

Journal of Electrical and Computer Engineering Innovations

Vol. 9; Issue 1: 2021

Table of Contents

A Transformer Self-Attention Model for Time Series Forecasting <i>R. Mohammdi Farsani, E. Pazouki</i>	1
Target Detection Using Multispectral Images, a Case Study: Wheat Detection in Chenaran County in Iran <i>M. Imani</i>	11
Improving the Security of a Low Cost Tag Search Protocol <i>S. Saderi Oskuiee, F. Moazami, G. Oudi Ghadim</i>	25
A Novel Method for Medical Image Segmentation Based on Convolutional Neural Networks with SGD Optimization <i>M. Taheri, M. Rastgarpour, A. Koochari</i>	37
Deep Neural Network with Extracted Features for Social Group Detection <i>A. Akbari, H. Farsi, S. Mohamadzadeh</i>	47
Fault Detection in Thermoelectric Energy Harvesting of Human Body <i>H. Yektamoghadam, A. Nikoofard</i>	57
NodeFetch: High Performance Graph Processing Using Processing in Memory <i>M.A. Mosayebi, M. Dehyadegari</i>	67
A New Routing Protocol in MANET using Cuckoo Optimization Algorithm <i>S. Tabatabaei, H. Nosrati Nahook</i>	75
Optimal PSS Parameters Design Based on a Novel Objective Function for Small Signal Stability Enhancement <i>P. Naderi, B. Ehsan Maleki, H. Beiranvand</i>	83
Fast and Efficient Hardware Implementation of 2D Gabor Filter for a Biologically-Inspired Visual Processing Algorithm <i>A. Mohammadi Anbaran, P. Torkzadeh, R. Ebrahimpour, N. Bagheri</i>	93
Modelling and Optimization of Channel Allocation for Power Line Communications Access Networks in the Presence of In-Line and In-Space Interference <i>M. Sheikh-Hosseini, S.M. Nosratabadi</i>	103
Stochastic Block NIHT Algorithm for Adaptive Block-Sparse System Identification <i>Z. Habibi, H. Zayyani, M.S.E Abadi</i>	115



Research paper

A Transformer Self-Attention Model for Time Series Forecasting

R. Mohammadi Farsani, E. Pazouki*

Artificial Intelligence Department, Faculty of Computer Engineering, Shahid Rajaei Teacher Training University, Tehran, Iran.

Article Info

Article History:

Received 14 May 2020
Reviewed 07 July 2020
Revised 18 September 2020
Accepted 21 November 2020

Keywords:

Time Series Forecasting (TSF)
Self-attention model
Transformer neural network

*Corresponding Author's Email Address:
ehsan.pazouki@sru.ac.ir

Abstract

Background and Objectives: Many real-world problems are time series forecasting (TSF) problem. Therefore, providing more accurate and flexible forecasting methods have always been a matter of interest to researchers. An important issue in forecasting the time series is the predicated time interval.

Methods: In this paper, a new method is proposed for time series forecasting that can make more accurate predictions at larger intervals than other existing methods. Neural networks are an effective tool for estimating time series due to their nonlinearity and their ability to be used for different time series without specific information of those. A variety of neural networks have been introduced so far, some of which have been used in forecasting time series. Encoder decoder Networks are an example of networks that can be used in time series forecasting. an encoder network encodes the input data based on a particular pattern and then a decoder network decodes the output based on the encoded input to produce the desired output. Since these networks have a better understanding of the context, they provide a better performance. An example of this type of network is transformer. A transformer neural network based on the self-attention is presented that has special capability in forecasting time series problems.

Results: The proposed model has been evaluated through experimental results on two benchmark real-world TSF datasets from different domain. The experimental results states that, in terms of long-term estimation Up to eight times more resistant and in terms of estimation accuracy about 20 percent improvement, compare to other well-known methods, is obtained. Computational complexity has also been significantly reduced.

Conclusion: The proposed tool could perform better or compete with other introduced methods with less computational complexity and longer estimation intervals. It was also found that with better configuration of the network and better adjustment of attention, it is possible to obtain more desirable results in any specific problem.

Introduction

A time series is a sequence of data that is usually collected at consecutive times. The time series is defined by

$$x(t), t = 0, 1, 2, \dots \quad (1)$$

where t denotes the elapsed time. Forecasting the time series means estimating the future of the series

according to its past (see in [1]).

In fact, time series forecasting means estimating the time series value in interval $t:T$ based on the past known interval $1:t-1$. Three common methods for forecasting time series are: Random models, Support Vector Machine (SVM) models and Artificial Neural Network models [1]. Autoregressive Integrated Moving Average (ARIMA) model is one of the most popular and

widely used random models which is used as a base model in most time series research (see in [2], [3] and [4]). This model assumes that the data are linear and follows a certain probability distribution such as normal. However, this model supports wide variety of time series problems. One of the important tools for solving forecasting time series problem is artificial neural network (ANN). Since ANN solves problem without any conditions or assumption, it is considered as an effective and widely used tool. Some samples of ANN that are used in this area are MultiLayer Perceptron (MLP) [5], Convolutional Neural Network (CNN) [6] Recurrent Neural Network (RNN) [7]. The SVM is another widely used tool to solve the time series forecasting problems [8]. Random modeling is one of the oldest forecasting tools in time series problem. Two common random model for forecasting time series are Autoregressive (AR) and Moving Averaging (MA) [2] and [3]. Also, two new models Autoregressive Moving Average (ARMA) and ARIMA are introduced by combining these two models (see in [2] and [3]). In AR(p) future point of a time series is calculated based on linear combination of p previous value of time series, a constant value and random noise that is shown in (2),

$$y_i = c + \sum_{i=1}^p \varphi_i y_{i-1} + \epsilon_i \quad (2)$$

where y_{i-1} are previous value of time series, ϵ_i is random noise, c is a constant value and φ_i are the linear parameters of the model. MA model is proposed with some change AR. In MA future point is calculated based on a linear combination of previous error that is shown in (3),

$$y_t = \mu + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t \quad (3)$$

where ϵ are the errors of time series with some statistical properties, μ is a mean value and φ_i are the linear parameters of the model. ARMA and ARIMA models are proposed by efficient combining of the AR and MA models. Method BJ has proposed an iterative approach for finding best model among ARIMA models without any assumption on time series problem [1].

Random models are used as benchmarks for comparisons in most studies because they are understandable. Random models perform well on linear data whereas most real time series data are nonlinear. Due to the limitations of random models in real data with special features such as missing values, multidimensional and nonlinear, they are less applicable.

Because of their desirable properties, neural networks are more suitable tools than random models. Neural networks are inherently data-driven and self-adaptive. Neural networks have the ability to learn a variety of linear and nonlinear data without the need to have expert knowledge and assumption of the data [9].

The neural network is a model of supervised learning so time series data must be transformed into this model. A sliding window is used to convert time series data into a supervised learning structure. A fixed length window slides over the time series data. The values in the window are considered as observations, which are represented by x . Values of the time series that are intended to be considered as neural network outputs represented by y . As a result, the $y = f(x)$ is modeled using a neural network.

A variety of neural networks have been introduced so far, some of which have been used in forecasting time series. One of the examples of ANN which is used in forecasting time series is MLP. TLNN neural network is a MLP ANN which is used for estimating unseen value of time series [10] which is used for estimating air line time series data. Some of the prominent features of MLPs that make them effective for time series estimation are nonlinear modeling, noise resistance, flexibility in the number of inputs and outputs that give flexibility to the forecasting length and the possibility of its development in multidimensional applications.

CNN are another set of neural networks that have a variety of applications in recent years. Numerous modeling using CNNs to estimate time series have also been proposed. In these models, CNNs filters learn periodic patterns and estimate time series with them. One-dimensional CNN is usually used in time series forecasting. The observed time series values are given as input to the network and by using a multilayer network unobserved values is estimated in the output [9].

Model WaveNet [11] is one of the CNN used for forecasting time series. In these model, casual convolution is used that means $t + 1$ value of times series is uncorrelated to x_{t+2}, \dots, x_{t+T} values. In this model pooling layer is not used and so the dimensions of the input and the output of the network is equal. This model is so fast because the recurrent connections are not used. In this model, the acceptance domain is increased by increasing the size of the layers and filters. For solving this problem a dilated convolution neural network was proposed. The dilated convolution increases the input domain by ignoring some input steps. The UFCNN is an other sample of CNN which is used for forecasting time series data [12]. The UFCNN is a kind of FCN that is used for semantic image segmentation. The UFCNN use undecimated wavelet transform undecimated convolution for forecasting time series data. In [13] a variety of deep learning based models for time series forecasting in financial applications are presented and compared according to their results.

The Recurrent Neural Network (RNN) such as Long Short-Term Memory (LSTM) is the other tools that is used for forecasting time series [14] and [15]. In this

network, the history of the inputs is used by using a recurrent connection. The LSTM give accurate estimation of time series data by using the historical state of the inputs and current values of the inputs simultaneously. Many various of LSTM models were proposed for forecasting time series. In [9] a multilayer neural network is used at the end of the based model of LSTM that estimates the next value of time series data based on the LSTM outputs. Stacked LSTM is another type of LSTM that uses a stack of LSTM to estimate more complex patterns [16]. Bidirectional LSTM is another variation of LSTM that is made to distinct LSTM [17]. One LSTM is used for forward estimating and other is used for backward estimating. In time series forecasting the bidirectional LSTM is used as a transformer from input to coded output. Each input is coded based on the future and the final coded output is given to the common LSTM for estimating as an input [18]. CNN-LSTM is a mixture of CNN neural network and LSTM neural network that the spatial model of the input is extracted with the CNN and temporal model of the input is extracted with LSTM. Using a mix of CNN and LSTM in time series forecasting is varied. A one-dimensional CNN is usually used to obtain spatial information, followed by a RNN to obtain temporal information [19], [20], [15] and [21]. In [15] a complete review of methods based on RNN for time series forecasting is provided. Due to the desirable performance of neural network-based models in the time series forecasting, in new and serious applications such as modeling severe epidemic and pandemic behaviour of Covid-19 are used [22] and [23]. In [22], the future conditions of novel Coronavirus is predicted by LSTM model. In [23] a LSTM model is used to predict the future mutation rate of Coronavirus virus which was able to provide acceptable results. In [24] a transformer neural network is used for time series forecasting where transformer-based model is used without change.

Encoder and Decoder networks are another approach of neural networks, an encoder network encodes the input data based on a particular pattern and then a decoder network decodes the output based on the encoded input to produce the desired output. Since these networks have a better understanding of the context, they provide a better performance. A wide variety of these types of networks are introduced in time series estimation. One of the most successful of these models was the LSTM sequence-to-sequence model with attention [25], [26] and [27]. In attention networks, the impact of coded inputs on outputs is controlled and intelligently applied in terms of other knowledge of the problem.

Neural networks are an effective tool for estimating time series due to their nonlinearity and their ability to be used for different time series without specific

information of those. Of course, challenges still remain, for example to provide good performance over long sequences, especially in RNN, which require further research. In this paper, one of the objectives is improving the performance of neural network-based tools for estimating long sequences.

Proposed Model

Since neural network has suitable features for time series forecasting, in this study, a model based on one of the desirable neural network models for estimating time series is proposed. The base model which is used in this study is the transformer model. In following describes how to apply this model to time series forecasting.

A. Transformer Model

The transformer model [28] was initially proposed for machine translation, but due to its high performance it was quickly incorporated into other areas such as image production [29], audio [30], text summarization [31], and music [32]. The transformer does not use recurrent and convolution, but uses attention in modeling. The transformer uses an encoder-decoder approach. Initially the input is entered into the encoder, after encoding, the output is generated according to the encoded input and the previous outputs in the decoder. The transformer network is used in proposed model as a based model.

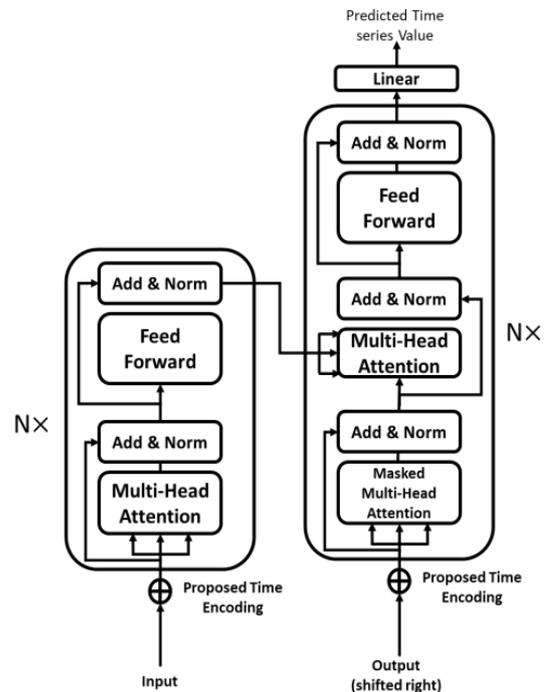


Fig. 1: Architecture of customized transformer neural network.

The encoder section contains a stack of encoders. The number of encoders in a stack is a free parameter, which is usually considered to be 6 layers [28]. In the decoder, a stack of decoders which is equal to the number of

layers in the encoder is used. Each encoding layer has its own parameters i.e. no weight is shared between these layers. Unlike recurrent networks, the transformer has no problem vanishing gradient and can access any point in the past regardless of the distance between words. This feature allows the transformer to find long-term dependencies. Also, unlike recurrent networks, the transformer lacks sequential computation and can run completely in parallel at high speeds. Because of the transformer design for machine translation, it cannot be directly used to forecasting time series. In following, changes made to the transformer to make it suitable for forecasting time series are given.

B. Transformer Model Customization

To adapt the transformer model with any custom time series some modifications to the base model are provided which is shown in Fig. 1. First the embedding layers of the model 's input related to NLP are omitted and the value of time series at any time from $z_t \in R^n$ is given to the model as input. The soft-max layer of the output that is used for classifying is also omitted and the mean square error (MSE) function which is related to regression is used as cost function according to (4).

$$e = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (4)$$

where y_i is the real output, \hat{y}_i is the generated output by the model. The transformer model uses Sin-Cosine method to encode time, which significantly reduces model accuracy in forecasting continuous time series. So, in proposed model, the time is coded by embedding a unique vector to the input data for each time. Vector with length n is used for time intervals with length n , for the value of time i , all its components are zero, except for position i th, whose value is one. The Fig. 2 illustrate the vectors added in this encoding. In this method each row is concatenated to the vector of the corresponding time.

1	0	...	0	0
0	1		0	0
.	.		.	.
.	.		.	.
.	.		.	.
0	0		1	0
0	0	0	1	

Fig. 2: Spatial encoder vectors.

In fact using this method for encoding time, a trainable vector is added to v , q and k vectors, letting the model to determine the value of the vectors by itself in contrary to the cosine and sine encoding method. Equation 6 illustrate the changes of the new encoding model in compare to the cosine-sine encoding method (5).

$$\begin{aligned}
k_i &= x_i W^k, v_i = x_i W^v, q_i = x_i W^q \\
x_i &= x_i^d + pos_i \\
q_i &= (x_i^d + pos_i) W^q = x_i^d \times W^q + pos_i \times W^q \\
v_i &= (x_i^d + pos_i) W^v = x_i^d \times W^v + pos_i \times W^v \\
k_i &= (x_i^d + pos_i) W^k = x_i^d \times W^k + pos_i \times W^k \quad (5) \\
e_{i,j} &= \frac{q_i \times k_j^T}{\sqrt{d}} \\
a_{i,j} &= \frac{\exp(e_{i,j})}{\sum_{k=1}^n \exp(e_{i,k})} \\
z_i &= \sum_{j=1}^n a_{i,j} \times v_j
\end{aligned}$$

where x_i^d represents the embedding vector and pos_i represents the cosine-sine vector in i time. Equation (5) is rewritten using new proposed time 's encoding in the form of (6) (only the modified parts of which are represented),

$$\begin{aligned}
x_i &= \text{concat}(x_i^d, \text{one} - \text{hot}(i)) \\
q_i &= (x_i^d) W^q + a_i^q \\
k_i &= (x_i^d) W^k + a_i^k \\
v_i &= (x_i^d) W^v + a_i^v \quad (6)
\end{aligned}$$

where x_i^d represents the time series vector, $z_{t=i} \in R^n$ and $\text{one} - \text{hot}(i)$ represents the time encoding vector in time i . The a_i^q , a_i^k and a_i^v are trainable vectors. In other words in (5) x_i^d is a trainable vector resulting from embedding and $pos_i * W$ vector is the time coder. While in (6) x_i^d is the time series value in i which is determined but the a_i^q , a_i^k and a_i^v time coder vectors are trainable vectors.

One of the problems in time encoding using the proposed *one - hot* method, is the input vector enlargement because of the increase the input steps. It happens in a way that even the size of time encoding vector becomes bigger than the input vector. To solve this problem, several different time encoding vectors are used for different time intervals. For example in a hourly data for one week instead of using a 168(24×7) vector, a 24 vector was using to specify the hour and 7 vectors for days of the week. The two vectors are concatenated so the time is coded in a 31($24 + 7$) vectors. In addition, instead of encoding the time absolutely from the beginning of the series, the typical date and time value is also used.

One of the reported research on transformer model, is replacing Cosine and Sine location encoding with relative vector [33] that increase the performance significantly. The presented solution of the reported research, using the relative vectors, is very similar to encoding time through concatenating a *one - hot* vector to the input vector. In the relative solutio, time is only coded in the v and k vector and the q vector only depends on the input vector independent of time which no explanation has been provided in the research. The relative solutio is also studied in proposed model.

