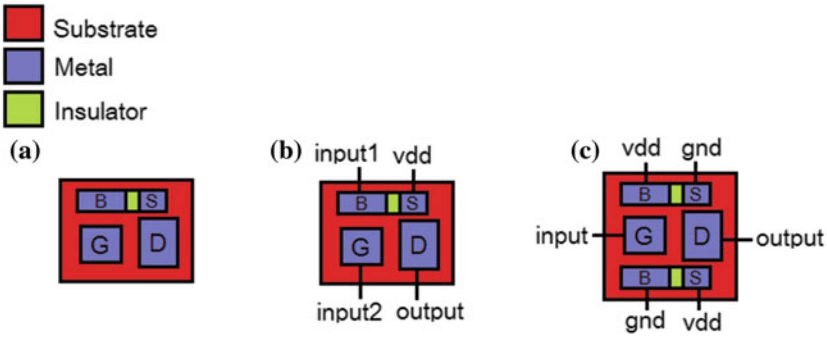


Fig. 2 Using body biasing. **a** Top view of design Improved by adding the bulk terminal for tuning pull-in voltage. **b** Another use of the bulk terminal to design an XOR gate. **c** Another example of using the bulk terminal to design a digital buffer

voltage increases, the pull-in and pull-out voltage increase linearly [5]. In digital circuits, tuning the pull-in voltage is important [5], for example, in an inverter, the only way to Symmetry the voltage transfer curve (VTC) is tuning the pull-in voltage by the voltage of the bulk terminal [5]. In addition by using the bulk terminal, logical circuits can be implemented in a simpler way such as Fig. 2b which is a digital XOR designed by using bulk terminal because the electrostatic force between the gate and bulk occurs only when one of the two inputs be VDD and the other be GND. It can also use two bulk terminals parallel to the gate terminal that drive two separate sources like Fig. 2c which is a digital buffer designed. Another design with the bulk and vertical beams is two gate terminals parallel to one bulk, which can implement more logical functions.

Design modeling

To simulate, we first need to choose the appropriate values of dimensions and parameters, so we need to specify the relationships between the output and inputs of the design. If we use the model of mass and spring for modeling, the electrostatic force and surface adhesion force want to connect the beams to each other, and the elastic force and damping force want to separate them [6]. We ignore the damping force and surface adhesion force, because the initial distance between the beams is large enough and they are smaller than the spring force and electrostatic force.

The relation between displacement and the gate voltage is according to the following equation [7]:

$$V = \sqrt{\frac{2k}{\varepsilon 0\,\mathrm{A}} d^2(\mathrm{d}0 - d)} \qquad (1)$$

where d0 is the initial distance between the beams and d is their displacement. K is the constant of the spring and A is the area of the parallel plates and Ɛ0 is permittivity of the air gap.

For all parameters, optimal values can be found [2]. As the Height of the beams is longer, the lower supply voltage is needed [2], but the chance of breaking of them becomes higher and the adhesion force becomes stronger. As the gate width is longer, the lower supply voltage is needed [3], but the area becomes larger, which is undesirable. As the more flexible metal we choose, it will require less supply voltage [3], but the mechanical strength and quality factor will be reduced [3]. The scale we choose for a design is less important, instead the shape of the design is important because the scales are convertible. For example, a design with 100 μm dimensions can be converted to 1 μm with a few change.

2 Simulation and Results

To simulate the layout in the COMSOL software, we put all dimensions in a parametric form so that we can find the best sizes for it by changing them. We chose titanium for this design in terms of mechanical strength and flexibility are in equilibrium [3]. Silicone oxide selected as insulator and dimensions of the design are selected as mentioned Fig. 3. Figure 3c shows the displacement with FEA (finite element analyze) in COMSOL, when the gate potential is 5 V, the electrostatic force only displaces the source and bulk plate by one tenth of a micrometer (because the gate and drain beams are thicker and have less flexibility) lead to connection between drain and source to each other.

After the drain and source connected, we need to make sure that the gate and bulk would not connect to each other, so once again we simulate the layout with connected drain and source. According to Fig. 4c, the corner of the bulk plate

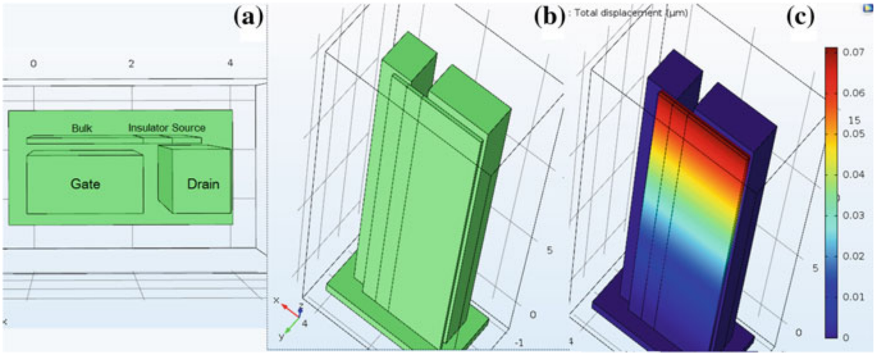


Fig. 3 Simulating in COMSOL **a** top view with electrodes name, **b** 3d view of design, **c** display of displacement in micrometer for gate_voltage = 5 V. Dimensions: beams height = 15 μm, gate_width = 2 μm, gate_depth = drain_width = 1 μm, insulator_width = source_width = 0.5 μm, bulk_depth = 0.1 μm, drain_depth = 1.1 μm, gate_source_gap = 0.2 μm, drain_source_gap = 0.1 μm

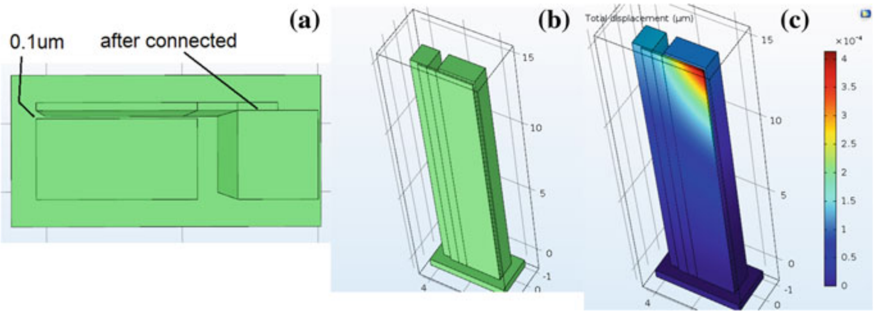


Fig. 4 Another simulation with connected drain and source to make sure that the gate and bulk are not connect to each other for gate_voltage = 5 V **a** top view, **b** 3d view of new simulation, **c** FEA shows that displacement of the bulk plate is not enough to connect to the gate electrode

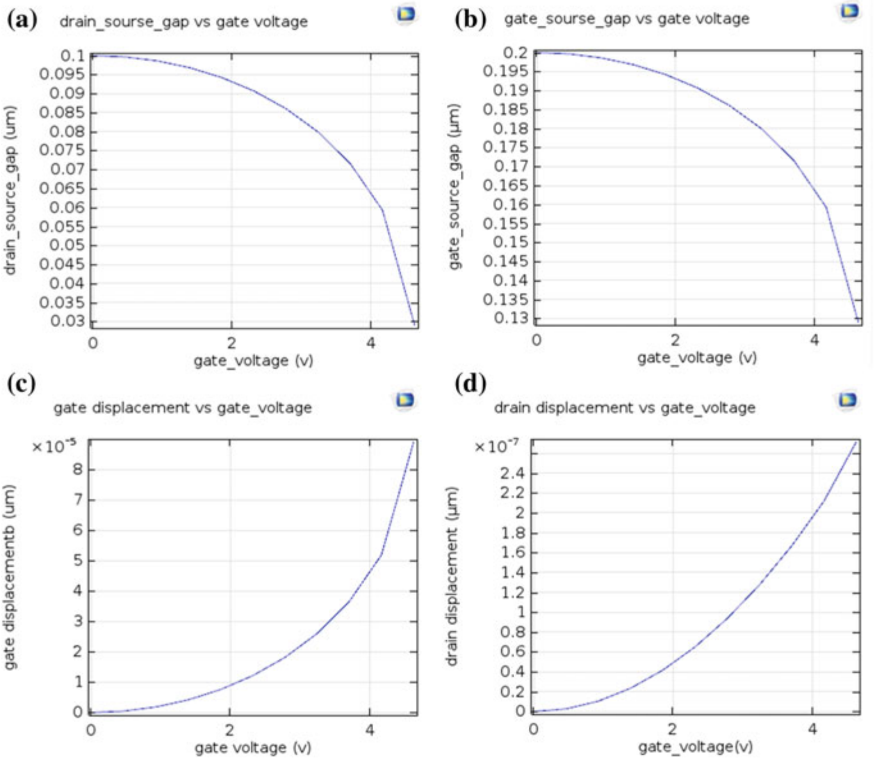


Fig. 5 Simulation results **a** reduction of the air gap between source and drain by increasing the gate voltage, **b** air gap between gate and source, **c** little displacement of the gate by the gate voltage, **d** little displacement of drain by the gate voltage that not interfere the operation of the switch

displaces less than 0.1 μm (0.0004 μm) that is not enough to connect to the gate electrode. Note that as the gate width increases, this displacement can be increased.

The curves shown in Fig. 5 are obtained from the finite element analysis in COMSOL, Fig. 5a, b illustrate the reduction of the air gap by increasing the gate voltage and at gate voltage of about 5 V, the air gap between the drain and source is zero.

Figure 5c, d shows that the gate and drain displacements by the gate voltage are small enough that not interfere the operation of the switch.

3 Conclusion

In this paper, an improved version of MEMS switch is presented. In the proposed MEMS device, vertical beams and the canal between drain and source place have been designed vertically on the substrate wafer. Furthermore, a bulk terminal is added to the proposed design in order to tune the threshold voltage to obtain the desired value.

This can reduce the occupied surface on the substrate wafer and thus more compact digital circuits and easier fabrication processes are achievable due to no required sacrificial layer which can be fabricated by bulk fabrication techniques.

References

1. Sinha N, Jones T, Guo Z, Piazza G (2010) Demonstration of low voltage and functionally complete logic operations using body-biased complementary and ultra-thin ALN piezoelectric mechanical switches. In: 2010 IEEE 23rd international conference on micro electro mechanical systems (MEMS), IEEE, pp 751–754
2. Akarvardar K, Eggimann C, Tsamados D, Chauhan YS, Wan GC, Ionescu AM, … Wong HSP (2008) Analytical modeling of the suspended-gate FET and design insights for low-power logic. IEEE Trans Electr Dev 55(1):48–59
3. Akarvardar K, Elata D, Parsa R, Wan GC, Yoo K, Provine J, … Wong HS (2007) Design considerations for complementary nanoelectromechanical logic gates. In: Electron devices meeting, IEDM 2007, IEEE international, IEEE, pp 299–302
4. Tabib-Azar M, Venumbaka SR, Alzoubi K, Saab D (2010) 1 volt, 1 GHz NEMS switches. In: Sensors, 2010 IEEE, IEEE, pp 1424–1426
5. Nathanael R, Pott V, Kam H, Jeon J, Liu TJK (2009) 4-terminal relay technology for complementary logic. In: Electron devices meeting (IEDM), 2009 IEEE international, IEEE, pp 1–4
6. Alzoubi K, Saab DG, Tabib-Azar M (2010) Circuit simulation for nano-electro-mechanical switches VLSI circuits. In: 2010 53rd IEEE international midwest symposium on circuits and systems (MWSCAS), IEEE, pp 1177–1180
7. Chakraborty S, Chaudhuri AR, Bhattacharyya TK (2009) Design and analysis of MEMS cantilever based binary logic inverter. In: International conference on advances in computing, control and telecommunication technologies, ACT'09, pp 184–188

Optimal Clustering of Nodes in Wireless Sensor Networks, Using a Gravitational Search Algorithm

Saeid Madadi barough and Ahmad Khademzadeh

Abstract Wireless sensors are among the most appropriate data collection solutions in the world. Information collected by sensors should be transmitted to a base station. In direct transmission, each sensor sends information directly to the center. Due to the high distance of the sensors from the center, they consume a lot of energy. In contrast, designs that create the shortest distance can prolong the life of the network. In these networks, sensor nodes are often confronted with limitations due to small size, such as processing power and limited power supply. These limitations have led researchers to carry out extensive studies in the design of these networks. The nodes are more distant than the nodes of the cluster, consuming more time and power to transmit data to the cluster. Therefore, better clustering of sensor nodes in wireless sensor networks to reduce the power consumption of sensor nodes through better connection of sensor nodes to cluster nodes using a gravitational search algorithm, and the results are optimized with artificial and optimal beehive algorithms. Particle swarm is compared.

Keywords Wireless sensor networks · Clustering · Gravitational search algorithm Evolutionary algorithms

1 Introduction

Recent advances in micro-electro-mechanical systems, smart sensors, wireless communications, and digital electronics make it possible to build small, low-power, low-cost sensor nodes that can communicate wirelessly [1]. These sensor nodes

S. Madadi barough (✉)
Computer Engineering Department, South Tehran Branch,
Islamic Azad University, Tehran, Iran
e-mail: S.madadi.b@gmail.com

A. Khademzadeh
Research and Communication Research Institute
(Iran Telecommunication Research Center), Tehran, Iran
e-mail: zadeh@itrc.ac.ir

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_79

consist of three sensor parts, information processing and wireless data transmission. Generally, a wireless sensor network contains a large number of these nodes, which is used to measure a parameter, and their data is considered collectively. That is, all data collected for a parameter is processed in a node of the network, usually called a wellhead node, and the actual value of that parameter is estimated fairly accurately [2].

In these networks, the failure of a network node has almost no effect on the estimated value. In wireless sensor networks, sensor nodes with a large number of intrinsic or near-desired sensor nodes are used to measure the parameter. The location of these nodes is not already designed. This simply helps to accommodate the sensor on the network, but instead, the protocols used for these networks should be self-regulating or self-organized, given that these sensors have a CPU inside. To reduce the amount of data transmission, these sensors only send the required data after processing the initial data. In wireless sensor networks, the main limitation for designing protocols is the limited energy of sensors. In fact, protocols that minimize power consumption in sensors are more relevant to wireless sensor networks. The nodes are more distant than the nodes of the cluster, consuming more time and power to transmit data to the cluster. Therefore optimal clustering of sensor nodes by connecting them to the nearest node of the cluster to reduce the power consumption of the network is raised as an optimization problem. Different algorithms are provided for optimal node connectivity in wireless sensor networks. But once these algorithms are used, the size of the grid and the number of nodes increases, and a lot of money is needed to get the optimal response. The timing of classical algorithms does not have the proper ability to solve this problem. Therefore, it is necessary to use other algorithms that, at an acceptable time, obtained an appropriate response to this problem. Among these suitable algorithms are evolutionary algorithms (Goyal). One of the evolutionary algorithms is the gravitational search algorithm. The continuation of the structure of this paper is that in the second section we will review the history of the research, and in the third part we will describe the evolutionary algorithms. In Sect. 4, the proposed model and the evaluation of the results of the use of evolutionary algorithms are presented for problem solving, and finally, the conclusion is drawn in the final section.

2 A Review of Past Research

In 2014, in a paper titled An Anthem Algorithm, Liu introduced a greedy migration mechanism to locate nodes in a sensor network to examine the problem of network coverage with minimal cost and reliable connectivity. The proposed method, using the ant colony algorithm, targeted the problem of network coverage and reduction of location cost. This algorithm adjusted the communication radius to improve the energy efficiency and network lifetime. The simulation showed that the proposed method reduces the cost of locating and establishes a balance between energy consumption between nodes and increases network lifetime in a sensor network [3].

In 2014, in a paper titled Comparison of Particle Swarm Optimization in Locating Sensor Network Nodes, Cao and colleagues studied various algorithms for optimizing particle swarm with different demographic topologies. The simulation results indicated that they are efficient in the area of the ring and square topology sensor network [4]. Brock et al., in 2014, in an article on placement for maximizing coverage in an asymmetric sensor network, using the genetic algorithm of a sensor network, which has an effective network coverage purpose in which asymmetric nodes have different strings. The simulation results showed that the proposed method provides an effective solution for nodes locating and eliminating overlap to maximize coverage [5]. Zhou provided a 2015 article on a locating method based on the particle swarm optimization algorithm and the Covey-Newton algorithm for sensor networks. In this method, the PSO algorithm was used to find optimal solutions and values, then these values were used in the repeat phase as input values for the Coasi-Newton algorithm. The simulation results indicate that the proposed algorithm has succeeded in locating the nodes [6].

Barlo et al., in an article in the article entitled "Locating nodes in wireless networking networks" in 2015, analyzing the results of the simulation of the WMN-GA system for various parameters and distributions, examined wireless networks. In this paper, they evaluated the efficiency of 4 different distributions (normal, uniform, exponential, and Wi-Fi) for lane routers, which took into account the proportion of data delivery and latency. To simulate a HWMP protocol. The simulation showed that the system used can be successfully used to locate routers [7]. Pong and his colleague in 2015, in an article entitled optimal locating algorithm based on the genetic algorithm in the sensor network, investigated the sensor networks for locating without a DV-Hub type.

In this paper, we used a DV-Hub non-range locating algorithm. This algorithm was optimized based on genetic algorithm. The simulation results indicated that the proposed method would increase the location accuracy compared to previous methods [8].

In 2015, Poja et al., in a method known as optimal positioning of sensors, use the BFA algorithm to examine sensor nodes as food-seeking bacteria, which are shown by the best possible connections. In this article, we used the assumption that a space of space can be covered by using a hexagon. As a result, using the BFA algorithm to optimize locations so that all nodes in the network move towards the vertices of the hexagon that are connected to each other, resulting in full coverage of the space [9]. San and et al., in an article on optimal location in the sensor network based on the algorithm of the mix of ant culture, in 2015, the evolutionary mechanism of the culture algorithm operates in a crowd-prone algorithm algorithm as a unitary evolutionary strategy, followed by searches It guides the population in the best choices and responses that makes the cultural search algorithm perform the optimal answer faster and more stable than other conventional algorithms [10].

3 Evolutionary Algorithms

The concept of optimization is such that, among the parameters of a function, we look for the values that minimize or maximize the function. All the appropriate values for this, the solutions possible and the best of these values, are called optimal solutions. Evolutionary algorithms cover both types of maximization and minimization issues. Optimization has always been accompanied with many problems. The former methods of solving optimization problems require countless computational efforts. Algorithms such as collective intelligence algorithms have partly solved this problem. By these algorithms, solutions are found that are almost close to the answer [11].

3.1 Gravitational Search Algorithm

Gravity Search Algorithm A gravitational search algorithm is optimized by means of a plan of gravitational and motion laws in a discrete-time synthetic system. The system environment is the same as the definition of the problem. According to gravity law, each mass perceives the location and condition of other bodies through gravitational gravity. Therefore, this force can be used as a means of exchanging information. From the optimized finder, it can be used to solve the optimization problem in which each answer to a problem is expressed as a distance. The amount of objects is determined according to the objective function. In the first step, the system space is determined. The environment contains a multidimensional coordinate system in the problem definition space. Every point of space is a solution to the problem. The search agents are a collection of objects. Each mass has four characteristics: (a) mass position, (b) active gravitational mass, (c) gravitational mass inactive (d) inertial mass. The above-mentioned objects are derived from the concepts of active gravitational mass and inertial mass in physics. Gravitational and inertial objects are defined inspired by the Newtonian physics concepts and are determined by the agility of the agents. Active gravity is a measure of the intensity of gravitational force around an object. Inactive gravitational mass represents the power of interaction in a gravitational field. The mass of inertia is also a measure of the resistance of the object against the change of location and movement. After the formation of the system, the rules governing it are determined. We assume that only the law of gravity and rules of motion dominate. The general form of these laws is almost similar to the laws of nature and are defined as follows. Imagine the system as a set of mass m. The position of each mass is the point of space, which is the answer to the problem. The position of the next d is shown from mass i with xdi. In this system at time t, for each mass i, by mass j in the direction d, a force is given as much as Fdij (t). The value of this force is calculated according to (1). Maj Major Gravity mass j, Mpi Mass gravity i and G (t) gravitational constant at time t and Rij are the distance between two masses i and j. Euclidean distance is used to determine the distance between objects according to (2).