Equation (7) shows the difference in calculating the q vector in this case. The other calculations are the same.

$$\begin{aligned} k_i &= x_i W^k, v_i = x_i^d W^v, q_i = x_i^d W^q \\ x_i &= \text{concat}(x_i^d, \text{one-hot}(i)) \\ q_i &= (x_i^d) W^q, v_i = (x_i^d) W^v + a_i^v, k_i = (x_i^d) W^k + a_i^k \end{aligned} \quad (7)$$

In this research, in addition to the cost function MSE another cost function is used based on a probabilistic approach. Suppose that $z_t \in R^n$ is the value of z series in t which is a n dimensional vector. if i , is the i th dimension of the time series, the aim is to model the conditional probability distribution of the following equation, $P(Z_{i,t_0:T} | Z_{i,1:t_0-1})$, which is shown the probability of future value of the i dimension. Equation (8) shows the observed interval of the time series and (9) shows unknown interval of the time series.

$$Z_{i,t_0:T} := [Z_{i,t_0}, Z_{i,t_0+1}, \dots, Z_{i,T}] \quad (8)$$

$$Z_{i,1:t_0-1} := [Z_{i,1}, Z_{i,2}, \dots, Z_{i,t_0-1}] \quad (9)$$

Equation (8) is realized based on the autoregression which is proposed in RNN architecture [34]. it is assumed for the sake of simplicity in calculations, the probabilities of each time series value based on its previous values are independent of each other. Accordingly, the conditional probability distribution in the form of (10) can be calculated.

$$Q_\theta(Z_{i,t_0:T} | Z_{i,1:t_0}) = \prod_{t=t_0}^T Q_\theta(Z_{i,t} | Z_{i,1:t-1}) = \prod_{t=t_0}^T l(Z_{i,t} | \theta(x_t)) \quad (10)$$

where x_t shows the output of decoder at time t . In (10) $l(Z_{i,t} | \theta(x_t))$ indicates the probability of occurrence of the time series value $Z_{i,t}$ with respect to the matched probability distribution parameters θ . The distribution function is assumed to be the Gaussian probability distribution. The parameters of the Gaussian probability distribution function are mean and standard deviation $\theta = (\mu, \sigma)$ and the equation of probability distribution function is presented in (11).

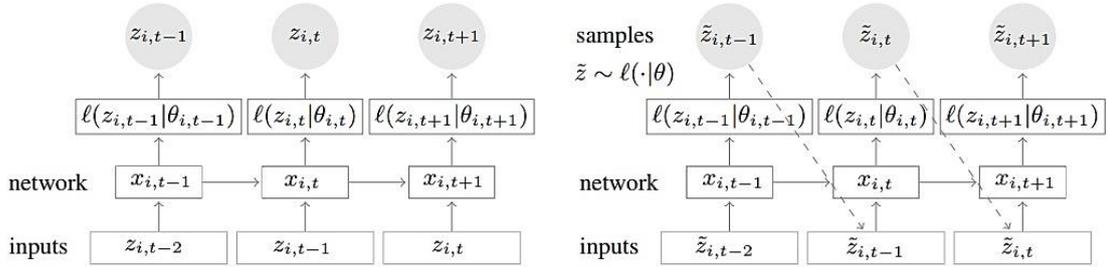


Fig. 3: Function of the model at training phase(left) and at testing phase(right) [35].

According to Fig. 3 the parameters of the distribution functions $l(Z_{i,t} | \theta(x_t))$ are estimated from the model output that is the estimated x_t . Then from the estimated distribution, the point with the highest

$$l_G(\mu, \sigma) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right) \quad (11)$$

$$\mu(x_{i,t}) = W_\mu^T x_{i,t} + b_\mu \quad (12)$$

$$\sigma(x_{i,t}) = \log(1 + \exp(W_\sigma^T x_{i,t} + b_\sigma)) \quad (13)$$

In order to estimate the probability distribution parameters based on the final output of the transformer neural network (12) and (13) are used. Probability distribution parameters proposed in (10) by maximizing probability logarithms is obtained as shown in (14). In other words, the probability distribution parameters are calculated to maximize the probability of joint distribution of the samples in the prediction interval.

$$L = \sum_{i=1}^N \sum_{t=t_0}^T \log(l(Z_{i,t} | \theta(x_t))) \quad (14)$$

Therefore, based on equations (10), (11) and (14) final cost function is shown as (15),

$$L = \sum_{i=1}^N \sum_{t=t_0}^T -\log(\sqrt{2\pi}) - \log(\sigma_{i,t}) - \frac{(z_{i,t} - \mu_{i,t})^2}{2\sigma_{i,t}^2}. \quad (15)$$

Since $\log(\sqrt{2\pi})$ value is constant, it can be ignored. Also, by considering the variance constant, the cost function will be converted to the MSE cost function, indicating that the MSE is in fact a special case of the maximum likelihood cost function, where the variance is constant.

In training phase, time series value is entered through the encoder-decoder attention layer from the observation section to the prediction section, and in the prediction or decoder interval, the network is trained by minimizing the $-L$ (see (15) function through the back propagation algorithm.

In testing phase, the network does not have access to the prediction interval. In this case, the network uses the output at any time as input at a later time. Figure 3 shows the model performance in the model training and evaluation process.

probability is considered as the output. Because Gaussian distribution considers the probability distribution function, the point has the maximum probability of being the mean point. The model is trained

by maximizing the joint probability of the output, and the model output at each step is used as input in the next step because of the prediction part is unknown.

Results and Discussion

A series of experiments have been performed to validate the proposed model. First, two datasets in different context that are used in the experiments are presented and quantitative validation metrics are introduced. Then the results of a variety of experiments with different purposes are presented. The results are discussed and summarized.

A. Datasets

The electricity consumption dataset in a power grid and traffic data of its application are used to validate the proposed model. The electricity consumption dataset was collected from monitoring of the electricity consumption of 370 households every 15 minutes from 01-01-2011 to 01-01-2015. By averaging every 4 consecutive time steps, the time intervals for both samples is converted from 15 minutes to 1 hour. So there are 24 observations per day. Also the model is trained by data from 01-01-2014 to 01-09-2014 and is validated by data from 01-09-2014 to 01-01-2015 data. In the training 7 consecutive days is used as input and 1 day later as output. In fact, the 168 time steps are considered as inputs and the next 24 steps as the estimation intervals. The traffic dataset was collected from 963 traffic monitoring in the city of San Francisco USA every 10 minutes from 01-01-2008 to 30-03-2009. By averaging every 6 consecutive time steps, the dataset is converted from 10-minute data to hourly data. Similar to the electricity consumption dataset, there are 24 observations per day, the last 7 days are used to train the model as input interval and the next 1 day as output or estimation interval. In fact we have 168 input steps and 24 output steps. In this dataset all available data prior to 15-06-2008 is used to train the model and the rest of the data is used as the validation data.

B. Evaluation Metrics

In this study two different metrics are used to evaluate the proposed model. Normalized Mean Square Error ($NMSE$) and Normalized Deviation (ND) are two metrics of network performance evaluation. The ND actually represents the standard ρ - risk metric that is well-known for predicting time series with $\rho = .5$. The equation of $NMSE$ and the equation of ND are given in (16) and (17) respectively.

$$NRMSE = \sqrt{\frac{\frac{1}{N(T-t_0)} \sum_{i,t} (z_{i,t} - \hat{z}_{i,t})^2}{\frac{1}{N(T-t_0)} \sum_{i,t} |z_{i,t}|}} \quad (16)$$

$$ND = \frac{\sum_{i,t} |z_{i,t} - \hat{z}_{i,t}|}{\sum_{i,t} |z_{i,t}|} \quad (17)$$

where N is a dimension of time series for each step,

$(T - t_0)$ is the estimation interval, $z_{i,t}$ is the actual value of time series data and $\hat{z}_{i,t}$ is the estimation value of the time series in time t and dimension i . Both metrics used are normalized, meaning that the effect of time series variations on them is diminished.

C. Results

Normalization is used in pre-processing the dataset. The values of the series are subtracted from the mean and divided by the standard deviation. The equation that is used for normalization is shown in (18),

$$y = \frac{x - \mu}{\delta} \quad (18)$$

$$\mu = \frac{1}{n} \sum_{i=1}^N x_i$$

$$\delta^2 = \frac{1}{n-1} \sum_{i=1}^N (x - \mu)^2$$

where y is the normalized value and μ is the mean and δ^2 is the variance. Also, in some experiments, the data is used that is transferred to the interval between zero and one according to the (19) as preprocessing. In the sliding window method, a dataset with length $T - w + 1$ is created from time series data with length T and window size w . w is the sum of the input and output intervals for each step of estimation.

$$y = \frac{x - \min}{\max - \min} \quad (19)$$

After normalization, the estimation problem should be converted to the problem with the observer. The step of process is performed with a sliding window that is divided in two section, the first is a observation and known values and the second one is the prediction section and unknown. In training phase two section of sliding window are known and in test phase the prediction section is estimated with the trained models and the actual values of this section is used for calculating evaluation metrics. This window moves forward one step at a time and generates new data. The results are obtained on a transformer with 2 layers, 8 heads, feed-forward layer with 512 neurons and 64 batch sizes after 15,000 training steps. A number of influential parameters of the network are discussed below. The experiments were performed by two transformer modes with the MSE objective function and the probabilistic objective function that are introduced. To validate the results, all the results were obtained from three different experiments and the mean and standard deviation of the different experiments were reported as a result. For comparing the results with other proposed methods the results which are reported in [35] are used. To compare the results, a similar datasets, training and testing procedures which are used in [35] are used. The obtained results from our proposed model, TRMF [36] and DeepAR [35] are shown in Table 1 and Table 2.

Table 1: The result of electricity consumption dataset

Metrics	TRMF	Deep AR	Transformer with MSE objective function	Transformer with Probabilistic objective function
NRMSE ↓	1.15	1	1.041±0.052	0.889±0.029
ND ↓	0.16	0.07	0.137±0.006	0.122±0.003

Table 2: The result of traffic dataset

Metrics	TRMF	Deep AR	Transformer with Probabilistic objective function
NRMSE ↓	0.43	0.42	0.521±0.047
ND ↓	0.18	0.17	0.24±0.012

As shown in Table 1 and Table 2 the transformer with the probabilistic objective function yields better results than the transformer with MSE objective function on both evaluation metrics. Also, the transformer with probabilistic objective function performs better on *NMSE* metric than on other models, but on *ND* metric it yields better than the TRMF model for electricity consumption dataset.

Each time step in the transformer attend to the other steps and each step is coded according to all other steps. By changing the mask vectors, it is possible to determine what other steps to pay attention to each time step. By changing the masks so that each time step is only causally attend on its preceding steps or the number of time steps ahead of itself and also the number of time steps r in its neighborhood, no change in network performance is observed. The results of the experiments shows that changes in the number of layers and attentions did not affect the network performance. One reason for this phenomenon is the difference between common time series with the concepts of natural languages time series. Generating attention vector(Mask vector), considering the context of each step, seems to be a good solution to this problem. In electricity consumption dataset based on the context behind the data, seasonal, weekly or daily steps can be defined. For example, by changing the definition of attention to how the encoder in each time step would only pay attention to the same day time steps and in decoder would only pay attention to same time step of previous day better results were obtained. Based on the results of this experiment that are shown in Table 3 and Table 4 performance has obviously increased with the context based change in attentions. Therefore, it can be concluded that a suitable change in attitudes can increase the efficiency of the transformer.

Another improvement which is introduced is encoding the time of values in time series with their

actual values instead of relative values. For example, in the relative encoding modes for encoding time in a daily time sequence with 3 consecutive steps one-hot 1, 2 and 3 vectors were attached to each step as the first to third days.

Table 3: The result of electricity consumption dataset after change in attention

Metrics	TRMF	Deep AR	Transformer with Probabilistic objective function	Transformer with attention change
NRMSE ↓	1.15	1	0.889±0.029	0.734±0.012
ND ↓	0.16	0.07	0.122±0.003	0.101±0.0004

Table 4: The result of traffic dataset after change in attention

Metrics	TRMF	Deep AR	Transformer with Probabilistic objective function	Transformer with attention change
NRMSE ↓	0.43	0.42	0.521±0.047	0.485±0.011
ND ↓	0.18	0.17	0.24±0.012	0.189±0.0007

Whereas in encoding by the actual value method, if the three consecutive steps are related to Tuesday through Thursday one-hot 4, 5 and 6 vectors were attached to each step. Encoding the location with the actual value did not cause a significant change in the electricity consumption dataset as before, but in the traffic data set the changes are significant which is shown in Table 5.

Table 5: The result of traffic dataset with actual value location encoding

Metrics	TRMF	Deep AR	Transformer with relative time encoding	Transformer with actual time encoding
NRMSE ↓	0.43	0.42	0.521±0.047	0.46±0.01
ND ↓	0.18	0.17	0.24±0.012	0.147±0.0005

Another investigated result is the use of normalization preprocessing based on (19) for data transfer between zero and one and the use of sigmoid function on linear transformation based (12) and (13) for estimates of the distribution parameters. The result of the estimation is given in Table 6 and Table 7.

Table 6: The result of traffic dataset with zero-one normalization and Sigmoid Function

Metrics	TRMF	Deep AR	Based model of Transformer	Transformer with new configuration
NRMSE ↓	0.43	0.42	0.521±0.047	0.433±0.009
ND ↓	0.18	0.17	0.24±0.012	0.134±0.0006

Table 7: The result of electricity consumption dataset with zero-one normalization and Sigmoid Function

Metrics	TRMF	Deep AR	Based model of Transformer	Transformer with new configuration
NRMSE↓	1.15	1	0.889±0.029	0.722±0.017
ND↓	0.16	0.07	0.122±0.003	0.93±0.004

One of the important features in time series prediction models is the ability of the model to predict long time. The performance of many models, such as the Deep AR model, is reduced by more than four times as the number of predictive steps is increased [37]. Changing the efficiency of the proposed transformer model with the probabilistic objective function in the RMSE metric when the prediction interval is increment from 24 steps to 48, 96 and 168 in the electricity consumption dataset is shown in Fig. 4.

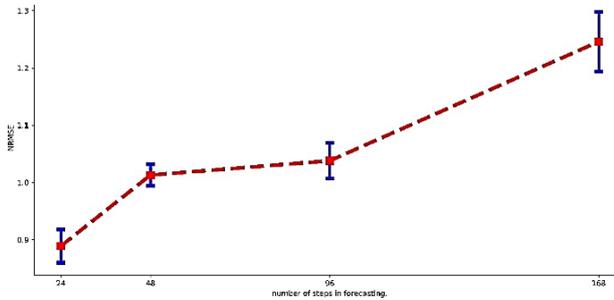


Fig. 4: The effect of increasing the estimation interval(Horizontal Axis) on NRMSE(Vertical Axis) in transformer.

As Fig. 4 shows, the error increases with the prediction interval increment, but the range of the error is from 0.9 to 1.3, while other models such as Deep AR is shown greater error by increasing prediction interval [37]. In other words, Fig. 8 shows that the proposed transformer is capable of long-term prediction.

Due to the lack of sequential computing, the transformer can be run completely in parallel. For this reason, the transformer has a higher speed than the recurrent neural network models. Also, self-attention has fewer computations than recurrent and convolution methods when the sequence length is less than the input dimension. On the other hand it lacks sequential computation. The restricted self-attention is a kind of attention that each word attend to only r neighbor words [28]. The implementation of the heads is done so that the calculations of several different heads are paralleled. Thus increasing the number of heads does not change the time complexity of the transformer. Also as shown in Table 9 as the input sequence increases, the computations increase quadratically. Increasing the input and output interval does not affect the number of network parameters. As the number of layers increases,

the number of parameters change linearly. In order to display the increase rates of the number of network parameters in Fig. 10 the number of parameters for different number of layers are provided for the sequence dimensions 963, feed dimensions 512, and dimension of v , q and k vectors 20.

Table 8. Compare the computational complexity for self-attention, where n is the length of input, d is the dimension of input, k is the kernel size on convolution and r is the number of neighborhood for self-attention [28]

Layer Type	Complexity per Layer	Sequential Operations	Maximum Path Length
Self-Attention	$O(n^2 \cdot d)$	$O(1)$	$O(1)$
Recurrent	$O(n \cdot d^2)$	$O(n)$	$O(n)$
Convolutional	$O(k \cdot n \cdot d^2)$	$O(1)$	$O(\log_k(n))$
Recurrent	$O(r \cdot n \cdot d)$	$O(1)$	$O(n/r)$

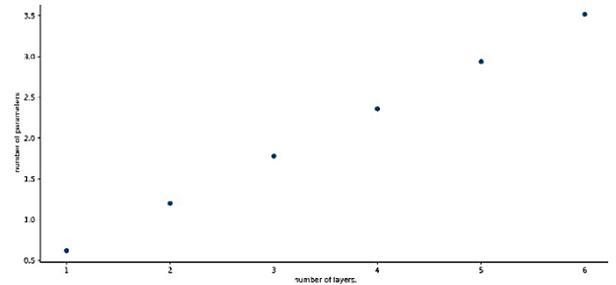


Fig. 5: Change the number of the network parameters (Vertical Axis) by changing the number of layers(Horizontal Axis).

According to Fig. 5, the network parameters vary linearly between 2 million to 12 million for 1 to 6 layers. The runtime of recurrent models, such as Deep AR, on the common dataset is approximately 7 hours [35], while with the transformer model achieving better results in $NMSE$ metric and competing results in ND metric in about 40 minutes.

The transformer model has the parameters of the number of layers, the dimensions of the feed-forward layer, the number of heads, and the dimensions of the vectors k , q and v . The dimensions of the feed-forward layer are typically much larger than the input dimension of the problem and are about twice that. In estimations As the number of layers increased, there was no change in the efficiency of the converter. The transformer uses a sine and cosine function to encode time. This method is less efficient in time series data, so this function was replaced by the time encoding method by adding a spatial vector. In training phase of the model Adam optimizer is used [38]. The rate of training on the network is very impressive and it is very difficult to determine. By examining different training rates and using different cut rates, finally the Noam scheduling training rate introduced in the transformer was used.

Also, to train the model, Google's research collaboration service with the K80 graphics computing unit was used.