$$\mathrm{F}_{\mathrm{ij}}^{\mathrm{d}}(\mathrm{t}) = \mathrm{G}(\mathrm{t})\frac{\mathrm{Ma_j(t) \times Mp_i(t)}}{\mathrm{R_{ij}(t)} + \varepsilon}(\mathrm{x}_{\mathrm{j}}^{\mathrm{d}}(\mathrm{t}) - \mathrm{x}_{\mathrm{i}}^{\mathrm{d}}(\mathrm{t})) \tag{1}$$

$$\mathrm{R_{ij}(t)} = \left\|\mathrm{X_i(t), X_j(t)}\right\|_2 \tag{2}$$

In (1), ε is a very small number. The force on mass i in the direction d is at the time t with Fdi (t), and is equal to the sum of all the forces that other bodies of the system impose on this mass. In (3), randj is a random number with uniform distribution in the interval [0, 1], which is considered to maintain the randomness of the search.

$$\mathrm{F}_{\mathrm{i}}^{\mathrm{d}}(\mathrm{t}) = \sum_{\mathrm{j=1,j\neq i}} \mathrm{rand_j(t)F_{ij}^{d}(t)} \tag{3}$$

According to Newton's second law, each mass is accelerated in the direction d, which is proportional to the force applied to the mass in that direction, divided by the mass of the mass inertia [relation (4)], in which the acceleration of mass i in the direction of dimension d in time t with adi (t) and the mass of mass inertia i with Mii (t) is shown [12].

$$\mathrm{a}_{\mathrm{i}}^{\mathrm{d}}(\mathrm{t}) = \frac{\mathrm{F}_{\mathrm{i}}^{\mathrm{d}}(\mathrm{t})}{\mathrm{M_{ii}(t)}} \tag{4}$$

$$\mathrm{V}_{\mathrm{i}}^{\mathrm{d}}(\mathrm{t}+1) = \mathrm{rand_i} * \mathrm{V}_{\mathrm{i}}^{\mathrm{d}}(\mathrm{t}) + \mathrm{a}_{\mathrm{i}}^{\mathrm{d}}(\mathrm{t}) \tag{5}$$

$$\mathrm{x}_{\mathrm{i}}^{\mathrm{d}}(\mathrm{t}+1) = \mathrm{x}_{\mathrm{i}}^{\mathrm{d}}(\mathrm{t}) + \mathrm{V}_{\mathrm{i}}^{\mathrm{d}}(\mathrm{t}+1) \tag{6}$$

The velocity of each mass is equal to the sum of the coefficients of the current velocity of mass and mass acceleration according to Eq. (5). The new position d of mass i is calculated according to (6) [2]. Randi is a random number with uniform distribution in the interval [0, 1], which is used to maintain the randomness of search, and $\mathrm{V}_{\mathrm{i}}^{\mathrm{d}}$ and $\mathrm{x}_{\mathrm{i}}^{\mathrm{d}}$ are the velocity and position of mass i in dimension d, respectively. Equation (7) is used to adjust the gravity coefficient. The gravity constant G is initially initialized and then reduced to over time to control the accuracy of the search [12].

$$\mathrm{G(t)} = \mathrm{G(G_0, t)} \tag{7}$$

Adjustment of the mass of the agents is done based on their target function, so that the factors with a better merit are attributed to a larger mass [Eq. (8)]. The mass size of the agents is normalized according to (9). Gravitational objects and inertia are considered equal to those in nature [Eq. (10)].

$$\mathrm{Mg_i} = \frac{\mathrm{fit_i(t) - worst(t)}}{\mathrm{best(t) - worst(t)}} \tag{8}$$

$$\mathrm{M_i(t)} = \frac{\mathrm{m_i(t)}}{\sum_{\mathrm{j=1}}^{\mathrm{N}} \mathrm{m_j(t)}} \tag{9}$$

$$\mathrm{M_{ij} = M_{gi} = M_{ai} = M_i} \tag{10}$$

In (8), fiti (t) represents the degree of fitness of mass i at time t. In minimization problems, we can use relationships (11) and (12) to calculate the best and worst.

$$\mathrm{best(t)} = \begin{matrix}\mathrm{minfit_j(t)} \\ \mathrm{j} \in (1,\ldots,\mathrm{N})\end{matrix} \tag{11}$$

$$\mathrm{worst(t)} = \begin{matrix}\mathrm{maxfit_j(t)} \\ \mathrm{j} \in (1,\ldots,\mathrm{N})\end{matrix} \tag{12}$$

At the beginning of the system, each object is randomly positioned in a point of space, which is the answer to the problem. At each moment of time, the objects evaluated then, the displacement of each crime is calculated after calculating the relations 1–12. System parameters include gravitational bodies, inertia mass and Newton gravity constant that are updated at each stage. The moratorium condition can be determined after a specified period of time [12].

3.2 Particle Swarm Optimization Algorithm

This algorithm is inspired by the collective movement of birds seeking food. A group of birds in the space are randomly looking for food. There is only one piece of food in the space discussed. None of the birds knows where food is. One of the best strategies can be to follow a bird that has the shortest distance to the food. This strategy is actually the source of the algorithm. The query space of the PSO algorithm is equivalent to the search space in the pattern of birds moving. Any solution called a particle in the PSO algorithm is equivalent to a bird and the number of particles (paths) is equivalent to the number of birds. Each particle has a suitability value calculated by a merit function, and the more particles in the search space reach the goal, the food is closer to the bird model, it is more qualified. Each particle also has a displacement that directs the motion of the particle and, with the aid of that, determines the next location of the particle. Each particle, by tracking the optimal particles in the current state, continues to move in the search space, in order to ultimately achieve the optimal response [13]. The main steps in the PSO algorithm are as follows:

- Initial preparation stage
- Repeat step

Initial preparation stage

At this stage, a population of particles is created, and for each particle a vector space and a displacement vector are considered. In the following, using the objective function, the target value of each particle is calculated and the best particle in terms of merit is maintained as the best global response.

Repeat step

In this step, the following operations will be repeated until one of the conditions for reaching the optimal answer or error rate is reached, or that the number of repetitions is completed. The operations performed at this stage are:

At each step of the repetition, the VI shift vector algorithm and the Xi location in the repetition (t + 1) of the relations (13) and (14) are determined:

$$\begin{aligned} \mathrm{V_i(t+1)} &= \omega\,\mathrm{V_i(t) + C_1 \times\ rand_1(pbest_i(t) - X_i(t)) + C_2} \\ &\quad \times\ \mathrm{rand_2(gbest_i(t) - X_i(t))} \end{aligned} \tag{13}$$

$$\mathrm{X_i(t+1) = X_i(t) + V_i(t+1)} \tag{14}$$

- Xi (t) is the current position of the i-th particle in the t again
- Vi (t) The current displacement vector of the i-th particle in the t-repeat
- pbesti (t) is the best place to experiment with i am to repeat t: gbesti (t) is the best place obtained between the neighbors of the particle i and the repetition t
- Rand1 and rand2: generate a random number between 0 and 1 [14].

3.3 Synthetic Bumblebee Algorithm

This algorithm was introduced in 2005 by Dervish Karbawa, which is based on the behavior of honey bees in finding optimum flowering trees for nectar gathering. An artificial bee colony algorithm is one of the algorithms based on collective intelligence and the result of the relationship between bees honeys [15]. In this algorithm, each bee alone cannot find the grass, but the collaboration and the exchange of information between a set of bees will lead to finding the appropriate bulb. In the colony algorithm, the artificial bee colony, community and colony of the bees are divided into three groups: recruited bees, search bees and watch bees. In this algorithm, each source of food represents a possible answer to the optimization problem, and the amount of nectar in each source represents the quality of that source. The number of honey bees or search bees is equal to the number of replies in the bee population. In the first step, a primary population is generated from the

SN response, which is the same as the food source location, where SN denotes the number of hired or searched bees. Each answer Xi (j = 1, 2,…, N) is a dimensional vector D, where D is the number of optimization parameters. The search bee selects a food source based on the probability. This choice is influenced by the quality of that food source. The probability of choosing each source is calculated by relation (15) [16].

$$p_i = \frac{fit_i}{\sum_{N=1}^{SN} fit_N} \tag{15}$$

In relation (15), fiti is the fitness value of the response i, and the selection of the new food source vij takes place according to the previous source of xij by relation (16).

$$v_{ij} = x_{ij} + \varphi_{ij}\left(x_{ij} - x_{kj}\right) \tag{16}$$

In the artificial colony algorithm, if a food source does not produce any improvement after the number of repeat steps, it is called an abandoned food source. The number of repeat steps is the limit parameter and is usually represented by the letter L. In this case, the observation bees will replace the new source by virtue of relation (17) and randomly.

$$\mathrm{x}_{\mathrm{i}}^{\mathrm{j}} = \mathrm{x}_{\mathrm{min}}^{\mathrm{j}} + \mathrm{rand}\left(\mathrm{x}_{\mathrm{max}}^{\mathrm{j}} - \mathrm{x}_{\mathrm{min}}^{\mathrm{j}}\right) \tag{17}$$

J is the number of optimization variables. In general, the optimization process can be summarized as follows:

- Creating a set of initial randomized responses to the search space for allocating bees to food sources (each bee is a source of food)
- Assigning search bees to food sources according to the amount of nectar
- Send watch bees to the search space to discover new food sources
- Remember the best food source
- Repeat steps above in such a way as to achieve the desired conditions.

4 The Proposed Method

A binary matrix n × n is used to model decision variables in the problem. In this matrix, the layers have the following concepts, definitions, and constraints:

- The xii represents the amount of the intersection on the matrix's original diameter. If xii = 0, that is, the iM node is in the sensor node network and does not serve any other node; otherwise, xii = 1 means that the node is a cluster and can serve sensor nodes.