Conclusion

Many issues in different domains can be modeled in the time series forecasting. Therefore, providing a desirable tool for estimating time series can have significant applications. In this paper, an effective tool for estimating time series is introduced. The objectives of this paper have been to provide a more accurate and long interval estimation tool. Neural networks are effective and useful tools for modeling various problems such as estimation time series problem. Introducing the transformer network as an effective tool in natural language processing has made a significant improvement in performance compared to previous methods in recent years. Since contextual data is a complex example of time series data, the transformer network seemed to be a useful tool in forecasting time series. In this paper, the modified transformer model use in a variety of time series with changes in the structure of this neural network. The natural language processing layers were removed, changes were made to the encoding temporal layer, and a new regression-based objective function was introduced. Experiments were performed using a modified transformer network on a variety of datasets and it was found that the tool could perform better or compete with other introduced methods with less computational complexity and longer estimation intervals. It was also found that with better configuration of the network and better adjustment of attention, it is possible to obtain more desirable results in any specific problem. Therefore, the proposed model has the potential to provide more desirable results with further study and it is a good base model for presenting a set of effective methods in time series forecasting.

Author Contributions

Reza Mohammadi Farsani designed and simulated and carried out the data analysis, and Ehsan Pazouki collected the data and interpreted the results and wrote the manuscript.

Acknowledgment

The authors gratefully thank the anonymous reviewers and the editor of JECEI for their useful comments and suggestions.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

References

- [1] R. Adhikari, R.K. Agrawal, "An introductory study on time series modeling and forecasting," *ArXiv Prepr. ArXiv13026613*, 2013.
- [2] G. E. Box, G. M. Jenkins, G. Reinsel, "Time series analysis: forecasting and control Holden-day San Francisco," *BoxTime Ser. Anal. Forecast. Control Holden Day1970*, 1970.
- [3] K.W. Hipel, A.I. McLeod, *Time series modelling of water resources and environmental systems*, 45, Elsevier, 1994.
- [4] G.P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, 50: 159–175, 2003.
- [5] F. Rosenblatt, "Principles of neurodynamics. perceptrons and the theory of brain mechanisms," *Cornell Aeronautical Lab Inc Buffalo NY*, 1961.
- [6] A. Krizhevsky, I. Sutskever, G.E. Hinton, "Imagenet classification with deep convolutional neural networks," *Comm. ACM*, 60(6): 1097–1105, 2012.
- [7] F.A. Gers, J. Schmidhuber, F. Cummins, "Learning to forget: continual prediction with LSTM," *Neural Comput.*, 12(10): 2451–2471, 2000.
- [8] L.-J. Cao, F.E.H. Tay, "Support vector machine with adaptive parameters in financial time series forecasting," *IEEE Trans. Neural Netw.*, 14: 1506–1518, 2003.
- [9] J. Brownlee, *Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python. Machine Learning Mastery*, 2018.
- [10] J. Faraway, C. Chatfield, "Time series forecasting with neural networks: a comparative study using the airline data," *J. R. Stat. Soc. Ser. C Appl. Stat.*, 47: 231–250, 1998.
- [11] A. Borovykh, S. Bohte, C.W. Oosterlee, "Conditional time series forecasting with convolutional neural networks," *ArXiv Prepr. ArXiv170304691*, 2017.
- [12] R. Mittelman, "Time-series modeling with undecimated fully convolutional neural networks," *ArXiv Prepr. ArXiv150800317*, 2015.
- [13] O.B. Sezer, M.U. Gudelek, A.M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," *Appl. Soft Comput.*, 90: 106181, 2020.
- [14] S. Hochreiter, J. Schmidhuber, "Long short-term memory," *Neural Comput.*, 9: 1735–1780, 1997.
- [15] H. Hewamalage, C. Bergmeir, K. Bandara, "Recurrent Neural networks for time series forecasting: current status and future directions," *Int. J. Forecast.*, 37(1): 388–427, 2021.
- [16] Z. Cui, R. Ke, Y. Wang, "Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction," *ArXiv Prepr. ArXiv180102143*, 2018.
- [17] M. Schuster, K.K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, 45: 2673–2681, 1997.
- [18] Z. Cui, R. Ke, Y. Wang, "Deep stacked bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction," presented at the 6th International Workshop on Urban Computing (UrbComp 2017), Nova Scotia, Canada, 2016.
- [19] H. Yao, X. Tang, H. Wei, G. Zheng, Y. Yu, Z. Li, "Modeling spatial-temporal dynamics for traffic prediction," *ArXiv Prepr. ArXiv180301254*, 2018.
- [20] J. Ke, H. Zheng, H. Yang, X. M. Chen, "Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach," *Transp. Res. Part C Emerg. Technol.*, 85: 591–608, 2017.
- [21] Y. Liu, H. Zheng, X. Feng, Z. Chen, "Short-term traffic flow prediction with Conv-LSTM," presented at International Conference on Wireless Communications and Signal Processing, Nanjing, China, 2017.

- [22] S. Shastri, K. Singh, S. Kumar, P. Kour, V. Mansotra, "Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study," *Chaos Solitons Fractals*, 140: 110227, 2020.
- [23] R.K. Pathan, M. Biswas, M.U. Khandaker, "Time series prediction of COVID-19 by mutation rate analysis using recurrent neural network-based LSTM model," *Chaos Solitons Fractals*, 138: 1-7, 2020,
- [24] N. Wu, B. Green, X. Ben, S. O'Banion, "Deep transformer models for time series forecasting: the influenza prevalence case," *ArXiv200108317 Cs Stat*, 2020.
- [25] M.T. Luong, H. Pham, C.D. Manning, "Effective approaches to attention-based neural machine translation," *ArXiv Prepr. ArXiv150804025*, 2015.
- [26] Y. Qin, D. Song, H. Chen, W. Cheng, G. Jiang, G. Cottrell, "A dual-stage attention-based recurrent neural network for time series prediction," *ArXiv Prepr. ArXiv170402971*, 2017.
- [27] Y. Liang, S. Ke, J. Zhang, X. Yi, Y. Zheng, "GeoMAN: multi-level attention networks for geo-sensory time series prediction," in *Proc. International Joint Conference on Artificial Intelligence: 3428-3434*, 2018.
- [28] N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, "Attention is all you need," *ArXiv Prepr. ArXiv1706.03762*, 2017.
- [29] N. Parmar, A. Vaswani, J. Uszkoreit, Ł. Kaiser, N. Shazeer, A. Ku, D. Tran, "Image transformer," *ArXiv Prepr. ArXiv180205751*, 2018.
- [30] D. Povey, H. Hadian, P. Ghahremani, K. Li, S. Khudanpur, "A time-restricted self-attention layer for asr," presented at *IEEE International Conference on Acoustics, Speech and Signal Processing, Calgary, AB, Canada*, 2018.
- [31] P. J. Liu, M. Saleh, E. Pot, B. Goodrich, R. Sepassi, L. Kaiser, N. Shazeer, "Generating wikipedia by summarizing long sequences," *ArXiv Prepr. ArXiv180110198*, 2018.
- [32] C.Z.A. Huang, A. Vaswani, J. Uszkoreit, I. Simon, C. Hawthorne, N. Shazeer, A.M. Dai, M. D. Hoffman, M. Dinculescu, D. Eck, "Music transformer: Generating music with long-term structure," *ArXiv Prepr. ArXiv180904281*, 2018.
- [33] P. Shaw, J. Uszkoreit, A. Vaswani, "Self-attention with relative position representations," *ArXiv Prepr. ArXiv180302155*, 2018.
- [34] A. Graves, "Generating sequences with recurrent neural networks," *ArXiv Prepr. ArXiv13080850*, 2013.
- [35] D. Salinas, V. Flunkert, J. Gasthaus, "DeepAR: Probabilistic forecasting with autoregressive recurrent networks," *ArXiv Prepr. ArXiv170404110*, 2017.
- [36] H.-F. Yu, N. Rao, I. S. Dhillon, "Temporal regularized matrix factorization for high-dimensional time series prediction," in *Advances in neural information processing systems*, 29: 847-855, 2016.
- [37] S. Li, X. Jin, Y. Xuan, X. Zhou, W. Chen, Y. X. Wang, X. Yan, "Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting," *ArXiv Prepr. ArXiv190700235*, 2019.
- [38] D.P. Kingma, J. Ba, "Adam: A method for stochastic optimization," *ArXiv Prepr. ArXiv14126980*, 2014.

Biographies



Reza Mohammadi Farsani is a graduate of Shahid Rajaee Teacher Training University in the field of computer engineering majoring in artificial intelligence at the master's level. (e-mail: r.mohammadifarsani@gmail.com)



Ehsan Pazouki is an Assistant Professor in the school of Computer Engineering at Shahid Rajaee Teacher Training University where he has been a faculty member since 2016. Ehsan completed his Ph.D. and M.S. at Amirkabir University. His research interests lie in the area of wide area surveillance, ranging from theory to design to implementation. He has collaborated actively with researchers in several other disciplines of computer science, particularly Cognitive Science. Ehsan has experience in various industries related to his specialization for more than 10 years. For additional information see <https://www.sru.ac.ir/en/school-of-computer/ehsan-pazouki/>

Copyrights

©2020 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

R. Mohammdi Farsani, E. Pazouki, "A transformer self-attention model for time series forecasting," *J. Electr. Comput. Eng. Innovations*, 9(1): 1-10, 2021.

DOI: [10.22061/JECEI.2020.7426.391](https://doi.org/10.22061/JECEI.2020.7426.391)

URL: http://jecei.sru.ac.ir/article_1477.html





Research paper

Target Detection Using Multispectral Images, a Case Study: Wheat Detection in Chenaran County in Iran

*M. Imani**

Department of Telecommunication, Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran.

Article Info

Article History:

Received 25 May 2020
Reviewed 04 July 2020
Revised 30 September 2020
Accepted 03 November 2020

Keywords:

Target detection
Multispectral image
Wheat field
Attribute filter
Collaborative representation

*Corresponding Author's Email
Address:
maryam.imani@modares.ac.ir

Abstract

Background and Objectives: Target detection is one of the main applications of remote sensing. Multispectral (MS) images with higher spatial resolution than hyperspectral images are an important source for shape and geometric characterization, and so, MS target detection is interested.

Methods: A target detector appropriate for multispectral (MS) images is selected among hyperspectral target detectors and redefined in this paper. Many target detectors have been proposed for hyperspectral images in the remote sensing filed. Most of these detectors just use the spectral information. Since, the MS images have higher spatial resolution compared to hyperspectral ones, it is proposed that select a target detector that uses both of the spectral and spatial features. To this end, the attribute profile based collaborative representation (AP-CR) hyperspectral detector is chosen for MS images. Shape structures extracted by flexible attribute filters can significantly improve the MS target detection.

Results: As a case study, the wheat fields in Chenaran County in Iran are chosen as targets to be detected. The image acquired by Landsat 8 is used for doing experiment. The results show the superior performance of AP-CR with 96.09 % accuracy for wheat detection using MS image of Landsat 8.

Conclusion: The high performance of AP-CR is due to extraction of flexible attribute characteristics and the use of collaborative representation for approximation of each image pixel. Although the AP-CR method provides the highest accuracy, it needs a high running time compared to other detectors.

Introduction

Identification and detection of target pixels such as man-made objects, special agriculture products and certain minerals is one of the main applications of remote sensing images. Target detection is interested among in various fields of agronomy, geology, mineralogy and so on [1], [2], [3] and [4]. Hyperspectral images by providing rich dimensionality for each pixel in the spectral domain allows for analyzing and discrimination of a variety of materials and land covers. A hyperspectral image is composed of hundreds images acquired in a wide range of electromagnetic spectrum where a high spectral

resolution about 10 nm can be provided. So, hyperspectral images with two spatial dimensions and a spectral dimension have a high capability in separation of targets from background. In addition to hyperspectral images, there are multispectral (MS) images that capture images from the ground in multiple (approximately about 10) spectral bands. Although MS images have lower spectral resolution compared to the hyperspectral images, but, they have higher spatial resolution that is important in characterization of shape, geometric and spatial structures of objects of image [5], [6] and [7]. So far many target detectors have been introduced that

most of them are applied for hyperspectral images. Table 1 provides a short review of them: Spectral matched filter (SMF) [8], Spectral angle mapper (SAM) [9], Matched subspace detector (MSD) [10], Orthogonal subspace projection (OSP) [11], Constrained energy minimization (CEM) [12], Adaptive subspace detector (ASD) [13], Sparsity based target detector (STD) [14], Kernel based SAM (KSAM) [15], Difference based target detection (DTD)-KSAM [16], Attribute profile based collaborative representation (AP-CR) [17] and some

other ones [18], [19], [20] and [21].

In this paper, we assess different hyperspectral target detectors to find which of them is appropriate for MS target detection. Since MS images have relatively high spatial resolution, AP-CR that utilizes the spatial information in addition to spectral one is selected for target detection.

AP-CR is a non-parametric approach with no need to estimate the data statistics.

Table 1: A short review of target detectors

Target detector	Year	Method
Spectral matched filter (SMF)	1992	A limiting case of generalized likelihood ratio test (GLRT) detector with simplified test statistic in the form of a normalized matched filter
Spectral angle mapper (SAM)	1993	It measures the spectral angle between the testing pixel and the spectral signature of target samples.
Matched subspace detector (MSD)	1994	An improved GLRT detector. Data is projected onto a low rank subspace to remove interference and remain signal. The image data projected onto the low rank subspace is matched to the image.
Orthogonal subspace projection (OSP)	1994	It projects the spectral vector of each pixel onto a subspace that is orthogonal to the undesired characteristics. It is an optimal process for interference suppression through the least squares approach.
Constrained energy minimization (CEM)	2000	It maximizes the response of the target signature while suppresses that of the undesired background.
Adaptive subspace detector (ASD)	2001	Based on theory of GLRT for adaption of MSD for unknown noise covariance matrices. It includes the adaptive cosine estimator (ACE) for situations that scaling of test samples may deviate from that of the training one.
Sparsity based target detector (STD)	2011	It sparsely approximate an image by a few target samples and also individually by using some background samples. Then reconstruction residuals are employed to obtain the output detection.
A supervised metric learning [18]	2014	A subpixel Target Detection for dealing with mixed signature of target spectrum and background pixels spectra through learning a distance metric.
A maximum margin metric learning [19]	2015	It transfers a hyperspectral target detection problem into a maximum margin problem.
Kernel based SAM (KSAM)	2016	The generalized version of SAM appropriate for nonlinear cases obtained by kernel function.
Median-mean line Reed-Xiaoli detector [20]	2017	An improved version of RX which utilizes the benefits of median-mean line metric.
Attribute profile based collaborative representation (AP-CR)	2018	It uses the attribute profiles for construction of target and background subspaces and then approximation of testing pixel based on target and background subspaces individually using collaborative representation. Then, reconstruction residuals are employed to obtain the output detection
Difference based target detection (DTD)-KSAM	2019	Output detection is result of difference between the distance of testing sample to the background signal and that to the target signal. KSAM is used as distance measurement.
A sparse and collaborative representation based detector [2]	2020	It utilizes the advantages of both sparse and collaborative representation for anomalous target detection.
Self-regularized weighted sparse model [21]	2021	An algorithm based on this hypothesis that data may come from multi-subspaces. It is proposed to detect infrared small targets in complex background.

It uses the benefits of attribute profile [22], [23] and collaborative representation (CR) [24], [25] to increase separation of targets from background using both of the spectral and spatial features. It redefined to use for MS images. One of the research domains in agriculture and geoscience field is detection of crop products such as wheat [26], [27] and [28]. Specially, in this paper, the AP-CR method is used for wheat detection using MS images acquired by Landsat 8 in Chenaran county fields in Iran.

Target Detector

Target detection algorithms have been often introduced for hyperspectral images. Hyperspectral images have a high spectral resolution that provide a worthwhile and useful source of spectral characteristics from background and targets. According to this reason, most of hyperspectral target detectors just use the spectral information while ignore the valuable spatial information. The spatial information can be very important for MS images that have lower spectral resolution and higher spatial resolution with respect to hyperspectral images. So, among different hyperspectral target detectors, the AP-CR method that uses the spatial information of images in addition to the spectral one is selected and redefined for MS target detection.

The geometrical structures and shape features have much perceptual significance that is useful for modelling of various objects and discrimination between different materials. Then, exploitation of these spatial features can be useful for different applications such as classification, object detection and target detection [29], [30]. Morphological operators, i.e., “opening” and “closing” filters by reconstruction extract spatial features such as shape characteristics and geometrical structures [31]. A structure element with a specific shape such as circle, rectangular and so on as a sliding window is considered in the morphological transformations. The extracted shapes and structures are substantially dependent on the shape of the considered structure element.

To extract structures with more generality that are not limited to some specific shapes, the attribute filters have been introduced. The attribute filters use one or several attributes instead of using a fixed structure element. Any attribute can be extracted from the image regions. The great flexibility in selection and defining the attributes allows a high capability for modelling of spatial information that are useful for discrimination of targets from background. The considered attributes can be textural such as standard deviation, entropy and range, geometric such as image moments, length of the perimeter, area, and shape factors, etc. Generally any measure computed from the image can be used as an attribute.

An Attribute profile (AP) is constituted through applying several attribute filters to a single band (gray level) image. The attribute filters merge connected components of the image at different levels. The filtering process is done as follows: a selected attribute denoted by a is calculated for each connected region R_i . The computed value indicated by $a(R_i)$ is then compared with a reference value λ . Two decision may be made [32]:

$a(R_i) \geq \lambda \rightarrow$ no change is applied to the connected region.

$a(R_i) < \lambda \rightarrow$ the grey level value of neighboring region with closer value is assigned to the connected region.

According to this approach, the neighboring connected components are merged. Merging with a neighboring region containing lower gray level is known as thinning and merging with the adjacent region of a higher gray level is called thickening. The attribute filters are applied to each band of MS individually and then, the filter outputs are stacked together to form the AP. Let consider m spectral bands for MS, s attributes and a sequence of thresholds $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. A thinning profile with attribute a_k is composed from applying a sequence of attribute thinning operators:

$$\begin{aligned} Thin_{a_k}(y_j) &= \{\gamma_n(y_j), \dots, \gamma_1(y_j)\}; \\ j &= 1, 2, \dots, m; k = 1, 2, \dots, s \end{aligned} \quad (1)$$

where γ_i indicates the thinning operator and y_j is the j th spectral band of MS image. Similarly, the thickening profile is generated by applying thickening operators:

$$\begin{aligned} Thick_{a_k}(y_j) &= \{\varphi_n(y_j), \dots, \varphi_1(y_j)\}; \\ j &= 1, 2, \dots, m; k = 1, 2, \dots, s \end{aligned} \quad (2)$$

where φ_i denotes the thickening operator. The extended thinning profile and the extended thickening profile of k th attribute are given, respectively by:

$$\begin{aligned} E_{Thin_{a_k}} &= \{Thin_{a_k}(y_1), Thin_{a_k}(y_2), \dots, Thin_{a_k}(y_m)\}; \\ k &= 1, 2, \dots, s \end{aligned} \quad (3)$$

$$\begin{aligned} E_{Thick_{a_k}} &= \{Thick_{a_k}(y_1), Thick_{a_k}(y_2), \dots, Thick_{a_k}(y_m)\}; \\ k &= 1, 2, \dots, s \end{aligned} \quad (4)$$

and finally the extended multi thinning profile and the extended multi thickening profile are provided as follows:

$$EM_Thin = \{E_Thin_{a_1}, E_Thin_{a_2}, \dots, E_Thin_{a_s}\} \quad (5)$$

$$EM_Thick = \{E_Thick_{a_1}, E_Thick_{a_2}, \dots, E_Thick_{a_s}\} \quad (6)$$

The extended multi thinning profile and the extended multi thickening profile are called for simplicity thinning profile and thickening profile, respectively. While the thinning profile contain the detail information of image, the thickening one has more homogeneous regions. From the other hand, thinning and thickening just contain the spatial features. To include the spectral features too, the MS bands are stacked to the thinning profile to compose the target subspace. Moreover, the spectral bands are stacked to the thickening profile to generate the background subspace. The obtained target and background subspaces are used for estimate of each testing pixel. An unknown pixel is approximated by each of target subspace and background subspace individually. The label of each subspace that can be better estimate the testing pixel is assigned to the testing pixel. So, a label of target or background is determined for each pixel of image. To approximate the pixel through a subspace, the collaborative representation (CR) method is used. According to this approach, each pixel can be approximated through a linear combination of samples of subspace where a distance weighted Tikhonov matrix is considered to adjust the contribution of each sample of the subspace. CR has two main advantages: 1- it is a nonparametric method for image modeling. In other words, it does not consider any assumption about statistical distribution of data, 2- due to using l_2 norm in the optimization problem, CR is simply implemented and achieves a closed form solution. By considering \mathbf{X}_b as the background subspace, estimate of the testing pixel denoted by \mathbf{y} is obtained by:

$$\hat{\mathbf{y}}_b = \mathbf{X}_b \hat{\mathbf{a}}_b \quad (7)$$

Similarly, by considering \mathbf{X}_t as the target subspace, the estimate of \mathbf{y} is given by:

$$\hat{\mathbf{y}}_t = \mathbf{X}_t \hat{\mathbf{a}}_t \quad (8)$$

where $\hat{\mathbf{a}}_b$ and $\hat{\mathbf{a}}_t$ are the weight vectors associated with the background and target subspaces, respectively. These weight vectors are computed through solving an optimization problem. For more explanations, the reader is referred to [17]. These vectors are obtained as follows:

$$\hat{\mathbf{a}}_b = (\mathbf{X}_b^T \mathbf{X}_b + \lambda_b \mathbf{\Gamma}_{y_b}^T \mathbf{\Gamma}_{y_b})^{-1} \mathbf{X}_b^T \mathbf{y} \quad (9)$$

$$\hat{\mathbf{a}}_t = (\mathbf{X}_t^T \mathbf{X}_t + \lambda_t \mathbf{\Gamma}_{y_t}^T \mathbf{\Gamma}_{y_t})^{-1} \mathbf{X}_t^T \mathbf{y} \quad (10)$$

where λ_b and λ_t are the regularization parameters and $\mathbf{\Gamma}_{y_b}$ and $\mathbf{\Gamma}_{y_t}$ are the Tikhonov matrices. These matrices are used to assign larger coefficients to more similar pixels to the testing pixel. In other words, they force that pixels of the subspace with more similarity to the testing pixel, have greater contribution in approximation of

testing pixel. The output of target detector is obtained by:

$$D(\mathbf{y}) = r_b(\mathbf{y}) - r_t(\mathbf{y}) \quad (11)$$

where $r_b(\mathbf{y})$ and $r_t(\mathbf{y})$ are the residual error of approximation computed in background and target subspaces, respectively:

$$r_b(\mathbf{y}) = \|\mathbf{y} - \hat{\mathbf{y}}_b\|_2 = \|\mathbf{y} - \mathbf{X}_b \hat{\mathbf{a}}_b\|_2 \quad (12)$$

$$r_t(\mathbf{y}) = \|\mathbf{y} - \hat{\mathbf{y}}_t\|_2 = \|\mathbf{y} - \mathbf{X}_t \hat{\mathbf{a}}_t\|_2 \quad (13)$$

Note that $D(\mathbf{y})$ is a scalar number assigned to each pixel \mathbf{y} . The values of $D(\mathbf{y})$ for all pixels of image is used to form a grey level image. A pseudo color image can be generated according to this grey level image where a larger value for a pixel indicates that the pixel belongs to the target class with a higher probability and a smaller value is corresponding to belonging to the background class. To generate a binary detection map, it is enough that consider a threshold value for target detection:

$$\text{if } D(\mathbf{y}) \geq \text{thr} \rightarrow \mathbf{y} \text{ is target}$$

$$\text{if } D(\mathbf{y}) < \text{thr} \rightarrow \mathbf{y} \text{ is background}$$

The block diagram of the proposed AP-CR method is shown in Fig.1.

Study area and Evaluation Measures

The study area is Chenaran County in Razavi Khorasan Province in Iran with the following geographic coordinates:

36° 34'45" N

59° 17'41" E

36° 47'27" N

58° 52'47" E

Chenaran is an agricultural city with many farms in the outskirts. The image is acquired by Landsat 8¹. An overview of the whole dataset is shown in Fig. 2. A part of this dataset, i.e., the region of Chenaran, is used for doing experiments in this work that the Google earth image of it is shown in Fig. 3. The grey level image of band 6 of Landsat 8 is also seen in Fig. 4. To assess the detection performance, the receiver operating characteristic (ROC) that plots a probability of detection P_d versus the false alarm rate R_f is used:

$$P_d = \frac{N_{cd}}{N_t} \quad (14)$$

and

$$R_f = \frac{N_{fd}}{N} \quad (15)$$

where N is the total number of image pixels and N_t is the total number of target pixels.

¹ <http://earthexplorer.usgs.gov>

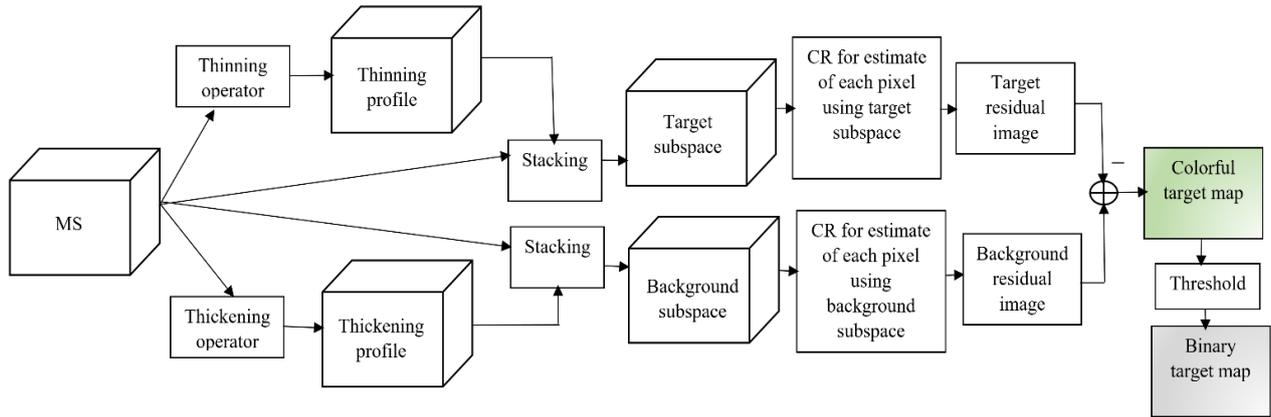


Fig. 1. Block diagram of the AP-CR method.

The number of pixels that are correctly detected is denoted by N_{cd} while the number of target pixels that are falsely detected is indicated by N_{fd} . Usually, in addition to plotting ROC, the area under curve (AUC) is computed. If a ROC curve be closer to the top left corner or its associated AUC has a higher value closer to 1, the detector has better performance. Some regions of wheat

fields are obtained through field operations as samples of the target class. The yellow regions shown in Fig. 5 are the wheat fields. The binary map of the target samples is also shown in Fig. 6. 10% of these regions are randomly selected and used as training samples. The chosen training samples are shown in Fig. 7. Similarly, 10% of background regions are used as background training samples.

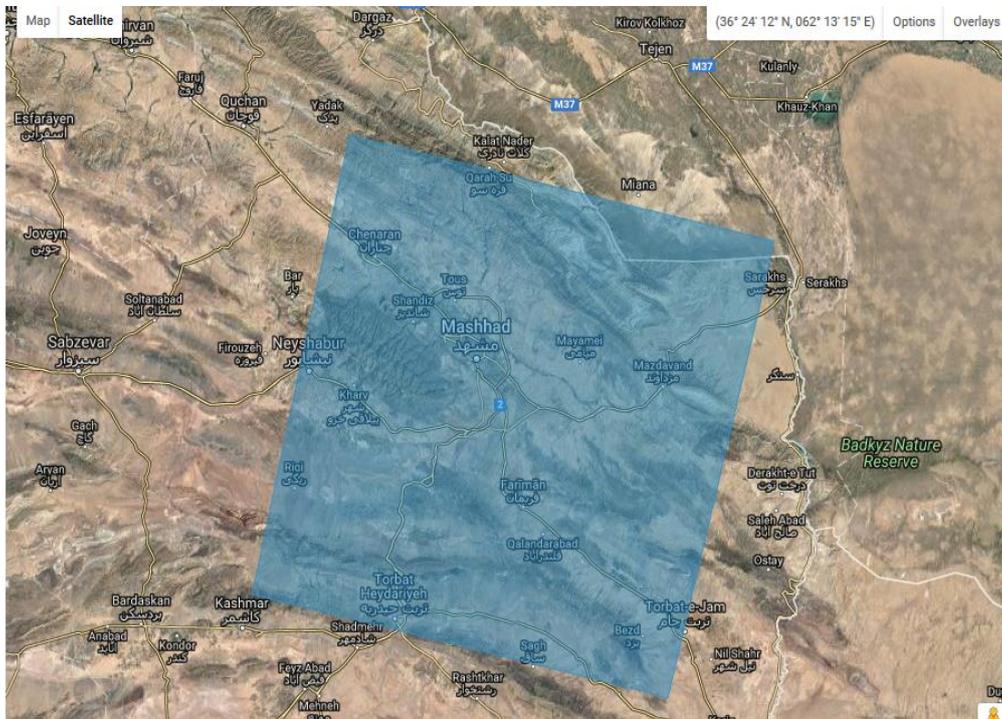


Fig. 2. An overview of Landsat 8 image outskirts of Chenaran County.

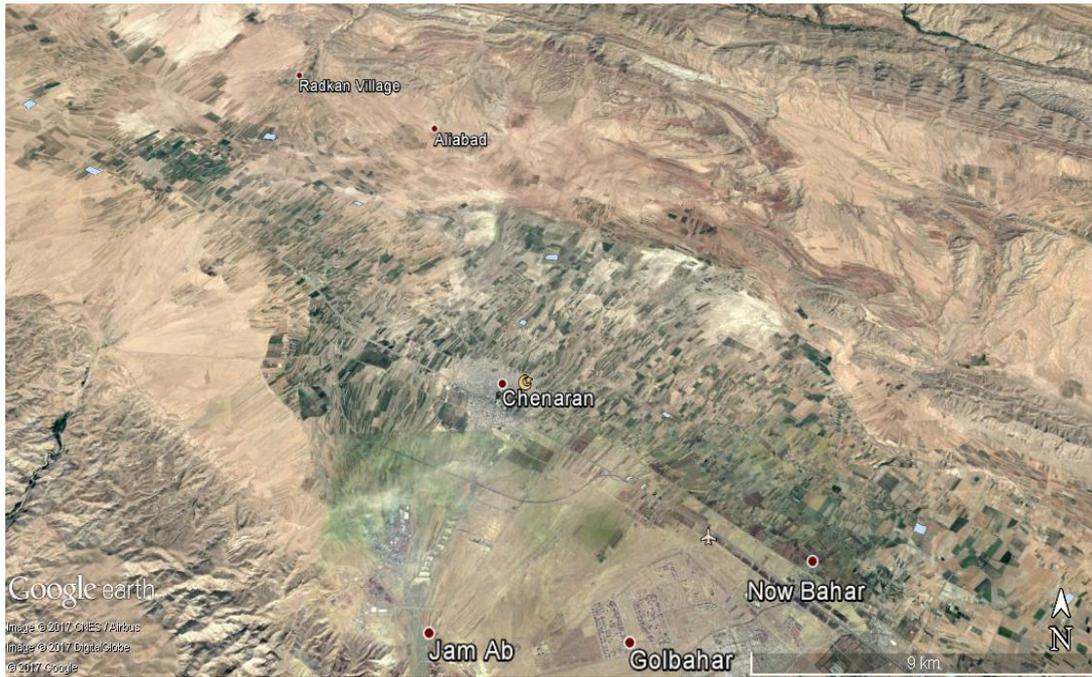


Fig. 3: Google earth image of the study area in Chenaran.



Fig. 4: Grey level image of band 6 of Landsat 8 of the study area in Chenaran.



Fig. 5: Some wheat fields as samples.



Fig. 6: Binary map of target regions.

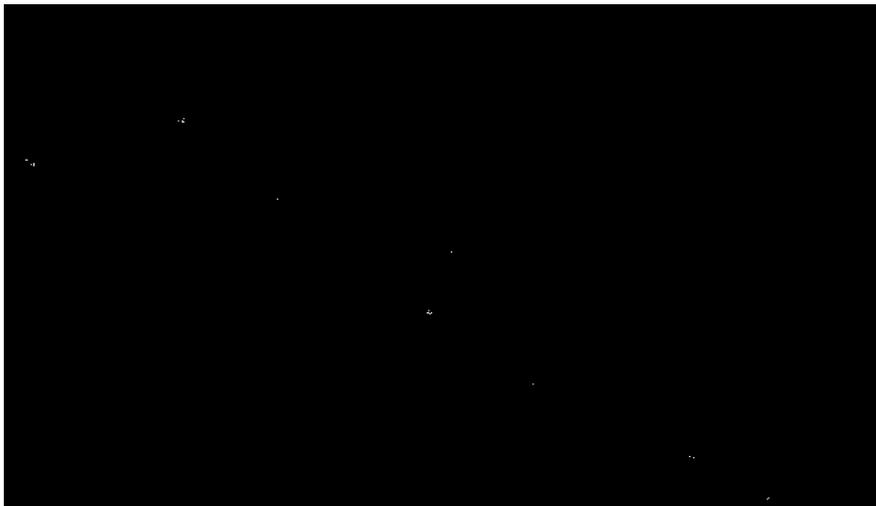


Fig. 7: Chosen training samples.

Results and Discussion

The detectors are firstly divided into two groups. The first group consists of the popular and widely used detectors (SMF, MSD, ASD, OSP, SAM, KSAM, CEM and STD).

The second group consists of two state-of-the-art recently detectors (DTD-KSAM and AP-CR). In the first experiment, the first group of detectors are applied to the MS image.

The AUC values in addition to the running time are reported in [Table 2](#). Among different methods of this group, CEM, SAM and KSAM provide the best results, respectively. STD despite spending the highest running time could not provide desired results. The ROC curves of different detectors of the first group are plotted in [Fig. 8](#). The detection map of the first rank methods are shown in [Fig. 9](#).

The map of wheat cover is obtained through Google earth monitoring.

Then, it is corrected according to field operations. The result map is shown in [Fig. 10](#) where wheat fields are illustrated with green color in the image. It can be used as a reference for comparison of the obtained target maps.

The different images obtained by applying thinning filters and thickening ones used in AP-CR method for composing the target and background subspaces are shown in [Fig. 11](#). In addition, the target residual image and the background residual image are shown in [Fig. 12](#). In the second experiment, the DTD-KSAM and AP-CR are assessed for wheat detection.

The comparison results (AUC and time) are reported in [Table 3](#). As seen, AP-CR significantly works better than DTD-KSAM for wheat detection using MS image of Landsat 8. To have a conclusion of the detection results, the AUC and running time of the best detectors of the first group are shown beside the DTD-KSAM and AP-CR detectors in [Table 4](#).

The results show that AP-CR is the best method although it needs the highest running time.

The high complexity of AP-CR is due to 1- applying the attribute filters and 2- solving the optimization problem in the collaborative representation for approximation of each image pixel.

As said before, the high ability of AP-CR is because of using valuable spatial features in addition to the spectral ones. Other competitor detectors, i.e., SAM, KSAM, CEM and DTD-KSAM just use the spectral information of the MS image. To assess the performance of these detectors in presence of spatial features, we do an experiment. In this experiment, a morphological profile containing 7 “opening” filters and 7 “closing” filters is constituted and then stacked on the original MS image to form a spectral-spatial cube.

Then, the obtained cube is fed to the SAM, KSAM, CEM, DTD-KSAM and AP-CR detectors as the input. The results are reported in [Table 5](#) and the ROC curves are shown in [Fig. 13](#).