- If xii = 0, node i is a sensor and cannot serve any other node, so for all xij in the i-th row, the values are zero.
- If xii = 1, the node is a cluster header and can serve other nodes. Therefore, in the i-th row of the matrix, one can find a column such as j, which is xij = 1, that is, the sensor node j receives the service node from the cluster node.
- In each row, the matrix xij $\leq$ xii always exists. In this situation, if xii = 0, in this case xij = 0 is necessary because node i is a sensor and cannot serve j. If xii = 1, in this case if xij = 0, that is, the sensor node j does not receive the service from the cluster node and if xij = 1, that is, the sensor node j receives the service node from the cluster node.
- The number of nodes in total is equal to n nodes and the number of nodes in the cluster is considered to be between 0.15n and 0.3n.

The longer the sensor node's distance from the cluster node is, the greater the power consumed to transmit the data from the sensor node to the cluster node, thus shortening network lifetime. Therefore, how to connect the nodes in proportion to the distance from the cluster head node, and the purpose of connecting the sensor node to the nearest node of the cluster is used to determine the distance between the cluster nodes i and the sensor node j from the Euclidean distance according to Eq. (18).

$$d(i,j) = \sqrt{(x_i - x_j)^2 + \left(y_i - y_j\right)^2} \tag{18}$$

(Xi, yi) is the coordinates of the node i, and (xj, yj) are the coordinates of the node j, and d (i, j) are the two knot spaces of each other. The cost of creating node nodes is also determined by the nodes, nodes, and node spacing. Therefore, relation (19) proposes the objective function of the problem.

$$f(x) = \text{Min} \sum_i \sum_j (C_{ik} + \alpha C_{kl} + C_{lj}) r_{ij} \tag{19}$$

In (19), Cik is the power consumption of the sensor node i to the cluster node k, and Ckl is the power consumption of the cluster node k to the cluster head node l, which consumes less power. Therefore, multiplied by a factor of less than 1 multiplied by α. Clj is the power consumption of connecting sensor node j to the cluster head node l. In the case of (19), the value of the parameter α is considered to be 0.7. Rij is also the number of requests between the two nodes i and j.

4.1 Using Evolutionary Algorithms for Problem Solving

In this section, we plan to examine the optimal clustering of sensor nodes using evolutionary algorithms. The algorithm studied in this section is the gravitational search algorithm, the artificial bison algorithm and the particle swarm optimization algorithm.

Results of the use of the gravitational search algorithm

In using the Gravitational Search Optimization algorithm to solve the problem, the population of the population is equal to 100, the frequency of the repetition of the algorithm is 400 and the gravitational constant is equal to 0.1. Table 1 presents the results of the implementation of the gravity search algorithm in terms of the best answer found, the average response in the last replication, the algorithm execution time, the number of nodes in the cluster, the number of sensor nodes, the number of assigned nodes, the number of assigned node nodes, Shows the precision and the time of convergence.

Figure 1 shows the convergence of the cost function and the mean cost function of population members in each repetition of the gravitational search algorithm.

Figure 2 shows the optimal clustering of sensor and cluster nodes in a wireless sensor network. The sensor nodes are blue and the nodes of the cluster are shown in red.

The results of using the particle swarm optimization algorithm.

In using the particle swarm optimization algorithm to solve the problem, the population of the population is 100, the frequency of the algorithm is 400, the weight of inertia 0.4, the personal learning coefficient 1.4962 and the national learning coefficient 1.3. Table 2 shows the results of implementing the particle swarm optimization algorithm in terms of the best answer found, the average response in the last replication, the execution time of the algorithm, the number of nodes in the cluster, the number of sensor nodes, the number of assigned nodes, the number of missed assignment nodes Findings show the accuracy and time of convergence.

Figure 3 shows the method of convergence of the cost function and the mean cost function of population members in each repetition of the particle swarm optimization algorithm.

Figure 4 shows the optimal clustering of sensor nodes and cluster nodes in the sensor network. The sensor nodes are blue and the nodes of the cluster are shown in red.

Results of the use of artificial beehive algorithm

In using the artificial bison algorithm to solve the problem, the population of the population is 100, the frequency of the algorithm is equal to 400, the number of test bees is equal to 100, the parameter of the limit (idle parameter) is 30 and the range of motion 1 is considered. Table 3 presents the results of the implementation of the artificial bison algorithm in terms of the best answer, the average response in the

Table 1 Evaluation of the results of using the gravity search algorithm

Best cost	Mean cost	Time	P	Node	True positive	False negative	Accuracy	Convergence
$10^7 \times 1.5757$	$10^7 \times 1.5757$	209	5	25	25	0	100%	216

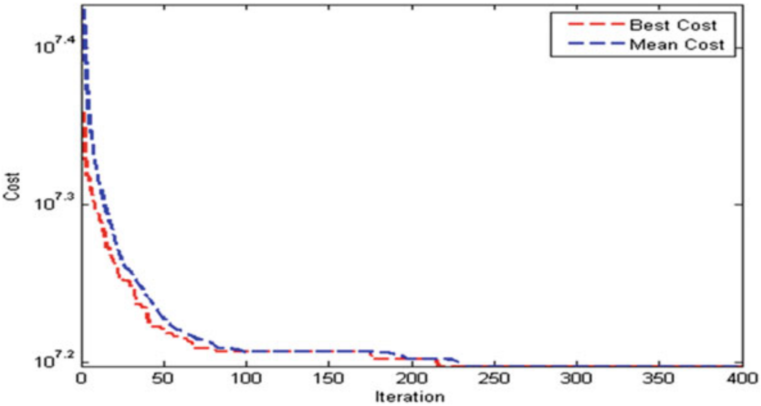


Fig. 1 The convergence method of the gravitational search algorithm

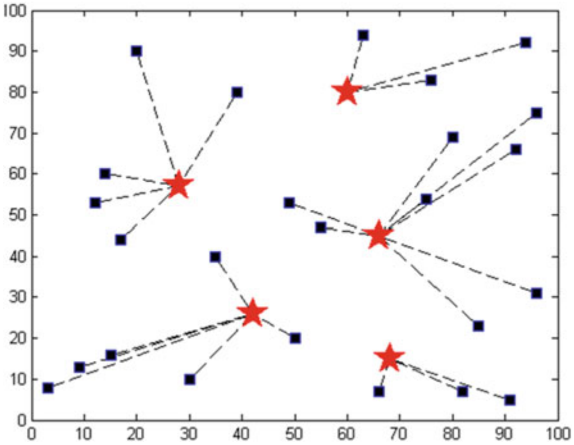


Fig. 2 Optimal clustering of sensor and cluster nodes in a sensor network in a gravitational search algorithm

Table 2 Evaluation of the results of using the particle swarm optimization algorithm

Best cost	Mean cost	Time	P	Node	True positive	False negative	Accuracy	Convergence
$10^7 \times 1.6246$	$10^7 \times 1.6246$	210	5	25	24	1	96%	274

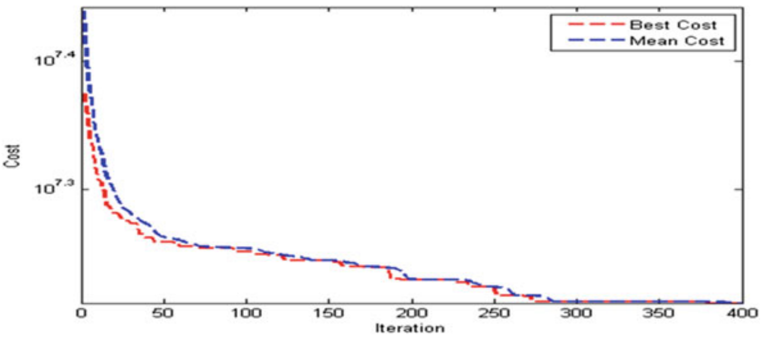


Fig. 3 Convergence mode of particle swarm optimization algorithm

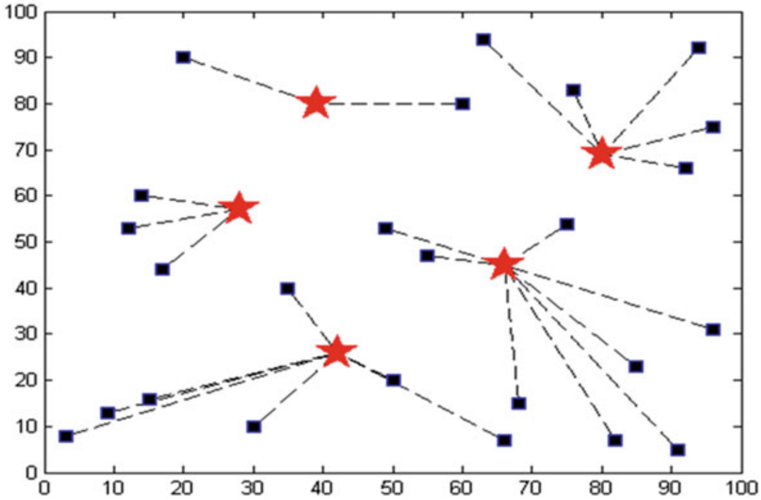


Fig. 4 Optimal clustering of sensor and cluster nodes in a sensor network in an optimization algorithm for particle swarm

Table 3 Evaluation of the results of the use of artificial beehive algorithm

Best cost	Mean cost	Time	P	Node	True positive	False negative	Accuracy	Convergence
10^7 × 1.6544	10^7 × 1.6544	215	5	25	23	2	92%	257

last replication, the algorithm execution time, the number of nodes, the number of sensor nodes, the number of assigned nodes, the number of assigned nodes, Accuracy and time of convergence show.

Figure 5 shows the convergence of the cost function and the average cost function of population members in each repetition of the artificial beehive algorithm.

Figure 6 shows the optimal clustering of sensor nodes and cluster nodes in the sensor network. The sensor nodes are blue and the nodes of the cluster are shown in red.