The detection maps are also illustrated in [Fig. 14](#). By comparison between the results of [Table 4](#) (obtained by just the spectral features) and the results of [Table 5](#) (obtained by the spectral and spatial features), it can be concluded that the SAM, KSAM and CEM detectors can provide better results when both of the spectral and spatial features are used.

But, efficiency of DTD-KSAM and AP-CR without using morphological features is better than when they are used.

Note that, AP-CR has itself spatial features obtained by attribute filters and when this detector is applied to a cube of spectral and spatial features of morphological profile, a redundancy of spatial information is generated that degrades the performance of AP-CR.

However, by a comparison among all detectors of [Table 4](#) and [5](#), it is found that the AP-CR method by using the MS image containing spectral features is the best detector with the highest AUC value.

In the following, by selection of AP-CR as the appropriate target detector for MS images, the binary detection map is generated.

Table 2: AUC values and running time of different detectors in the first group

Method	SMF	MSD	ASD	OSP	SAM	KSAM	CEM	STD
AUC	55.85	49.98	49.23	19.50	70.54	69.21	82.51	50.00
Computation time (seconds)	11.24	17.16	20.43	6.43	8.41	40.92	52.19	30142.19

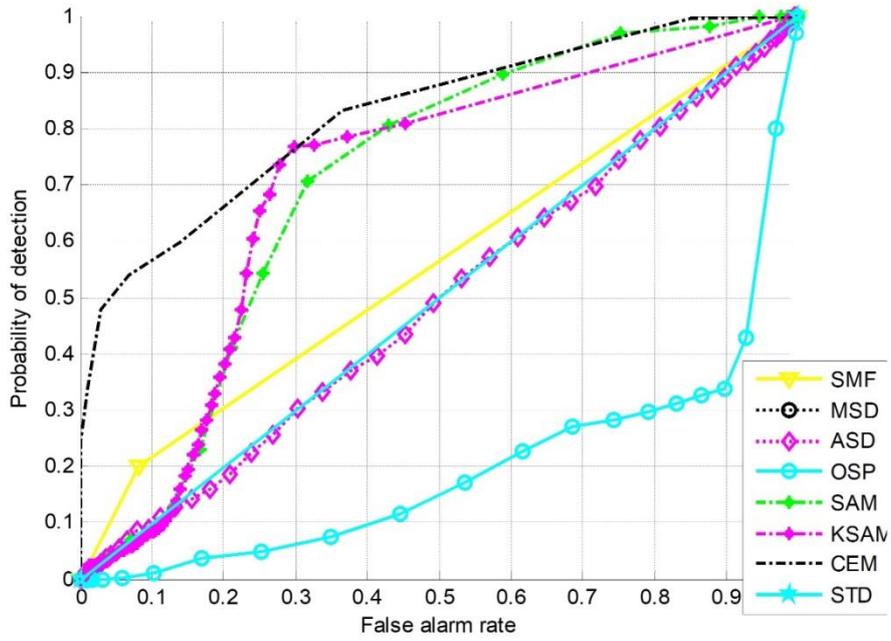


Fig. 8: ROC curves of different methods of the first group of detectors.

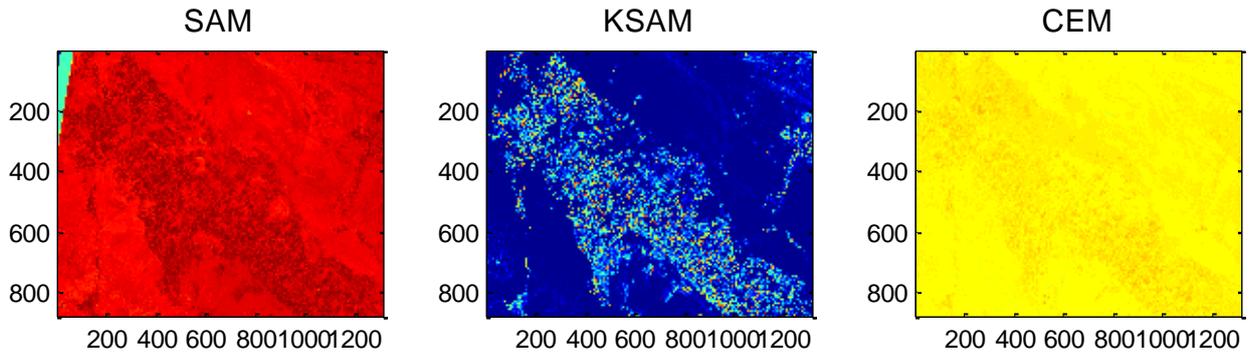


Fig. 9: Detection maps of SAM, KSAM and CEM detectors.

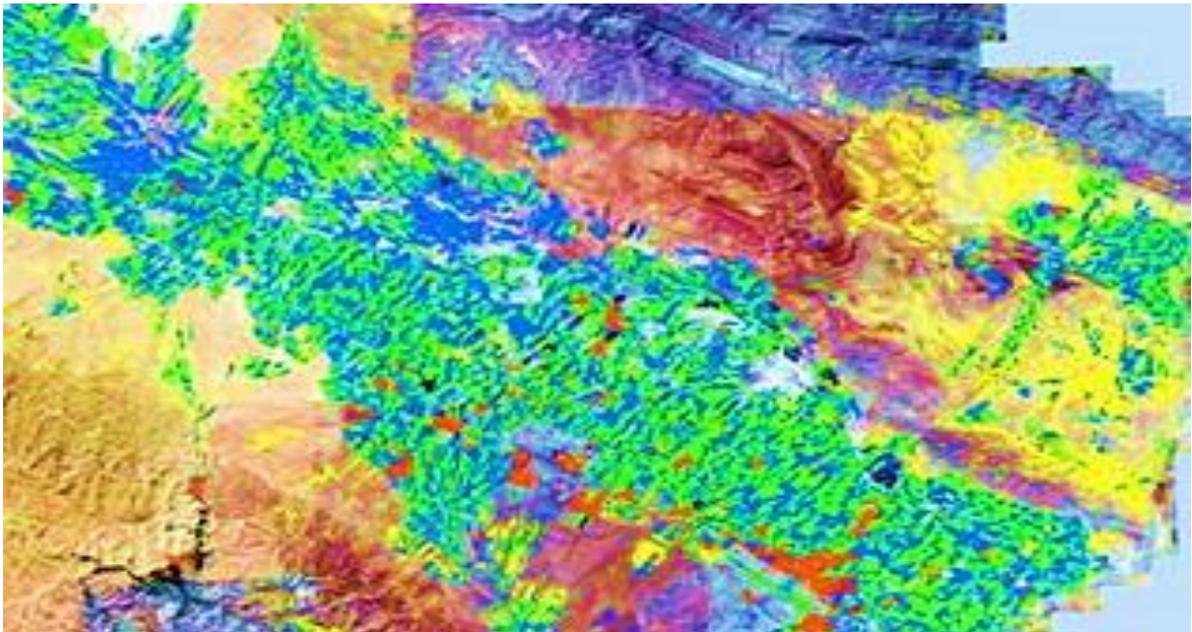


Fig. 10: The wheat cover map as reference for target detection where wheat fields are shown with green color.

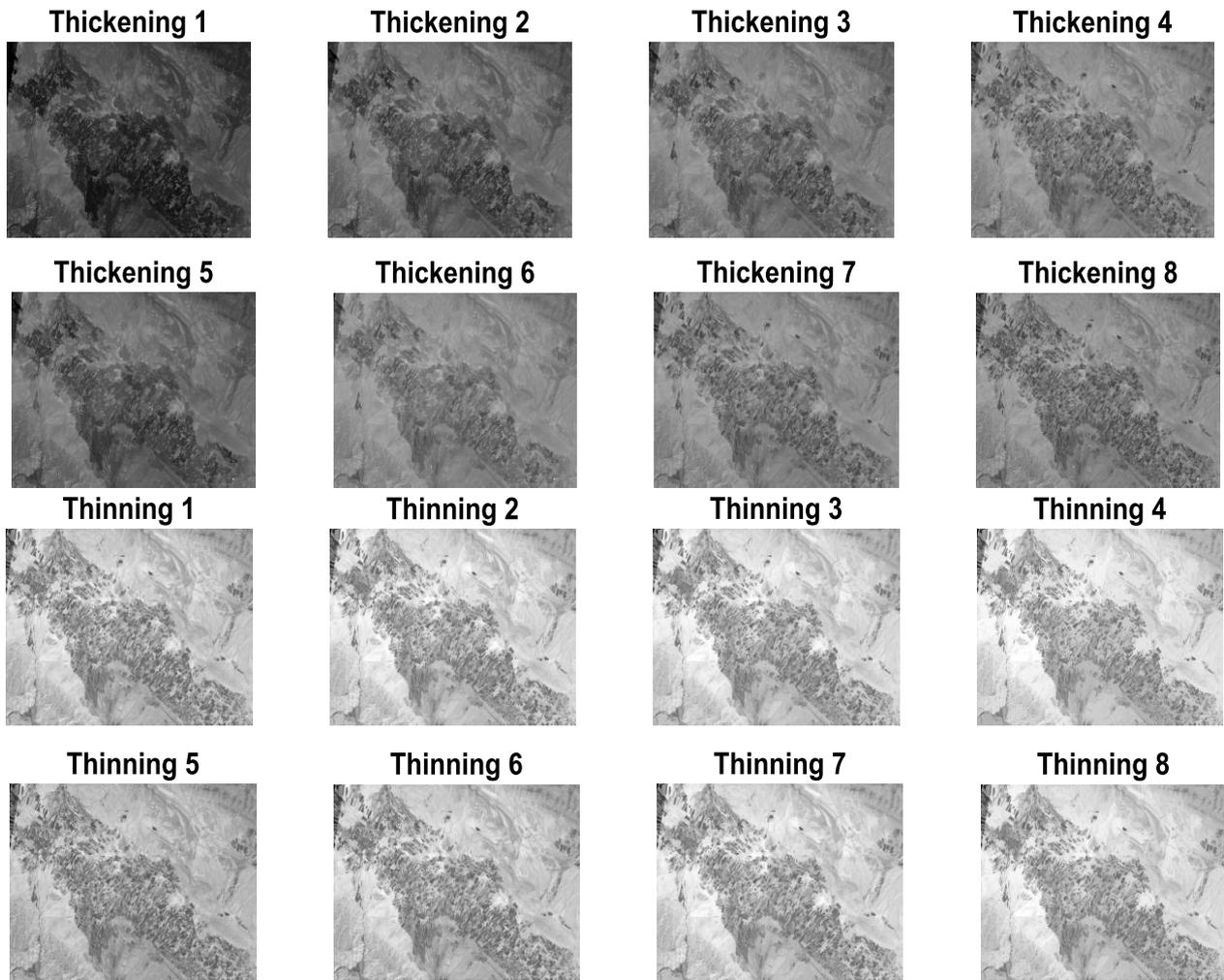


Fig. 11: Output images of thinning and thickening filters.

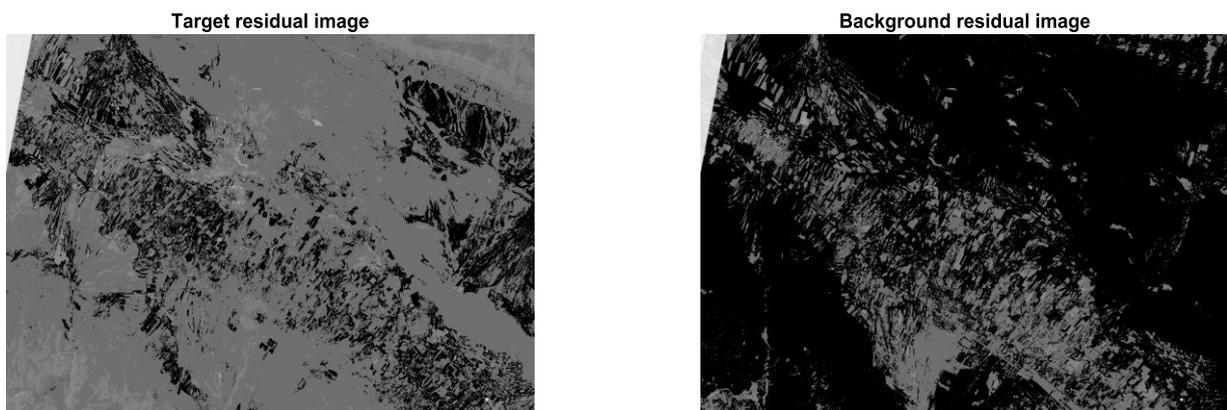


Fig. 12: The residual images.

Table 3: AUC values and running time of different detectors in the second group

Method	DTD-KSAM	AP-CR
AUC	79.88	96.09
Computation time (seconds)	68.44	6428.38

Table 4: AUC values and running time of different detectors obtained by spectral cube

Method	SAM	KSAM	CEM	DTD-KSAM	AP-CR
AUC	70.54	69.21	82.51	79.88	96.09
Computation time (seconds)	8.41	40.92	52.19	68.44	6428.38

Table 5: AUC values and running time of different detectors obtained by spectral-spatial cube

Method	SAM	KSAM	CEM	DTD-KSAM	AP-CR
AUC	75.62	72.57	85.06	78.47	93.20
Computation time (seconds)	8.85	38.07	94.51	65.38	6509.83

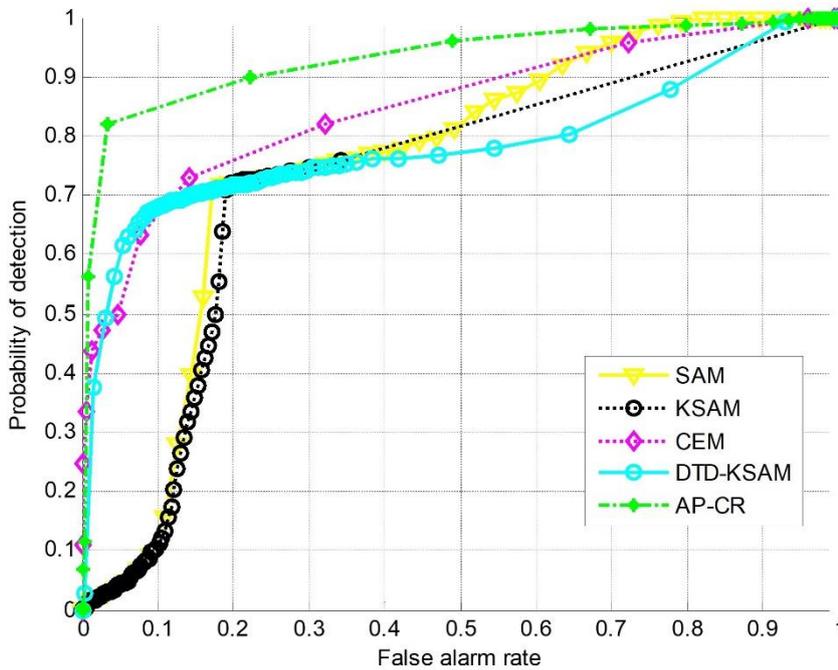


Fig. 13: ROC curves obtained by spectral-spatial cube.

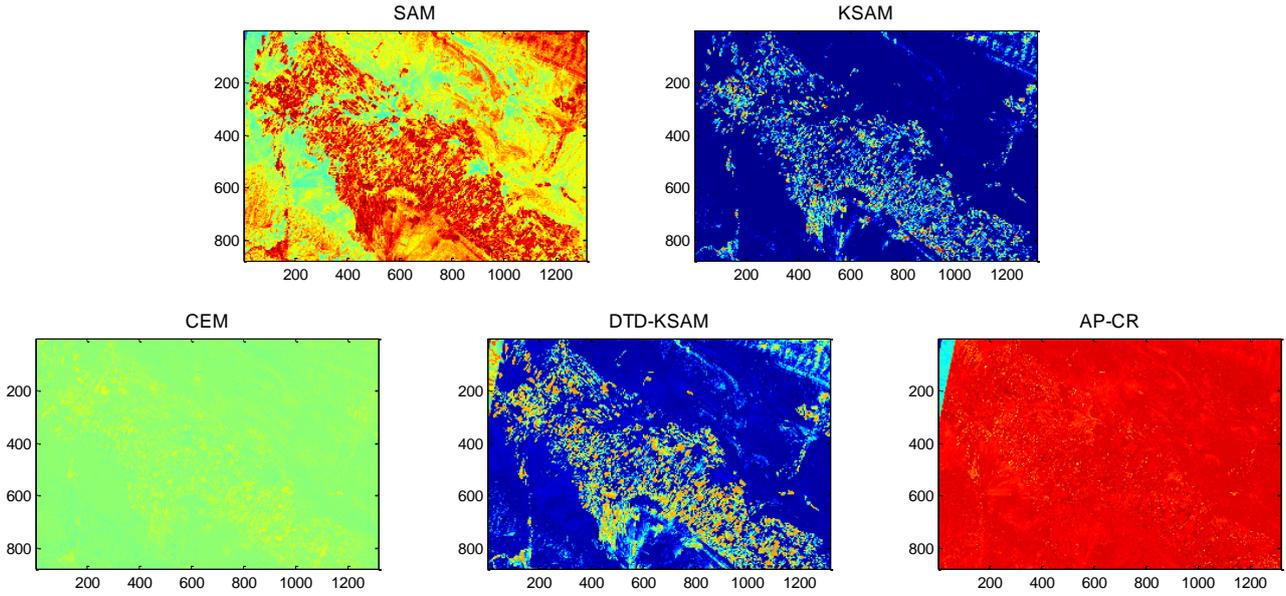


Fig. 14: Detection maps obtained by spectral-spatial cube.

As said before, the binary detection map is obtained through applying a threshold value. With considering each threshold value, a different binary map is achieved. Let $r1 = [D(y_1), D(y_2), \dots, D(y_N)]$ be the output detection for all N pixels of image. At first, the threshold value is selected by:

$$thr = \max(\max_{i=1:N} D(y_i)) \times const \quad (16)$$

where $const$ is a constant value. For instance, the binary detection maps of AP-CR obtained by $const = 0.001, 0.005, 0.01, 0.05, 0.1, 0.5$ are shown in Fig. 15. As

seen, by selection of a smaller threshold value, more details of image are labeled as target while by considering a larger value for threshold, lower points are assigned to the target class. By doing more experiments, another formula for selection of appropriate threshold is obtained:

$$thr = \min_{i=1:N} D(y_i) + 0.91(\max_{i=1:N} D(y_i) - \min_{i=1:N} D(y_i)) \quad (17)$$

The obtained binary detection map by using the above threshold is shown in Fig. 16 that is very close to the reference image of Fig. 10.