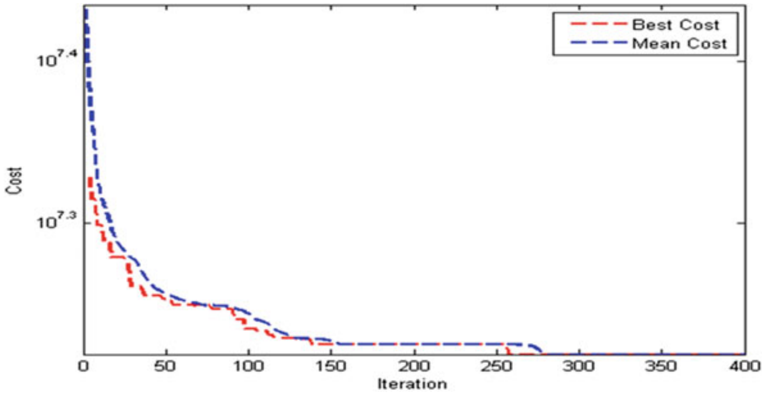


Fig. 5 Synthesis method of artificial beehive algorithm

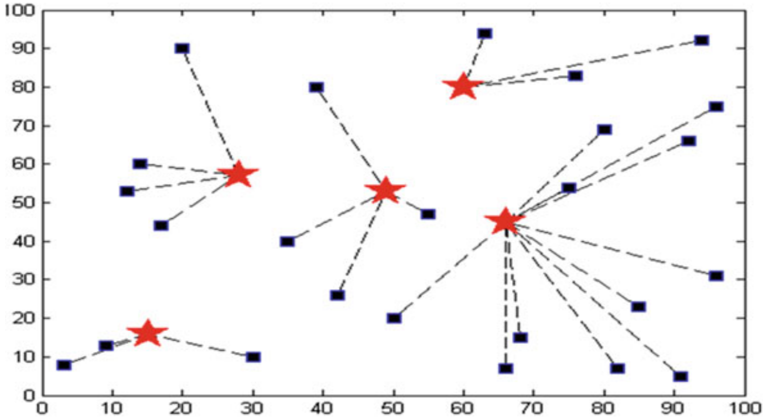


Fig. 6 Optimal clustering of sensor and cluster nodes in the sensor network in the artificial beehive algorithm

5 Conclusion

This paper is based on the optimal clustering of sensor nodes in order to reduce the power consumption of sensor nodes through the optimal connection of sensor nodes to cluster nodes. In the second section, we proceeded with the research into wireless sensor networks and optimal clustering methods for sensor nodes, and we used the algorithms and evolution algorithms to solve the clustering problem and the nodes in sensor networks. And benefited. In the third section, we introduced evolutionary algorithms. Gravitational optimization algorithms, particle swarm optimization and artificial insemination were investigated. In the fourth section, we were asked to model and implement the problem, and quantitative and qualitative results of the use of evolutionary algorithms were presented. The results indicate that the particle swarm optimization algorithm with a power consumption of $10^7 \times 1.6246$, a runtime of 210 s, and a 96% accuracy of sensor cluster clustering, perform better than the artificial bucket algorithm with power consumption of $10^7 \times 1.6544$, Runtime is 215 s and accuracy is 92%. The gravity search algorithm with power consumption of $10^7 \times 1.5757$, run time 209 s, and accuracy of 100% clustering of sensor nodes, performed in terms of power consumption, runtime, and accuracy better than particle swarm optimization algorithm and artificial beehive algorithm. It is chosen as the best algorithm for solving the optimal clustering problem of nodes in the sensor network to reduce the power consumption of the network through the optimal connection of the nodes to the node nodes.

References

1. Neto JHB, Rego A, Cardoso AR, Celestino J Jr (2014) MH-LEACH: a distributed algorithm for multi-hop communication in wireless sensor networks. ICN 2014, 55–61
2. Rajagopalan R, Varshney PK (2006) Data aggregation techniques in sensor networks: a survey
3. Liu X, He D (2014) Ant colony optimization with greedy migration mechanism for node deployment in wireless sensor networks. J Netw Comput Appl 39:310–318
4. Cao C, Ni Q, Yin X (2014) Comparison of particle swarm optimization algorithms in wireless sensor network node localization. Paper presented at the 2014 IEEE international conference systems, man and cybernetics (SMC)
5. Brar GK, Virk AK (2014) Deployment of nodes for maximum coverage in heterogeneous wireless sensor network using genetic algorithm. Int J 2(6)
6. Zou Z, Lan Y, Shen S, Wang R (2014) Node localization based on optimized genetic algorithm in wireless sensor networks. In: Advances in wireless sensor networks
7. Barolli A, Oda T, Ikeda M, Barolli L, Xhafa F, Loia V (2014) Node placement for wireless mesh networks: analysis of WMN-GA system simulation results for different parameters and distributions. J Comput Syst Sci
8. Peng B, Li L (2015) An improved localization algorithm based on genetic algorithm in wireless sensor networks. Cogn Neurodyn 9(2):249–256
9. Nagchoudhury P, Maheshwari S, Choudhary K (2015) Optimal sensor nodes deployment method using bacteria foraging algorithm in wireless sensor networks. Paper presented at the

emerging ICT for bridging the future-Proceedings of the 49th annual convention of the computer society of India CSI Volume 2

10. Sun X, Zhang Y, Ren X, Chen K (2015) Optimization deployment of wireless sensor networks based on culture–ant colony algorithm. Appl Math Comput 250:58–70
11. Shilane D, Martikainen J, Dudoit S, Ovaska SJ (2008) A general framework for statistical performance comparison of evolutionary computation algorithms. Inf Sci 178(14):2870–2879
12. Rashedi E, Nezamabadi-Pour H, Saryazdi S (2009) GSA: a gravitational search algorithm. Inf Sci 179(13):2232–2248
13. Clerc M, Kennedy J (2002) The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE Trans Evol Comput 6(1):58–73
14. Clerc M (2012) Standard particle swarm optimisation
15. Karaboga D, Akay B (2009) A comparative study of artificial bee colony algorithm. Appl Math Comput 214(1):108–132
16. Milan T (2013) Artificial Bee Colony (ABC) algorithm with crossover and mutation. Appl Soft Comput, pp 687–697
17. Goyal RA (2015) Review on energy efficient clustering routing protocol in wireless sensor network

A Bee Colony (Beehive) Based Approach for Data Replication in Cloud Environments

Saedeh khalili azimi

Abstract Cloud computing refers to applications and services which run on a distributed network, they also represent the ability of information technology as a service to network users. Data replication is an important way of managing mass data in a distributed manner; data replication is seen as one of the important issues in distributed systems which are usually undertaken for increasing the efficiency, availability, and security of information. Data replication's core idea is developing methods for putting repetitions in different places, so that there are multiple iterations of the specific file in different sites. A key issue in data managing is the manner that system deals with duplicates. This included steps such as: which files are replicated, when is the data replication done and where in the system are they to be placed. In this study, the proposed method has been implemented using the MATLAB environment and the results have showed that the performing time of the proposed method is much lower as compared to previous methods and it had improved performance time compared to the previous methods.

Keywords Replicated · Distribution network · Efficiency · Availability
Distributed network

1 Introduction

Cloud computing refers to programs and services that are implemented in a developed network and use virtual references, which are also available through usual internet protocols and network standards. its concept lies in the fact that in cloud computing used references are unlimited and represent characteristics of the hardwares (which the softwares are operating on them) are completely independent of the user system. Data replication is one of the important issues in distributed

S. khalili azimi (✉)
Department of Computer Science and Engineering,
Islamic Azad University, Tabriz Branch, Tabriz, Iran
e-mail: s.khaliliazimi@gmail.com

S. Montaser Kouhsari (ed.), *Fundamental Research in Electrical Engineering*, Lecture Notes in Electrical Engineering 480,
https://doi.org/10.1007/978-981-10-8672-4_80

systems which is usually undertaken for increasing the efficiency, availability, and security of information. Managing compatibility is the problem of managing/ calibrating the compatibility between dubbed versions. This refers to the fact that every change which occurs in the dubbed version must also be undertaken in the copy version. If not, the dubbed versions would not be the homogenous and therefore nullify the very concept of replication. In most of the capability models, it is assumed that target processes simultaneously allot availability to one common reference. Here capability refers to the fact that every process knows that in that moment other processes allot availability to the same data reference (either reading or changing) as it is generally assumed. In large scale distribution systems implementing capability models for common data in an efficient way was difficult [1]. Data replication is a tool for achieving high availability and managing errors in systems and distributed databases. Achieving high availability is one the most important issues for users. Data replication is composed of replication selection and replication placement stages. In the replication selection stage, it is decided which file needs to be replicated. And then, in the replication stage the best replication site for to the data chosen according to the intended work. The problem of this stage is selecting the best replication site so that the users will be able to access data through operation. The replication selection guideline chooses a site for recycling data based on limitations [2]. The best possible node is selected on the basis of network delay time and user requests for copying data. The replicated version of data is placed in the most appropriate node. If the best site does not have enough space for saving the new replication, the mechanisms/functions for freeing up space for the data replication replacement becomes activated [3]. In the replication replacement stage the replication version of data is placed in the most convenient node which is based on network delay time and user requests for copying data. By the time which replication replacement decides where and when the copy will be placed [4]. If it is not supposed to replicate, the file will be read from another place. In data replication the appropriate place for accessing data for the intended work is chosen. These two stages are combined to detect the best site for replication selection and placement [1]. Park et al. represented a replication algorithm based on internet hierarchy known as replication based on hierarchical bandwidth. In which, sites are grouped according to the suggested algorithm in different areas in a way that the sites inside the area are linked with high bandwidth and the two other areas of the sites are linked by low bandwidth [5]. This method consists of three stages of hierarchical replication. The hierarchical replication method places replications on sites, and bandwidth is its key parameter for replication selection and removing files. When a work is timed a replication manager must transfer all the requested files which are not available before performing the work. Therefore, data replication efficiency increases through decreasing workflow time [6]. Dynamic hierarchical replication method predicts future needs of the network and their replication is observable before their request. This prediction is done by paying attention to hierarchical access to past files and subsequently, responding time, delay time, this in turn considerably decreases bandwidth consumption [7, 8]. In this algorithm after allocating the work to the appropriate site by timing, any file which is requested by work and is unavailable on the local site should first be sent to the demanding site and then replicated. By having created different replications and through emphasizing attention on the large size of the files and thelong file transferring time [9, 10]. Then selecting the appropriate site gains more urgency regarding a distribution system's efficiency. It is better to select the appropriate replication for transferring to a local site by using the hierarchical method. First it should be investigated if there is a replication of the file in local sites or not [7, 11, 12].