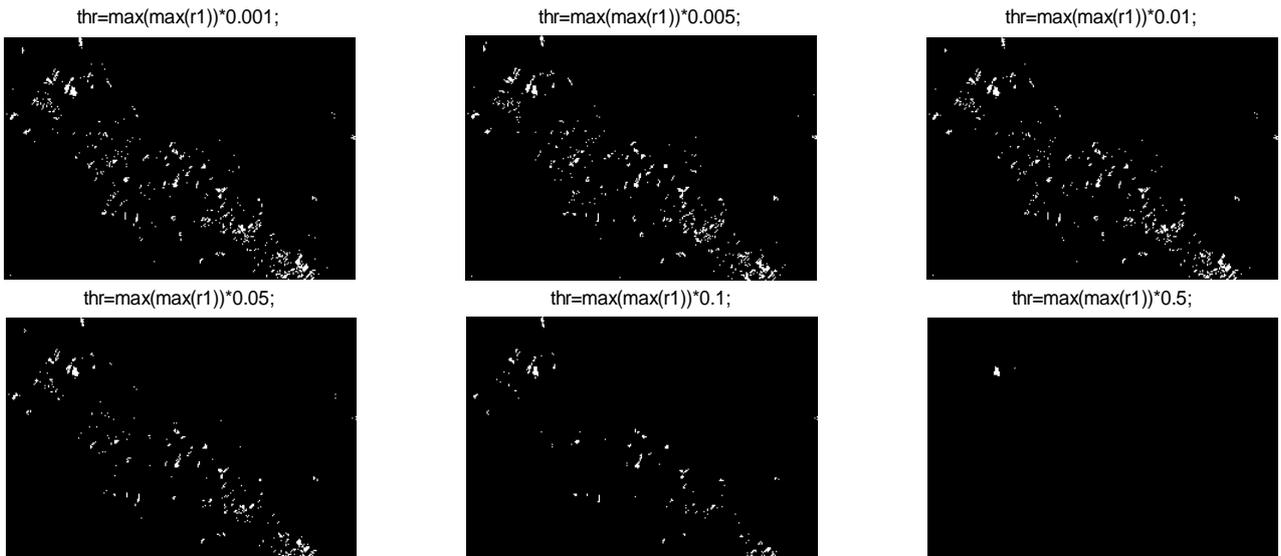


Fig. 15: Binary detection maps of AP-CR obtained by different thresholds.

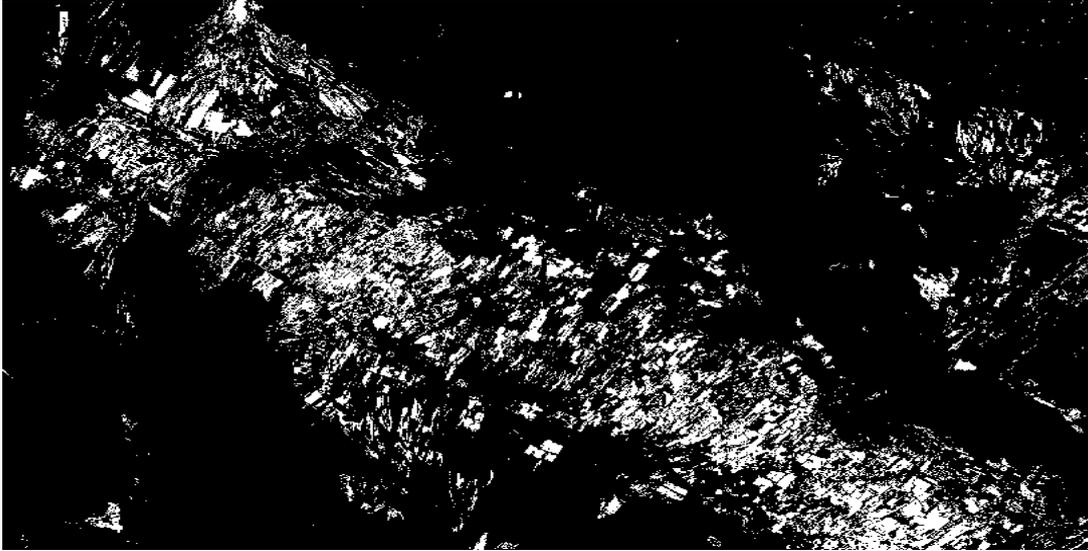
Binary Detection Map, $\text{thr}=\text{minr}+0.91*(\text{maxr}-\text{minr})$;

Fig. 16: Binary detection maps of AP-CR obtained by the chosen threshold.

Conclusion

Different hyperspectral target detectors such as SMF, MSD, ASD, OSP, SAM, KSAM, CEM, STD, DTD-KSAM and AP-CR are assessed for MS target detection in this work. Among different detectors, AP-CR is selected for MS target detection. The high performance of AP-CR is due to extraction of flexible attribute characteristics and the use of collaborative representation for approximation of each image pixel. In the collaborative representation, two individual subspaces for target and background classes are considered. The thinning profile containing details is used for composing the target subspace while the background subspace is generated by the thickening profile. The experimental results show the superior performance of AP-CR compared to other detectors whether they are applied to the original MS cube (containing just spectral features) or to the spectral-spatial cube that is result of stacking morphological profile on the MS image. Although the AP-CR method provides the highest accuracy, it needs a high running time compared to other detectors.

Author Contributions

Maryam Imani has all roles of ideation, data collecting, experiments designing, data analyzing, results interpreting, and writing the manuscript.

Acknowledgment

The author would like to thank the anonymous reviewers for their valuable comments for improving the paper.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the

ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

<i>MS</i>	Multispectral
<i>SMF</i>	Spectral matched filter
<i>SAM</i>	Spectral angle mapper
<i>MSD</i>	Matched subspace detector
<i>OSP</i>	Orthogonal subspace projection
<i>CEM</i>	Constrained energy minimization
<i>ASD</i>	Adaptive subspace detector
<i>STD</i>	Sparsity based target detector
<i>KSAM</i>	Kernel based SAM
<i>DTD</i>	Difference based target detection
<i>AP-CR</i>	Attribute profile based collaborative representation
<i>ROC</i>	Receiver operating characteristic

References

- [1] S.W. Murphy, C.R.D.S. Filho, R. Wright, G. Sabatino, R.C. Pabon, "HOTMAP: Global hot target detection at moderate spatial resolution," *Rem. Sens. Environ.*, 177: 78-88, 2016.
- [2] M. Imani, "Sparse and collaborative representation-based anomaly detection," *SIViP*, 14: 1573-1581, 2020.
- [3] T. Guo, F. Luo, L. Zhang, B. Zhang, X. Tan, X. Zhou, "Learning Structurally Incoherent Background and Target Dictionaries for Hyperspectral Target Detection," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 13: 3521-3533, 2020.
- [4] U. Kanjir, H. Greidanus, K. Oštir, "Vessel detection and classification from spaceborne optical images: A literature survey," *Remote Sens. Environ.*, 207: 1-26, 2018.
- [5] Y. Ma, J. Han, J. Liu, Y. Zhang, L. Bai, "A multi-scaled hierarchical structure model for multispectral image detection," *Signal Process. Image Commun.*, 47: 193-206, 2016.
- [6] B. Melville, A. Fisher, A. Lucieer, "Ultra-high spatial resolution fractional vegetation cover from unmanned aerial multispectral

- imagery," *Int. J. Appl. Earth Obs. Geoinf.*, 78: 14-24, 2019.
- [7] Z. Zhao, Y. Zhang, L. Bai, Y. Zhang, J. Han, "Multispectral target detection based on the space-spectrum structure constraint with the multi-scale hierarchical model," *Signal Process. Image Commun.* 68: 58-67, 2018.
- [8] F.C. Robey, D.R. Fuhrmann, E.J. Kelly, R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, 28(1): 208-216, 1992.
- [9] F.A. Kruse et al., "The spectral image processing system (SIPS) interactive visualization and analysis of imaging spectrometer data," *Rem. Sens. Environ.*, 44(2): 145-163, 1993.
- [10] L.L. Scharf, B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.* 42(8): 2146-2157, 1994.
- [11] J.C. Harsanyi, C.I. Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sens.*, 32(4): 779-785, 1994.
- [12] C.-I. Chang, D.C. Heinz, "Constrained subpixel target detection for remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, 38(3): 1144-1159, 2000.
- [13] S. Kraut, L.L. Scharf, L.T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, 49(1): 1-16, 2001.
- [14] Y. Chen, N.M. Nasrabadi, T.D. Tran, "Sparse representation for target detection in hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, 5(3): 629-640, 2011.
- [15] G. Camps-Valls, "Kernel spectral angle mapper," *Electron. Lett.*, 52(14): 1218-1220, 2016.
- [16] M. Imani, "Difference based target detection using mahalanobis distance and spectral angle," *Int. J. Remote Sens.*, 40(3): 811-831, 2019.
- [17] M. Imani, "Attribute profile based target detection using collaborative and sparse representation," *Neurocomputing*, 313: 364-376, 2018.
- [18] L. Zhang, L. Zhang, D. Tao, X. Huang, B. Du, "Hyperspectral remote sensing image subpixel target detection based on supervised metric learning," *IEEE Trans. Geosci. Remote Sens.*, 52(8): 4955-4965, 2014.
- [19] Y. Dong, L. Zhang, L. Zhang, B. Du, "Maximum margin metric learning based target detection for hyperspectral images," *ISPRS J. Photogramm. Remote Sens.*, 108: 138-150, 2015.
- [20] M. Imani, "RX anomaly detector with rectified background," *IEEE Geosci. Remote Sens. Lett.*, 14(8): 1313-1317, 2017.
- [21] T. Zhang, Z. Peng, H. Wu, Y. He, C. Li, C. Yang, "Infrared small target detection via self-regularized weighted sparse model," *Neurocomputing*, 420: 124-148, 2021.
- [22] K. Bhardwaj, S. Patra, "An unsupervised technique for optimal feature selection in attribute profiles for spectral-spatial classification of hyperspectral images," *ISPRS J. Photogramm. Remote Sens.*, 138: 139-150, 2018.
- [23] M. Dalla Mura, J.A. Benediktsson, B. Waske, L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," *IEEE Trans. Geosci. Remote Sens.*, 48(10): 3747-3762, 2010.
- [24] L. Wang, M. Li, H. Ji, D. Li, "When collaborative representation meets subspace projection: A novel supervised framework of graph construction augmented by anti-collaborative representation," *Neurocomputing*, 328: 157-170, 2019.
- [25] M. Imani, H. Ghassemian, "Weighted joint collaborative representation based on median-mean line and angular separation," *IEEE Trans. Geosci. Remote Sens.*, 55(10): 5612 - 5624, 2017.
- [26] M. Azadbakht, D. Ashourloo, H. Aghighi, H., S. Radiom, A. Alimohammadi, "Wheat leaf rust detection at canopy scale under different LAI levels using machine learning techniques," *Comput. Electron. Agric.*, 156: 119-128, 2019.
- [27] J.G.A. Barbedo, J. G.A., E.M. Guarienti, C.S. Tibola, "Detection of sprout damage in wheat kernels using NIR hyperspectral imaging," *Biosyst. Eng.*, 175: 124-132, 2018.
- [28] C.E. Dogan, N. Cebi, A. Develioglu, E.O. Olgun, E. O., O. Sagdic, "Detection of cysteine and cysteine in wheat flour using a robust LC-MS/MS method," *J. Cereal Sci.*, 84: 49-54, 2018.
- [29] Y. Seo, K.-s, Shin, "Hierarchical convolutional neural networks for fashion image classification," *Expert Syst. Appl.*, 116: 328-339, 2019.
- [30] Y. Zhong, X. Han, L. Zhang, "Multi-class geospatial object detection based on a position-sensitive balancing framework for high spatial resolution remote sensing imagery," *ISPRS J. Photogramm. Remote Sens.*, 138: 281-294, 2018.
- [31] E. Ogunlu, K. Islam, D. Perez, V.J. Hill, W.P. Bissett, R.C. Zimmerman, J. Li, "Detection of seagrass scars using sparse coding and morphological filter," *Remote Sens. Environ.*, 213: 92-103, 2018.
- [32] M. Imani, H. Ghassemian, "Attribute profile based feature space discriminant analysis for spectral-spatial classification of hyperspectral images," *Comput. Electr. Eng.*, 62: 555-569, 2017.

Biographies



Maryam Imani received the B.Sc. and M.Sc. degrees in electrical engineering from Shahed University, Tehran, Iran, and the Ph.D. degree in electrical engineering from Tarbiat Modares University, Tehran, Iran in 2009, 2011, and 2015 respectively. She continued her research in Tarbiat Modares University as a postdoc. She is an Assistant professor of the Faculty of Electrical and Computer Engineering at Tarbiat Modares University, Tehran, Iran. Her research interests include pattern recognition, signal and image processing, information analysis, and remote sensing.

Copyrights

©2020 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

M. Imani "Target detection using multispectral images, a case study: wheat detection in chenaran county in iran," *J. Electr. Comput. Eng. Innovations*, 9(1): 11-24, 2021.

DOI: [10.22061/JECEI.2020.7455.394](https://doi.org/10.22061/JECEI.2020.7455.394)

URL: http://jecei.sru.ac.ir/article_1479.html





Research paper

Improving the Security of a Low Cost Tag Search Protocol

S. Saderi Oskuiee, F. Moazami*, G. Oudi Ghadim

Department of Content Transfer Technology, Cyberspace Research Institute, Shahid Beheshti University, Tehran, Iran.

Article Info

Article History:

Received 21 May 2020
Reviewed 07 July 2020
Revised 11 September 2020
Accepted 12 November 2020

Keywords:

RFID systems
Reader
Tag
Server
Search protocol
Eavesdropping
Tracing

*Corresponding Author's Email
Address:

f_moazemi@sbu.ac.ir

Abstract

Background and Objectives: Radio Frequency Identification (RFID) systems use radio frequency waves to exchange information between a legitimate sender and a receiver. One of the important features of RFID systems is to find and track a specific tag among a large number of tags. Numerous works have been done about authentication and ownership protocols, but the number of researches done in the tag searching area is much less. Although security is a paramount factor in search protocols, but these days designers are looking for a secure search protocol that is also low cost. One way to have a low cost search protocol is that to be compatible with EPC C1G2 standard, which is an electronic product code class 1 generation 2 that works in the 860-960 MHz frequency range.

Methods: Most recently, Sundaresan et al. have proposed an RFID tag search protocol based on quadratic residues and 128-bit pseudo random number generators and XOR operation that can be easily implemented on passive tags and is compatible with EPC C1G2 standard. We show that this protocol is not immune against tag tracing, and try to improve the protocol in a way that traceability attack will not be applicable and the protocol stays low cost and EPC compatible.

Results: Since the problem in Sundaresan et al.'s search protocol is due to the tag not being able to recognize the used queries from the new ones, we improved the protocol using a counter within the queries, so the tag will realize that the query is used or not. Then we analyze the security of the improved protocol and prove its formal and informal security against known attacks.

Conclusion: In this paper, we firstly analyze the security of Sundaresan et al.'s search protocol and show that the search protocol is vulnerable to traceability attack with two different scenarios. Then we propose an improved search protocol that is secure against tracing the tags. Following that, we analyze the security of the improved search protocol.

Introduction

Radio Frequency Identification (RFID) system is a wireless technology that consists of three key parts: tags, readers, and a back-end server [1].

RFID system uses electronic and electromagnetic waves to make a conversation between a tag and a reader. A tag has an antenna and uses it for receiving and transmitting a Radio Frequency (RF) signal. It also has an integrated circuit that modulates and

demodulates these signals [2].

RFID system can easily search for a particular tag among other tags by using RF signals. A reader will send its request for finding a specific tag and the wanted tag will answer the reader's request if it was the wanted tag.

An RFID system could use an authentication protocol to find a specific tag. But if the number of tags in the searching area increases, it will be costly to use an authentication protocol [3], so it is better to use a search

protocol if we want to find a specific tag among a vast number of tags.

Early works in tag searching area were based on cryptographic functions such as hash function [3], AES [4] and symmetric encryption function [5], meaning the reader sends its request in the form of a cryptographic function and the tag also answers to reader's request with the same function.

One of the first works in the tag searching area is the serverless search protocol that has presented by Tan et al. [3]. In their protocol, tags should only answer to the requests of authenticated readers, and readers should only query authenticated tags. In the proposed protocol, the reader queries the tag using a hash function, and the tag responds to the reader using a hash function as well. The authors state that an adversary can identify the tag in this search protocol. To improve the search protocol, authors suggested different solutions. At first, they oblige the reader to use a different random number within each query and also made the tags to store the received random numbers from previous queries. Authors also state that an opportunistic adversary can track at least one tag after a specific number of queries. The second solution is to query the first m bits of tag's id (id_j) alongside with reader's and tag's random numbers n_r, r_r . Then the tag checks m bits of its own id with the one in the reader's query. If it holds, it will answer to the query. This solution does not work well when the id for each tag is structured (it means few first bits of the id is for product code, and the next bits are for the tag's origin). Another solution is when each tag receives the query from the reader, it checks the query and if the query does not belong to that specific tag, it will answer the query by probability of λ . This way, the adversary will not be able to realize if the wanted tag is present in that area or not. This protocol uses hash functions, so it is not accordant with EPC C1G2 standard. In 2011, Kim et al. [6] showed that Tan et al.'s search protocol is vulnerable against reader tracing, and then they proposed a serverless search protocol. In their protocol, the server provides an access list for each reader, and each access list contains the group of tags that are authorized to search. The proposed protocol has three stages: setup, authentication, and search. At the first stage, the server gives an access list of the tags that are authorized to search to the reader. Tags are divided evenly into small groups by server and each group has its own identity that is set by the server. Search protocol uses hash functions for queries and answers. Sundaresan et al. stated in their papers that an adversary easily could impersonate the reader and compromise tag location privacy [7], [8], [9].