2 Suggested Method

The suggested algorithm is the new improved form of a dynamic hierarchical replication algorithm. This algorithm has been implemented in Matlab environment. In the suggested method, the bee algorithm was used for both the selection and the replacement of data replication. As described in the following section.

2.1 Artificial Bee Colony (ABC) Algorithm

In the artificial bee colony algorithm, for the first iteration, half of the bees function as workers and half of them act as guardians. For every food recourse there is only one worker bee. In other words the number of worker bees is equal to the number of available food around the colony [13, 14]. In the artificial bee colony algorithm, every cycle of searching consists of three stages. First sending the worker bee to food recourses and then secondly evaluating the amount of their nectar, Thirdly food recourse selection is done by guardian bees after sharing information with worker bees and evaluating the nectar of foods, determining the destination of the bees and sending them to food recourses. In the initial evaluating stage a complex of food recourse positions which are selected randomly by worker bees are determined. Afterwards, these worker bees go back to the colony and share nectar information of every recourse with the bees which are waiting in the dance region inside the colony. In the second stage after sharing information every worker bee goes to food recourse within a region that the bee itself has seen in past cycle and that recourse is in its memory and then a new food recourse becomes selected [13, 15].

2.2 Suggested Algorithm

The hierarchical topology used for organizing the intended system is shown in Fig. 1. Its structure is made of three levels.

The first level includes clusters of personal computers and these nodes are connected together by a high bandwidth in a local network which is called the local

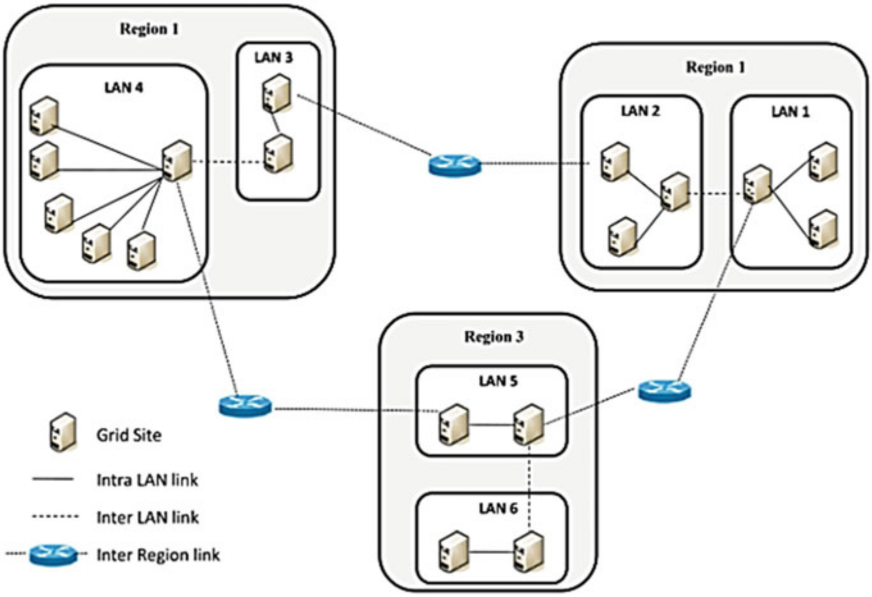


Fig. 1 Network topology

area network level. This region is the next level in which every region includes local area network that is connected together with lower bandwidth in comparison to first level. In the suggested algorithm after allocating work to a convenient site which is done by using timing, every file which is requested by work and isn't available in the local site, first should be sent to the demandant site, then replicated. By having different replications in a cloud environment and through emphasizing attention on the large size of the files and the long file transferring time. It should be checked initially, if there is a copy of the file on this site, where the job is located, or not. If the file is not on the intended site, in the next stage the sites which are in the local area network will be checked using the bee algorithm. If there was a copy in the local area network the best copy amount of all the copies will be selected by the bee algorithm and will be transferred to the demandant site. If there is no copy of the intended file in the local area network, in the next stage the copy should be searched in other local area networks by implementing the bee algorithm. Also among the existing copies the best one should be selected and transferred to the demandant site. In the next stage, in case of not finding the requested file in the intended other local area networks. Other regions should be searched and in case of finding the intended file it should be then transferred to the demandant site. After finding the intended file the next stage is finding best site for saving the copy. The newly suggested algorithm is in fact the improved form of a dynamic hierarchical algorithm. This algorithm was implemented in the Matlab environment. Table 1 shows the simulation parameters. Taking this point to the consideration is important as expansibility problems can be solved and system efficiency may become

Table 1 Simulation parameters

Parameter	Value
Number of works	1500
Size of every file (GB)	1
Maximum size of site (GB)	50
Work delay (ms)	2500

increased by data replication. But keeping the compatible copies need to be synchronized, which is costly in terms of efficiency. For decreasing the costs, in most of the represented algorithm the files are considered to be read only files. Some sites which represent the food recourses in the bee colony algorithm have been considered for simulation, so that the suggested algorithm could be implemented in the Matlab environment. In the site selection section for placing intended copies the bee colony algorithm was used. In these sites the possibility of the files' existence is provided randomly. A history of request numbers of every file is kept in every site. Worker bees place existing sites in three levels, local, region, and other regions. In the bee colony algorithm if new positions or new food regions have better quality or more nectar, the bee stays in new region if they don't come back from the past region and one unit will be added to its trial index. Here, our meaning of quality is the fitness of solutions that is the value of target function. f\Function fitness is achieved by paying attention to connected environments as it is shown below:

$$f(x) = \sum_{i=1}^{n} x_i^2 \tag{1}$$

Which in it x_i represent existing neighborhoods around food recourse. By paying attention to existing data, this function is considered as a formula (1) the probability of existing files in sites gets included in evaluating the fitness of that site. The more likely the file is to be in a site the more chance that site has to be selected as the best site for replication. This function in the suggested method is calculated as following:

$$z = \mathrm{sum}(x(i) * f(i)^2) \tag{2}$$

The $f(i)$ is the probability of a file's existence in a site and x_i represents neighborhoods around a food recourse. After the completion of searching stages by worker bees for selecting the best site, by paying attention to the number of requests for a file, a convenient site will be selected and after confirming its superiority by the guardian bees of that intended site, by paying attention to its fitness, it will be selected for replication transferring. In the next stage the selected file is place in the best site. The more the number of the requests for a site are, the more likely it is to be selected as the best place for replacing the intended replication and the site, that has reviewed more requests for the intended file will be selected as the best site for saving the file replication. For saving a new replication first the feasibility of replication in the selected site should be checked by paying attention to this point that limited saving space is purposed for every site, if the size of the file is more than the available space, then file replication is impossible and the file should be accessed indirectly by the time of replication, if there is not enough space for saving, the new file will be substituted for the file that has received the least amount of requests for it, is recorded. while the information about the amount of requests are recorded in the past stage and they can also be used in this stage. This algorithm was tested in three stages and the results are discussed below.

2.2.1 Implementing Suggested Algorithm

The flowchart of the suggested algorithm of this stage is shown in Fig. 2. And in the configuration of the implemented network the structure of bandwidth configuration is illustrated in Table 2. In this stage the number of requests for the intended file from every site should be clarified. First, the number of requests must be checked among local sites and then sites of other regions are searched by the bee colony algorithm. The site which has the most requests for the intended file will be selected as the best site for saving. In the last stage after finding the best place for saving the new replication, first the feasibility of replication in the selected site will be checked, if the size of the file is more than the size of saving space, then it is not possible to replicate the file and the file should be accessed indirectly. but if there is enough space on the site, then the file will be replicated otherwise, if there is a replicate of the file in the local area network, respectively the file will not get replicated. incoherent But if the file is to be replicated, other files should be removed and of course the files that should be removed may not have any replicates in this local area network and in case of a need in the future it will take a long time to transfer them from other local area networks. Therefore, it is necessary to find better ways for replacing files. The site which has the most requests for the intended file is selected as the best site for saving the file's replicate. By the time of replication, in case of not having enough free space for saving the new file, the new file will be substituted for the file which has the least number of requests is recorded for it. The flowchart of this stage is demonstrated in Fig. 3.

3 Results Evaluating

In this section the simulation and bandwidth parameters are defined. Both Tables 1 and 2 are related to these values, as shown in the following section.

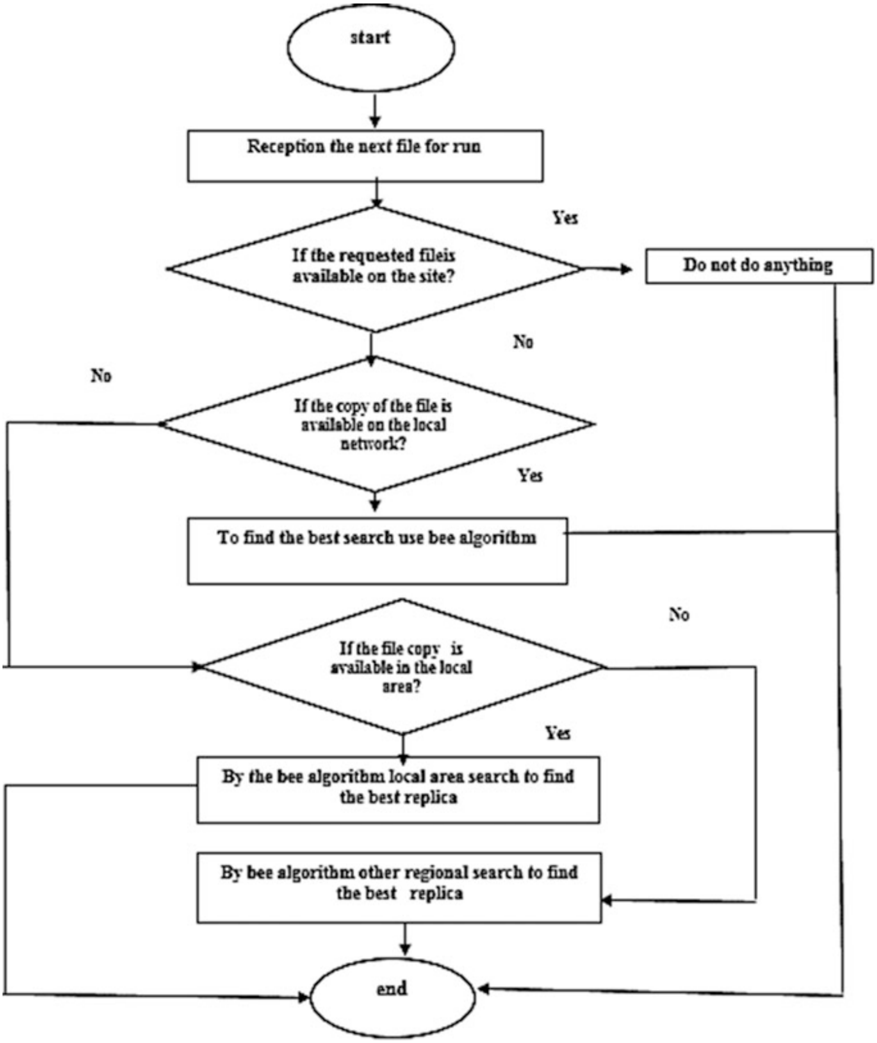

start
Reception the next file for run
If the requested fileis available on the site?
Yes
Do not do anything
No
If the copy of the file is available on the local network?
No
Yes
To find the best search use bee algorithm
No
If the file copy is available in the local area?
Yes
By the bee algorithm local area search to find the best replica
By bee algorithm other regional search to find the best replica
end

Fig. 2 Replication selection trend in suggested algorithm

Table 2 Band width configuration

Parameter	Value
Band width of local area network	1000
Band width between local area networks	100
Band width of regions	10

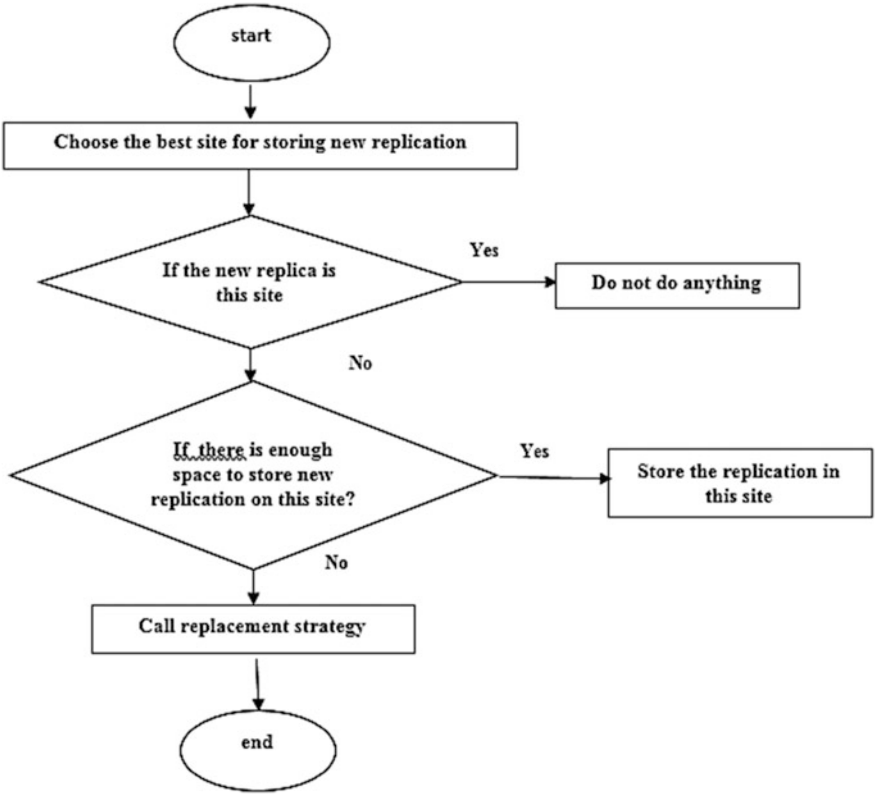

start
Choose the best site for storing new replication
If the new replica is this site
Yes
Do not do anything
No
If there is enough space to store new replication on this site?
Yes
Store the replication in this site
No
Call replacement strategy
end

Fig. 3 Replication replacement trend in suggested algorithm

3.1 Simulation Results

The suggested algorithm was implemented in the Matlab environment by using defined datasets. According to the results the performing time of this algorithm is at 1600 s, this proves that the performance time has decreased in comparison to past methods. the problem of using memory and the challenge of memory occupation which is one of the problems of data replication, somewhat decreases because in the suggested method the suitability of sites for selection of replication is evaluated according to the fitness function and there was no need to keep lists of sites which have a copy of the intended file, also in the stage of file replacement in sites for removing extra files for freeing up space. there was no need to make a queue and to keep the information of files and sites. In the represented charts the suggested algorithm has been tested on 1500 works and the results have been revealed in Figs. 4, 5, 6 and 7.

The present results have indicated that the time of performance has relatively decreased in comparison to past methods. The bar chart determines the mean

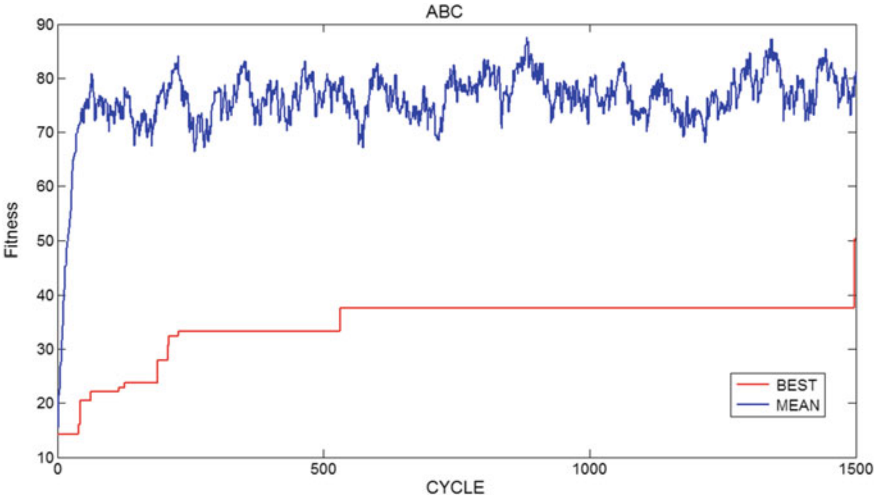


Fig. 4 Simulation results based on bee colony algorithm

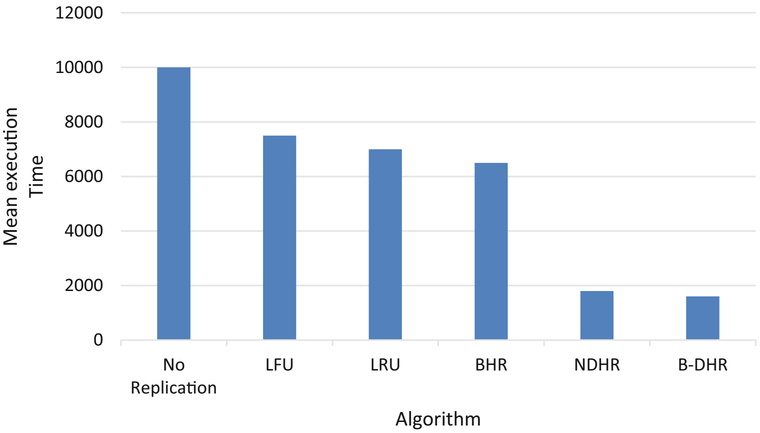


Fig. 5 Mean job time for different replication algorithm

performing time of some different replication algorithms. The next linear chart indicates change in a number of works over different methods, in which we can see a relative decrease in performing time in both charts. In the suggested algorithm two important points were considered. The first point is when the different sites have a copy of the intended file, in this time selection of the best copy among them is a very effective step in optimizing efficiency. The replicate selection algorithm selects the best replicate by using the bee algorithm and using this algorithm has led to increased efficiency and accessibility to the best possible site as a result of crossing the local optimums. The second stage is when there is not enough space for saving

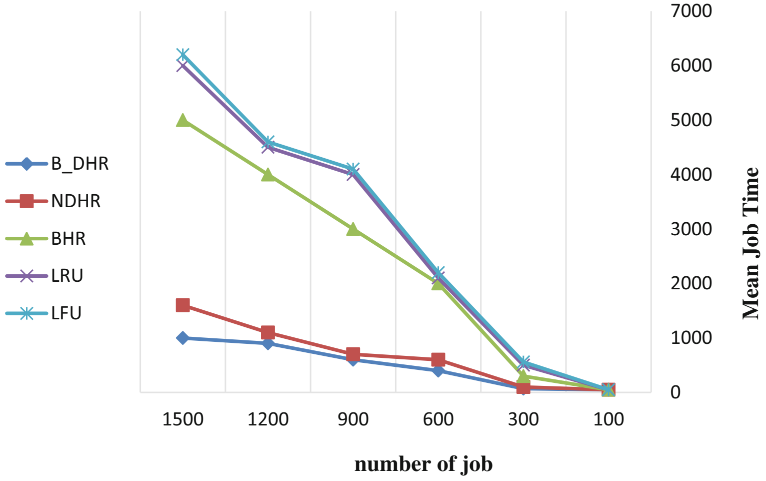


Fig. 6 Mean job time by changing number of works

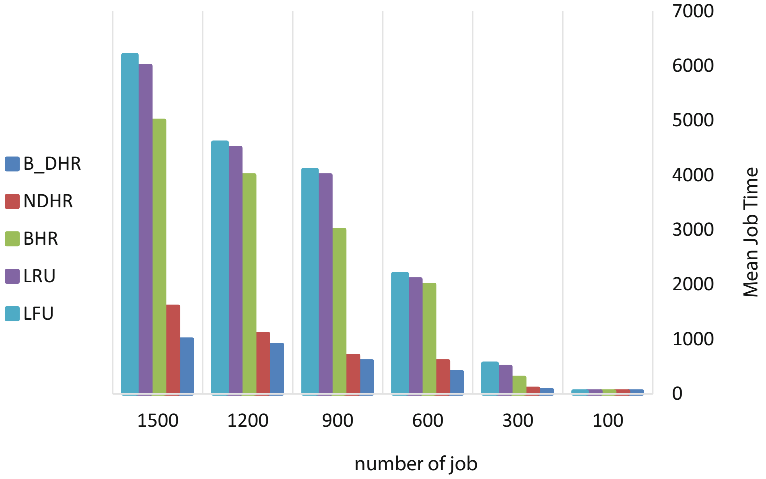


Fig. 7 Mean job time by changing number of works

the new replicate therefore some of the replicates should be removed. The suggested algorithm chooses the file for elimination that has the amount least of requests is recorded for it, and since, there is no need to queue then the site's saving list became relatively more efficient. Also memory usage in this method had somewhat decreased in contrast with other methods. Moreover, the problem regarding removing files which will be needed in future has been already solved in this method (Figs. 8, 9 and 10; Tables 3, 4 and 5).