In 2012, Safkhani et al. [10] showed that Tan et al.'s search protocol [3] is vulnerable against traceability

attack. They stated that in the first and third search protocols a tag could be traced.

In 2009, Lin et al. [11] proposed a serverless authentication and search protocol, which was simply an improved version of Tan et al.'s [3] search protocol. They use hash function within their protocol and aim to reduce the amount of computation in Tan et al.'s protocol. The proposed protocol is vulnerable to replay and impersonation attacks [12].

Won et al. [4] proposed a search protocol that was based on 128-bit AES cryptography function and timestamp, without a central database. In this protocol, the reader encrypts the query with 128-bit AES and sends it with a timestamp to the group of tags, tags that receive the query will check the timestamp. This way, the tags will be sure that it is not the query of the previous sessions. If the timestamp of the received query is smaller than timestamp of previous query it shows that there was a replay attack, and the tag denies the query. Otherwise, the query is correct and the tag will decrypt the query. Then, if the query is sent for that specific tag, it will answer to it. The answer to the query will also be in an encrypted form. Since all the queries are in an encrypted form, an adversary cannot get any information by eavesdropping. This protocol prevents illegal tracking of the tags and also provides secure privacy for them. DoS and de-synchronization attacks are impossible in this protocol. Since this protocol uses one way hash functions, it is not compatible with EPC C1G2 standard.

Ahamed et al. [13] proposed a serverless tag search protocol. They presented three search protocols but claim that the last one is immune against the attacks applicable to the previous two protocols. They used a pseudo random number generator P that takes a seed as an argument and a function M that generates the next random number. The reader generates a random number with P and sends it to the tag. The tag checks the validity of the received number. If the query is valid, the tag updates its own number using M . Otherwise, it will replay with probability of λ . They state that the proposed search protocol is immune against tracking, de-synchronization, and also cloning attacks.

Zuo proposed a search protocol in which the reader uses a hash function and a shared secret key to encrypt its own hashed random number along with the wanted tag's id, and the tag evaluates the query by decrypting it with its own secret key [14]. He used noise tags to guarantee that there would be an answer to the received queries. In Zuo's protocol, important secrets and data are stored in the reader. If an attacker steals the portable reader, he can perform cloning and impersonation attacks [8].

In 2011, Chun et al. [5] proposed a search protocol

that uses symmetric encryption function. Yoon showed that this protocol is vulnerable against DoS attacks [15]. Also, Chun's protocol is not compatible with EPC C1G2 standard.

Mtita et al. [12] proposed a serverless mutual authentication and search protocol in which the reader firstly downloads a list of tags from the server that are authorized to search.

Then it uses an HMAC function and a timestamp to query a tag and the tag answers the query using a random number and an HMAC function as well. However, Sundaresan et al. stated that their protocol is susceptible to DoS and de-synchronization attacks [9]. Since Mtita's protocol uses HMAC, it is not compatible with EPC C1G2 standard.

Some of the researchers tried to propose lightweight search protocols that use lightweight functions such as Physical Unclonable Function (PUF) [16], Linear Feedback Shift Register (LFSR)[16], and Nonlinear Feedback shift Register (NLFSR)[16].

Kulseng et al. [16] use LFSR to generate random numbers when sending a query and also when tag answers the query, and use PUF to authenticate the tags. Later on, Lv et al. showed that Kulsang et al.'s search protocol is vulnerable against tracing attack [18].

In 2019, Eslamnejhad namin et al. [19] proposed a lightweight search protocol that uses an encryption technique called Authentication Encryption (AE), that guarantees confidentiality and integrity at the same time. In their search protocol to query a tag, the reader firstly increases its counter and generates a random number and uses XOR to hide the wanted tag's id, the shared key between the reader and the tag, the random number and also encrypts the computed value and the counter. The tag evaluates the received query by decrypting the received query and checks if the id is also a valid one. If the query is not valid, the tag will answer with probability of λ .

Although using these cryptographic and lightweight functions in tag search protocols made them secure against some of the possible attacks on RFID protocols [20], such as eavesdropping, physical attacks, DoS attacks, and traceability.

They were not compatible with EPC C1G2 standard. To have a low cost search protocol that is compatible with EPC C1G2 standard, the tag that concerns us is the passive one, and since it has no battery inside, it is cheaper and simpler [9].

EPC C1G2 standard is an electronic product code class 1 generation 2 that works in the 860-960MHz frequency range [21]. Besides, this standard uses FHSS (Frequency Hopping Spread Spectrum). FHSS is a method of transmitting radio signals by rapidly switching a carrier among many frequency channels. When a radio signal is

transmitted, it can read the tags at slightly different frequencies to get the best possible read from the tags. Protocols that are compatible with EPC C1G2 standard use simple functions, so they are low cost protocols.

Recently, some of the researchers favor protocols that do not utilize hash functions and are compatible with EPC C1G2 standard. In 2012, Sundaresan et al. [7] proposed a search protocol that is based on quadratic residues and 128-bit pseudo random number generators and XOR operation that can be easily implemented on passive tags and is compatible with EPC C1G2 standard. But it is not secure enough. They showed later in 2017 [9] that it is not forward secure.

In 2015, Sundaresan et al. [8] proposed another search protocol that was based on a 128-bit pseudo random number generator and XOR operation and was compatible with EPC C1G2 standard, but in 2016 Jannati and Bahrak showed that it was vulnerable against de-synchronization and impersonation attacks, and tag location privacy is not satisfied [22]. Also, in 2018, Eslamnejhadnamin et al. [23] showed that the search protocol proposed by Sundaresan et al. [8] is not safe against traceability attack.

In 2017, Sundaresan et al. [9] proposed a search protocol that is based on quadratic residues and 128-bit pseudo random number generator and XOR operation that can be easily implemented on passive tags and is compatible with EPC C1G2 standard. It was the improvement of their work from 2012 [7]. We will show in this paper that this protocol is vulnerable to traceability attack.

The rest of the paper is organized as follows. We will briefly review the Sundaresan et al.'s latest search protocol. Then, we will propose a traceability attack on Sundaresan et al.'s search protocol with two different scenarios. In the next section, we will propose an improvement on their protocol. Then we analyze the security of the improved search protocol. The last section concludes the paper.

Review of Sundaresan search protocol

Sundaresan et al. [9] proposed a search protocol that is based on quadratic residues by using basic MOD, XOR, and 128-bit PRNG operation. For additional security, the protocol hides the random number generated by PRNG function in queries. The proposed protocol is consists of two phases – the setup phase and the secure search phase. In the setup phase, the server gives the reader an access list *AL* which contains the tags that are authorized to search. The second phase is where a secure search happens with the proposed protocol. Table 1 shows notations that are used in proposed protocol. In this section, we describe two phases of the protocol in details as follow:

A. Setup Phase

Assume the channel between server and reader is safe, and an adversary cannot obtain any information from conversation between the server and the reader. Server S , at first, authenticates the reader R and then gives an access list AL to the reader. This list contains all the tags that are authorized to search by the reader. This list does not include any information about secrets of the tag and its ID. And the reader only has $h(TID, t_s)$. The server also determines rts , which is a shared secret between the tag and the reader.

The reader has to store $m = g * h$ (g and h are two big prime numbers) and $h(RID)$. The tag also has to store hashed form of tag ID, and the random number s , and $n = p * q$ (p and q are two big prime numbers). The server stores TID , $h(TID)$, RID , $h(RID)$, t_s , and prime numbers p , q , g , and h , current and previous shared secret between the server and the tag s , s^{-1} and also ctr and $ctrmax$.

Table 1: Notations

Notations	Descriptions
R, S, T	Represents Server, Reader and Tag respectively
AL	Access List for the Reader
$TID, h(TID)$	Unique Tag ID and hash value of TID
$RID, h(RID)$	Unique Reader ID and hash value of RID
t_s	Secret key unique for each tag in the system, used to generate $id = h(TID, t_s)$; known only to the server
s, s^{-1}	Random number generated by server and previous value of s
p, q, g, h	Four large prime numbers generated by the server
m, n	$m=g.h$ stored in reader and $n=p.q$ stored in the tag
k	Number of the readers that can access a tag
l	Number of tags a reader is authorized to search
R_{TID}	Computed as $R_{TID} = h(TID) \oplus s$
R_{TID}^{-1}	Computed as $R_{TID}^{-1} = h(TID) \oplus s^{-1}$
rts, rts^{-1}	Shared secret between reader and a tag; previous value of rts
r_r, δ	Random numbers generated by the reader
t_r	Random number generated by the tag
$ctr, ctrmax$	Current and maximum value for counter
λ	Probability that a tag replies if query is not for that tag
\oplus, \parallel	Exclusive-OR Function (XOR) and Concatenation of two values

ctr is 0 at the beginning and will increase by 1 after each successful search and when it reaches the $ctrmax$,

the reader has to get another search authorization from the server.

B: Search Phase

The search phase is shown in Fig. 1, and it has six steps. In this phase, the reader sends its query, x and y , to the tag. If the tag is the wanted tag, it answers with α'' and t_r'' and if it is not the wanted tag, it answers with a random number with probability of λ .

Then reader sends $\alpha'', t_r'', \mu'', \delta'', r_r''$ to the server. Server validates the reader and the tag, and sends ACK' to the reader. Reader validates ACK' , and sends ACK to the tag.

Traceability Attack on Sundaesan et al. Protocol

In this section, we show that traceability attack is applicable on Sundaesan et al. [9] search protocol with two different scenarios. In the first scenario, an adversary will listen to the reader's query and saves x and y , and will block tag's response to the reader. The reader will not receive the response and will not update its own rts . At the moment, the tag will update its own rts^{-1} and rts as follows:

$$rts \rightarrow rts^{-1}, \quad (1)$$

$$PRNG(rts) \rightarrow rts. \quad (2)$$

The adversary sends the captured queries to the tags. The target tag will check the validity of captured x and y sent by the adversary with both rts^{-1} and rts . Hence it will always consider them as valid queries, therefore the target tag responds to the requests with probability of 1. On the other hand, tags that are not the wanted tag will replay to the requests with probability of λ .

In this scenario, which is inspired by Eslamnejhadnamin et al.'s work [23], suppose that there are N tags in the reader's searching area, the adversary will send a captured valid queries m times to these tags. Assume the random variable X denote the number of received answers by the adversary from the tags. If the wanted tag is not present, then the random variable X follows the binomial distribution with parameters $m \cdot (N - 1)$ and λ . Hence the expected value of X is $m \cdot (N - 1) \cdot \lambda$ and the variance is $m \cdot (N - 1) \cdot \lambda \cdot (1 - \lambda)$. The adversary counts the number of received answers. If for a real number k , $|X - m \cdot (N - 1) \cdot \lambda| < k\sigma$, then the adversary concludes that the target tag is not present. To find the optimal value of k and m such that traceability attack performs with high probability, the adversary can use Chebyshev's inequality.

By Chebyshev's inequality, if X is a random variable with expected value μ , and non-zero variance σ^2 , then for any real number $k > 0$, we have:

$$\Pr(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}.$$

So, by choosing the appropriate value of parameters k and m , the opportunistic adversary can be successful with high probability.

In the second scenario, which is inspired by Jannati and Bahrak's attack [22] on Sundaresan et al. [8], to perform the traceability attack, we assume the adversary

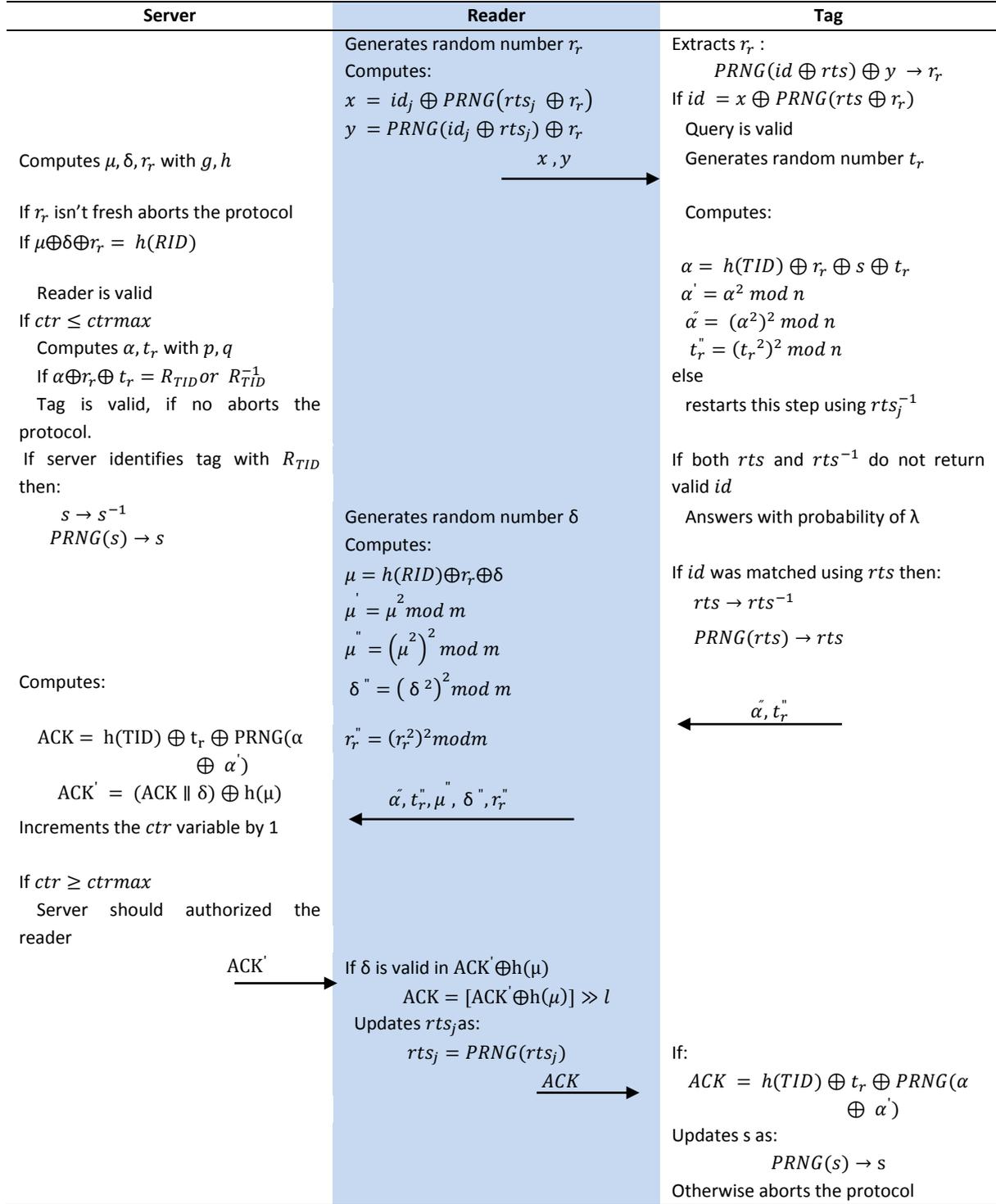


Fig. 1: Sundaresan et al. search protocol [9].

eavesdrops on the communication between the reader and the tag. Since the channel between reader and tag is insecure, adversary can store reader's queries $x = id_j \oplus PRNG(rts_j \oplus r_r)$ and $y = PRNG(id_j \oplus rts_j) \oplus r_r$ and block tag's respond.

So, the reader will not receive the respond and will not renew its own rts and at the moment, the tag will update rts^{-1} and rts . To trace a specific tag, the adversary will send reader's previous queries to tags and receive responses from them.

The target tag will check the validity of captured x and y sent by the adversary with both rts^{-1} and rts . Hence it always will consider them as valid queries, and if a tag does not answer, then the adversary concludes that it is a noise tag. Adversary repeats this process and tries to find the wanted tag.

Assume N tags exist in the reader's searching area. We know that noise tags will answer the query with probability of λ . Also, suppose that T_1 is the wanted tag and T_2, T_3, \dots, T_N , are noise tags, and all of them are available.

Let an adversary sends the captured queries m times, and let A_i be the event that T_i answers to all the received queries, hence $\Pr(A_i) = \lambda^m$. In this scenario, the traceability attack will fail if at least one of the noise tags answers all received queries. On the other hand, the probability of the attack's failure is equal to $\Pr(A_2 \cup \dots \cup A_N)$.

Also, we have:

$$\Pr(A_2 \cup \dots \cup A_N) \leq \sum_{i=2}^N \Pr(A_i) = (N-1) \lambda^m \quad (3)$$

So, if the adversary wants to limit the probability of the attack's failure to γ , the number of queries that has to send is $m = \left\lceil \frac{\log(\gamma) - \log(N-1)}{\log(\lambda)} \right\rceil$.

In this section, we show that Sundaresan et al. [9] search protocol is unprotected against traceability attack. In the next section, we will improve their protocol to prevent tracing the tag.

The traceability attack is applicable because the tag cannot recognize if the received query is reused or not. So, to improve Sundaresan et al. [9] search protocol and prevent tractability attack, we use a counter within the queries. This way the tag will recognize if a query is fresh or not.