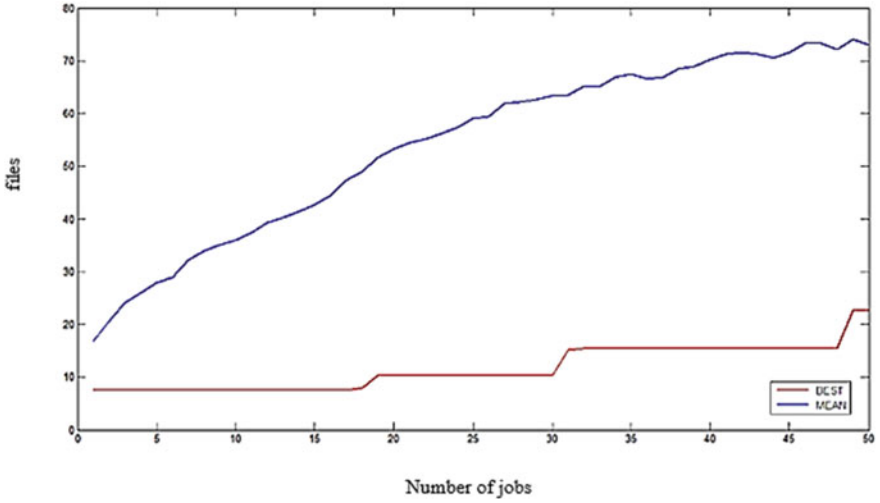


Fig. 8 Simulation results based on bee colony algorithm with different simulation parameters in test 1

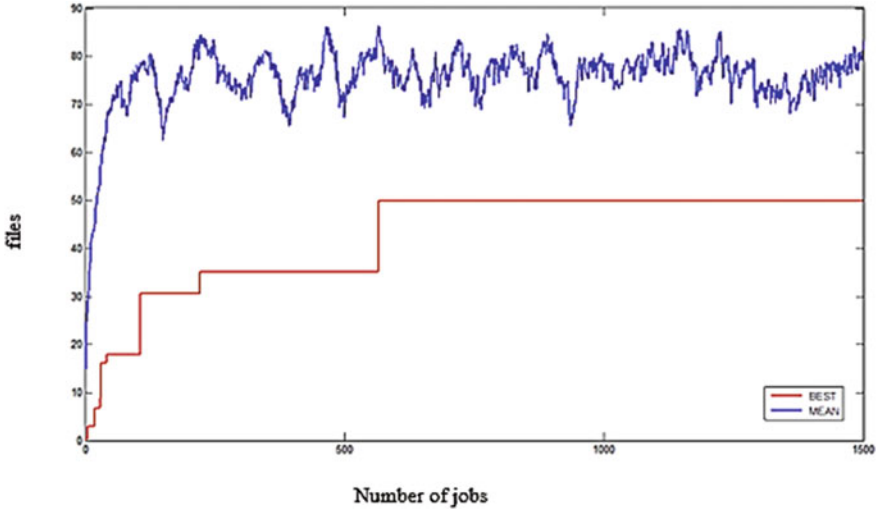


Fig. 9 Simulation results based on bee colony algorithm with diffrent simulation parameters in test 2

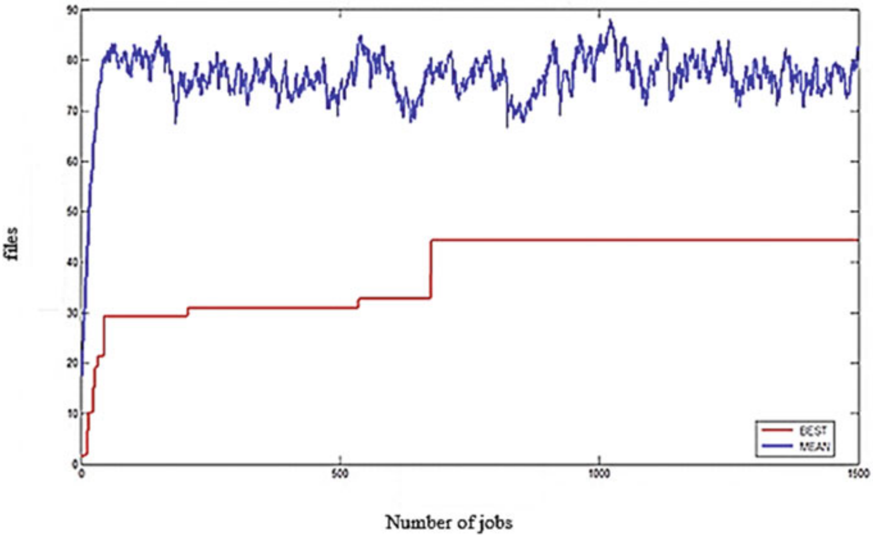


Fig. 10 Simulation results based on bee colony algorithm with diffrent simulation parameters in test 3

Table 3 Simulation parameters for test 1

Parameter	Value
Number of works	50
Size of every file (GB)	1
Maximum size of site (GB)	50
Work delay (ms)	2500

Table 4 Simulation parameters for test 2

Parameter	Value
Number of works	1500
Size of every file (GB)	1
Maximum size of site (GB)	10
Work delay (ms)	2500

Table 5 Simulation parameters for test 3

Parameter	Value
Number of works	1500
Size of every file (GB)	10
Maximum size of site (GB)	50
Work delay (ms)	2500

References

1. Jeon M, Lim K, Ahn H, Le B (2010) Dynamic data replication scheme in the cloud computing environment. In: IEEE Second Symposium on network cloud Computing and Applications, vol 3. pp 40–47
2. Teng M, Junzhou L (2005) In: International conference on advanced information and applications
3. Mansouri N, Dastghaibyfard G (2012) A dynamic replica management strategy in data grid. J Netw Comput Appl 35:1297–1303
4. Zaha W, Xu X, Wang Z (2008) A weight-based dynamic replica replacement strategy in data grids. IEEE Asia-pacific service computing conference
5. Park S, Kim J, Ko Y, Yoon W (2004) Dynamic data replication strategy based on internet hierarchy. Springer-Verlag, Berlin, pp 838–846
6. Mansouri N, Dastghaibyfard G, Mansouri E (2013) J Netw Comput Appl 36(2):711–722
7. Rahmani A, Fadaie Z, Chronopoulos AT (2015) Data placement using dewey encoding in a hierarchical data grid. J Netw Comput Appl 88–98
8. Sakr S, Liu A, Batista D, Alomari M (2010) A survey of large scale data management approaches in cloud environments. IEEE communications surveys & tutorials, accepted for publication, Manuscript received 28 Jun 2010; Revised 2 December
9. Sato H (2008) In: International conference on grid computing, vol 1, pp 250–257
10. Abdurrab A, Xie T (2010) In: International conference in clustering computing and the grid, vol 10, pp 215–223
11. Taheri J, Lee Y, Zomaya A, Siegel H (2013) A bee colony based optimization approach for simultaneous job scheduling and data replication in grid environments. Comput Oper Res 1564–1578
12. Sashi K, Thanamani A (2011) Dynamic replication in a data grid using a modified BHR region based algorithm. Future Gener Comput Syst 202–210
13. Xu X, Liu Z, Wang Z, Sheng Q, Yu J, Wang X (2017) S-ABC: a paradim of service domain-oriented artificial bee colony algorithm for service selection and composition. Future Gener Comput Syst 68:304–319
14. Mansouri N, Dastghaibyfard G (2013) Enhanced dynamic hierachical replication and weighted scheduling strategy in data grid. J Parallel Distrib Comput 534–543
15. Gao K, Suganthan P, Pari Q, Tasgetirn M, Sadollah A (2016) Knowl-Based Syst 109:1–16
16. Horri A, Sepahvand R, Dastghaibyfard G (2008) IJCSNS Int J Comput Sci Netw Secur 8:8
17. Camman M, Stockinger K (2002) In: International symposium on cluster computing and the grid, pp 340–345
18. Sashi K, Santhanam T (2013) ARPN J Eng Appl Sci 8(2)
19. Abawajy J, Member S (2014) IEEE Trans comput 63:2975–2987
20. Grossman R, Gu Y, Mambretti J, Sabala M, Szalay A, White K (2010) An overview of the open science data cloud, HPDC’10, Chicago, Illinois, USA
21. Kingsy R, Manimegalai R (2014) J Parallel Distrib Comput 74(2):2099–2108
22. Tanenbaum AS, van Steen M (2006) Distributed systems: principles and paradigms. Vrije Universitet Amesterdam, The Netherlands
23. Mansouri N (2016) Adaptive data replication strategy in cloud computing for performance improvement. Front Comput Sci
24. Kumar KA, Quamar A, Deshpande A, Khuller S (2014) SWORD: workload aware data placement and replica selection for cloud data management systems. VLDB J 23(6):845–870
25. Boru D, Kliazovich D, Granelli F, Bouvry P, Zomaya AY (2015) Energy efficient data replication in cloud computing datacenters. J Cluster Comput 18(1):385–402

26. Janpet J, Wen YF (2013) Reliable and available data replication planning for cloud storage. In: Proceedings of the IEEE 27th international conference on advanced information networking and applications (AINA), pp 772–779
27. Leesakul W, Townend P, Xu J (2014) Dynamic data reduplication in cloud storage. In: Proceedings of the IEEE 8th international symposium on service oriented system engineering (SOSE), pp 320–325
28. Huang K, Li D, Sun Y (2014) CRMS: a centralized replication management scheme for cloud storage system. In: Proceedings of the IEEE/CIC international conference on communications in China (ICCC), pp 344–348