Proposed Improved Search Protocol

The setup phase in improved protocol is like Sundaresan et al. [9] search protocol's setup phase. Also, in this phase server gives a counter (ctr_j) to the reader. The tag and the reader will increase the amount of counter after each query. After setup phase, reader starts to search for the wanted tag. At first, reader will increase its counter by 1 ($ctr_j = ctr_j + 1$), then calculates A , B and C as follow:

$$A = id_j \oplus PRNG(rts_j \oplus r_r) \oplus PRNG(ctr_j \oplus r_r) \quad (4)$$

$$B = PRNG(id_j \oplus rts_j) \oplus r_r \quad (5)$$

$$C = ctr_j \oplus PRNG(r_r \oplus id_j) \quad (6)$$

Reader then sends A , B and C to the tags in the area. Each tag will extract pseudo random number from received B using its own id and rts :

$$PRNG(id_j \oplus rts_j) \oplus B \rightarrow r_r. \quad (7)$$

Then the tag will extract the counter from received C using extracted pseudo random number and its own id :

$$C \oplus PRNG(r_r \oplus id_j) \rightarrow ctr_j. \quad (8)$$

If the counter (ctr_j) extracted from query is smaller than the one in tag's memory, the tag will notice that either replay or impersonation attack took place. Otherwise, tag will check if the id in A is equal to its own id as follows:

$$id_j = A \oplus PRNG(rts_j \oplus r_r) \oplus PRNG(ctr_j \oplus r_r) \quad (9)$$

If above check fails, tag will repeat these steps using rts_j^{-1} . If tag's id didn't match using both rts and rts_j^{-1} , then the tag will respond with probability of λ . The improved search protocol is shown in

Fig. 2.

Informal Security Analysis

In this section, we informally prove that our improved search protocol is resistant to replay, traceability, desynchronization, and Dos attacks. We also indicate that our improved protocol provides tag and reader anonymity and location privacy.

A. Replay Attack

To perform a replay attack, the adversary stores legitimate queries sent from the reader during a session then uses this information to query the tag. Since both the reader and the tag use random numbers and encipher them properly, and the reader uses a counter (ctr) to create a query. Also, both the reader and the tag update their counter after a query. So, the tag will understand that if the received message is fresh or not and will answer the old query with probability of λ .

So, the replay attack is not applicable, since every legitimate query will be fresh.

B. Traceability Attack

The weakness of Sundaresan et al. [9] search protocol is that the freshness of the queries is not guaranteed. In our improved search protocol, the reader uses a counter (ctr) in each message and increases its counter by 1 after each query.

The wanted tag also increases its counter after receiving a legitimate query. Thus, each query will be fresh, and the attacker will not notice that if the wanted tag is present or not.

So, performing a traceability attack is impossible, and tag location privacy is satisfied.

C. Tag Anonymity

In improved search protocol, the attacker cannot detect the value of tag's unique id , since it is well hidden in $h(TID, t_s)$. So, the adversary cannot obtain any information about the tag, and its anonymity is guaranteed.

D. De-synchronization Attack

If the adversary blocks answer from the tag $\{\alpha'', t_r''\}$ or the answer gets lost during the communication, this way the tag updates its own rts but the reader does not, and this causes de-synchronization of the keys. Since the tag saves its own rts_j^{-1} from the last session and checks the validity of the received query with it, a de-synchronization attack is not applicable.

E. DoS Attack

If an attacker blocks the message $\{ACK\}$ or forges it, it can cause de-synchronization between the tag and the server leading to DoS attack. If the message $\{ACK\}$ is blocked, it cannot lead to DoS attack.

Since the server keeps both s and s^{-1} , and it will be able to validate the next answers by R_{TID}^{-1} .

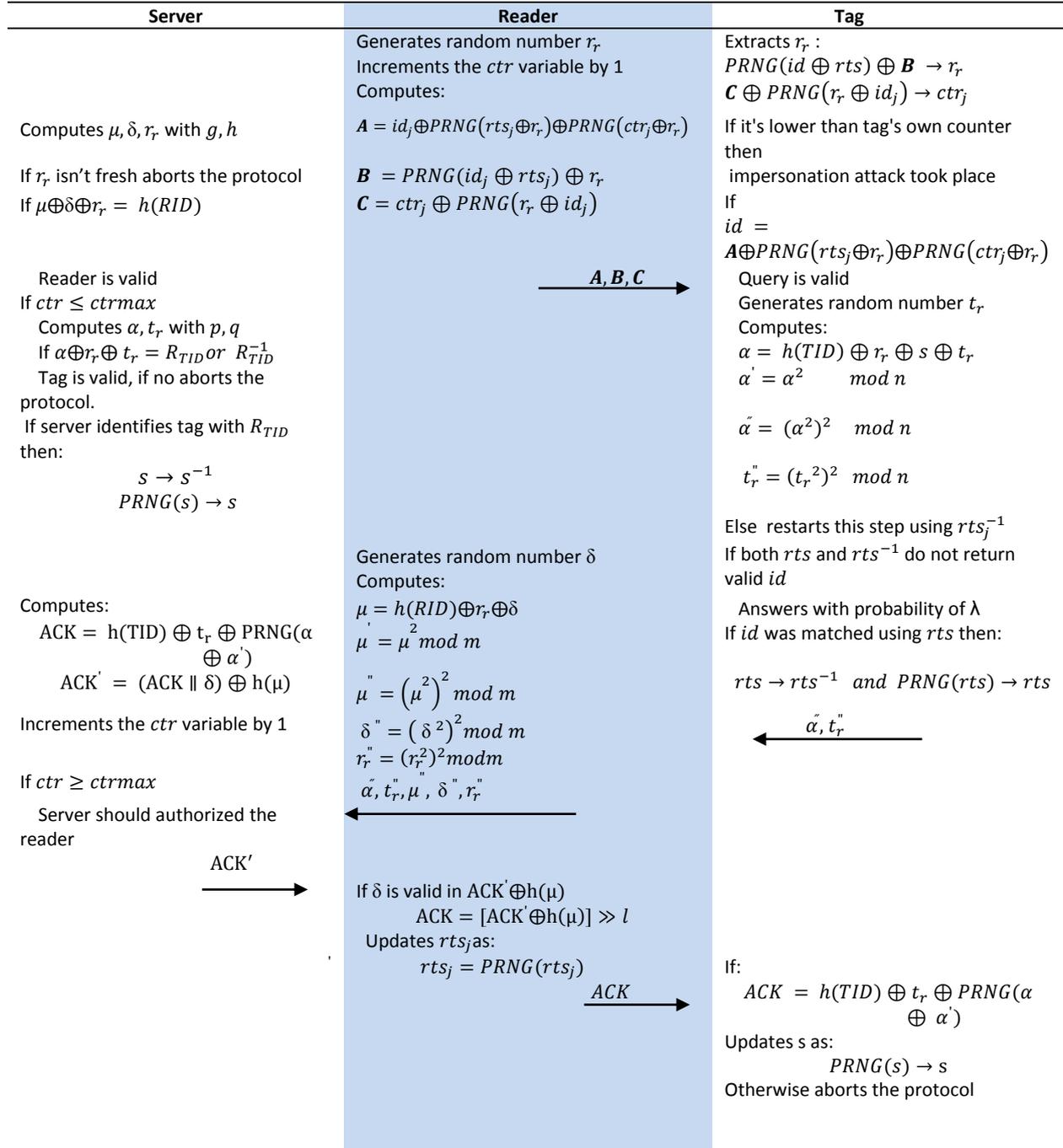


Fig. 2: Improved Search Protocol.

To forge the message $\{ACK\}$, the attacker has to know both the reader and the tag's secrets. Since it is

not possible to obtain any information about secrets as they are well hidden during the sessions, only a

legitimate server can calculate the message $\{ACK\}$. Therefore it is impossible to perform a DoS attack.

F. Reader Anonymity and Location Privacy

The reader's id is sent in a hash form $h(RID)$, and it is properly enciphered in the μ . So, if the attacker does not know the value of prime numbers g and h , he will not obtain any information. Therefore, the search protocol provides reader anonymity and location privacy.

Formal Security Analysis

In this section, we use GNY logic [24], which is a formal method to evaluate the security of improved search protocol. Table 2 shows the rules of GNY logic.

In the first step, we translate the messages of the protocol to the GNY logic parser:

Message1: $R \rightarrow T: \mathbf{A}$:

$T \triangleleft * (id_j \oplus PRNG(rts_j \oplus r_r) \oplus PRNG(ctr_j \oplus r_r))$ (The tag T receives message \mathbf{A}).

Message2: $R \rightarrow T: \mathbf{B}$: $T \triangleleft * (PRNG(id_j \oplus rts_j) \oplus r_r)$ (The tag T receives message \mathbf{B}).

Table 2: GNY logic rules [24]

Rules	Description
T1	A principal is being told of a "not-originated-here" formula.
P1	A principal is capable of possessing a formula he is told.
P2	If a principal possesses two formulas, then he is capable of possessing a function F of them.
F1	If P believes a formula X is fresh, then he is entitled to believe that any formula of which X is a component is fresh.
I1	Suppose that for principal P, all of the following conditions hold: (1) P receives a formula consisting of a X encrypted with key K and marked with a not-originated-here mark; (2) P possesses K; (3) P believes K is a suitable secret for himself and Q; (4) P believes formula X is recognizable; (5) P believes that K is fresh or that X is fresh. Then P is entitled to believe that (1) Q once conveyed X; (2) Q once conveyed the formula X encrypted with K; (3) Q possesses k.
J1	J1 states that if P believes that Q has jurisdiction over some statement C and that Q believes in C, then P ought to believe in C as well.

Message3: $R \rightarrow T: \mathbf{C}$: $T \triangleleft * (ctr_j \oplus PRNG(r_r \oplus id_j))$ (The tag T receives message \mathbf{C}).

Message4: $T \rightarrow R: \alpha'$:

$R \triangleleft * ((h(TID) \oplus r_r \oplus s \oplus t_r)^4 \bmod n)$ (The reader R receives message α').

Message5: $T \rightarrow R: t_r'' = R \triangleleft * ((t_r)^4 \bmod n)$ (The reader R receives message t_r'').

Message6: $R \rightarrow S: \alpha''$

$S \triangleleft * ((h(TID) \oplus r_r \oplus s \oplus t_r)^4 \bmod n)$ (The server S receives message α'')

Message7: $R \rightarrow S: t_r'' : S \triangleleft * ((t_r)^4 \bmod n)$ (The server S receives message t_r'')

Message8: $R \rightarrow S: \mu'' : S \triangleleft * ((h(RID) \oplus r_r \oplus \delta)^4 \bmod m)$ (The server S receives message μ'')

Message9: $R \rightarrow S: \delta'' : S \triangleleft * ((\delta)^4 \bmod m)$ (The server S receives message δ'')

Message10: $R \rightarrow S: S \triangleleft * ((r_r)^4 \bmod m)$ (The server S receives message r_r'')

Message11: $S \rightarrow R: ACK'$

$S \triangleleft * (ACK' = (h(TID) \oplus t_r \oplus PRNG(\alpha \oplus \alpha') \parallel \delta) \oplus h(\mu))$ (The reader R receives message ACK')

Message12: $R \rightarrow T: ACK: T \triangleleft * (ACK' \oplus h(\mu)) \gg l$ (The tag T receives message ACK)

Then we write the assumptions used in the improved protocol that are going to use to evaluate the correctness of the protocol:

A1: $R \ni r_r$: The reader R possesses r_r .

A2: $R | \equiv \#r_r$: The reader R believes that r_r is fresh.

A3: $R \ni rts$: The reader R possesses rts .

A4: $R | \equiv \#rts$: The reader R believes that rts is fresh.

A5: $R \ni ctr$: The reader R possesses ctr .

A6: $R | \equiv \#ctr$: The reader R believes that ctr is fresh.

A7: $T \ni t_r$: The tag T possesses t_r .

A8: $T | \equiv \#t_r$: The tag T believes that t_r is fresh.

A9: $T \ni rts$: The tag T possesses rts .

A10: $T | \equiv \#rts$: The tag T believes that rts is fresh.

A11: $T \ni ctr$: The tag T possesses ctr .

A12: $T | \equiv \#ctr$: The tag T believes that ctr is fresh.

A13: $R | \equiv R \overset{rts}{\leftrightarrow} T$: The reader R believes that rts is a suitable secret between the reader R and the tag T.

A14: $T | \equiv T \overset{rts}{\leftrightarrow} R$: The tag T believes that rts is a suitable secret between the tag T and the reader R.

A15: $T \ni s$: The tag T possesses s .

A16: $T | \equiv \#s$: The tag T believes that s is fresh.

A17: $S \ni s$: The sever S possesses s .

A18: $S | \equiv \#s$: The server S believes that s is fresh.

A19: $R \ni \delta$: The reader R possesses δ .

A20: $R | \equiv \#\delta$: The reader R believes that δ is fresh.

A21: $T | \equiv T \overset{t_s}{\leftrightarrow} S$: The tag T believes that t_s is a suitable secret between the tag T and the server S.

A22: $S | \equiv S \overset{t_s}{\leftrightarrow} T$: The server S believes that t_s is a suitable secret between the server S and the tag T.

Then we described the security correctness goals:

B1: $T | \equiv R | \sim \#id_j \oplus PRNG(rts_j \oplus r_r) \oplus PRNG(ctr_j \oplus r_r)$:

The tag T believes that the reader R conveys the formula $(id_j \oplus PRNG(rts_j \oplus r_r) \oplus PRNG(ctr_j \oplus r_r))$.

B2: $T | \equiv R | \sim \#PRNG(id_j \oplus rts_j) \oplus r_r$: The tag T believes

that the reader R conveys the formula $(PRNG(id_j \oplus r_{ts_j}) \oplus r_r)$.

B3: $T \equiv R | \sim \#(ctr_r \oplus PRNG(r_r \oplus id_j))$: The tag T believes that the reader R conveys the formula $(ctr_r \oplus PRNG(r_r \oplus id_j))$.

B4: $R \equiv T | \sim \#(h(TID) \oplus r_r \oplus s \oplus t_r)^4 \bmod n$: The reader R believes that the tag T conveys the formula $((h(TID) \oplus r_r \oplus s \oplus t_r)^4 \bmod n)$.

B5: $R \equiv T | \sim \#((t_r)^4 \bmod n)$: The reader R believes that the tag T conveys the formula $((t_r)^4 \bmod n)$.

B6: $S \equiv R | \sim \#(h(TID) \oplus r_r \oplus s \oplus t_r)^4 \bmod n$: The server S believes that the reader R conveys the formula $((h(TID) \oplus r_r \oplus s \oplus t_r)^4 \bmod n)$.

B7: $S \equiv R | \sim \#((t_r)^4 \bmod n)$: The server S believes that the reader R conveys the formula $((t_r)^4 \bmod n)$.

B8: $S \equiv R | \sim \#(h(RID) \oplus r_r \oplus \delta)^4 \bmod m$: The server S believes that the reader R conveys the formula $((h(RID) \oplus r_r \oplus \delta)^4 \bmod m)$.

B9: $S \equiv R | \sim \#((\delta)^4 \bmod m)$: The server S believes that the reader R conveys the formula $((\delta)^4 \bmod m)$.

B10: $S \equiv R | \sim \#((r_r)^4 \bmod m)$: The server S believes that the reader R conveys the formula $((r_r)^4 \bmod m)$.

B11: $R \equiv S | \sim \#(h(TID) \oplus t_r \oplus PRNG(\alpha \oplus \alpha') \parallel \delta) \oplus h(\mu)$: The reader R believes that the server S conveys the formula $(h(TID) \oplus t_r \oplus PRNG(\alpha \oplus \alpha') \parallel \delta) \oplus h(\mu)$.

B12: $T \equiv R | \sim \#(ACK' \oplus h(\mu)) \gg l$: The tag T believes that The reader R conveys the formula $(ACK' \oplus h(\mu)) \gg l$.

-Proofing B1, B2 and B3:

D1: Assume that the tag is being told of messages A, B and C (T1).

D2: by D1, assume that the tag possesses A, B and C (P1).

D3: by D2, assume that the tag believes that A, B and C are fresh (F1).

D4: From D3 (after series of steps D1, D2, D3) and assumptions A2, A4, A6, and A13 and applying postulates I1 and P2, B1 is achieved.

D5: From D3 and assumptions A2, A4, A13 and applying postulates I1 and P2, B2 is achieved.

D6: From D3 and assumptions A2, A6 and applying postulates I1 and P2, B3 is achieved.

-Proofing B4 and B5:

D7: Assume the reader is being told of messages α'' and t_r'' (T1).

D8: By D7, assume that the reader possesses α'' and t_r'' (P1).

D9: By D8, assume that the reader believes α'' and t_r'' are fresh (F1).

D10: B4 is achieved From D9 (after series of steps D7, D8, D9) and assumptions A2, A8, A16 and applying

postulates I1 and P2.

D11: B5 is achieved From D9 (after series of steps D7, D8, D9) and A8, and applying postulates I1 and P2.

-Proofing B6, B7, B8, B9 and B10:

D12: Assume the server is being told of messages α'' , t_r'' , μ'' , δ'' and r_r'' (T1).

D13: By D12, Assume the server is possesses messages α'' , t_r'' , μ'' , δ'' and r_r'' (P1).

D14: By D13, assume server believes that messages α'' , t_r'' , μ'' , δ'' and r_r'' are fresh (F1).

D15: B6 is achieved From D14 (after series of steps D12, D13, D14) and A2, A8, A16 and applying postulates I1, J1 and P2.

D16: B7 is achieved From D14 (after series of steps D12, D13, D14), A8 and applying postulates I1, J1 and P2.

D17: B8 is achieved From D14 (after series of steps D12, D13, D14), A2, A20 and applying postulates I1, J1 and P2.

D18: B9 is achieved From D14 (after series of steps D12, D13, D14) and A20 and applying postulates I1, J1 and P2.

D19: B10 is achieved From D14 (after series of steps D12, D13, D14) and A2 and applying postulates I1, J1 and P2.

-Proofing B11:

D20: Assume the reader is being told of message $\{ACK'\}$ (T1)

D21: By D20, assume that the reader possesses $\{ACK'\}$ (P1).

D22: By D21, assume that the reader believes message $\{ACK'\}$ is fresh (F1).

D23: From D22 and assumptions A8, A20 and applying postulates I1 and P2, B11 is achieved.

-Proofing B12:

D24: Assume the tag is being told of message $\{ACK\}$ (T1).

D25: By D24, assume that the tag possesses $\{ACK\}$ (P1).

D26: By D25, assume that the tag believes message $\{ACK\}$ is fresh (F1).

D27: From D26 and assumptions A8, and applying postulates I1 and P2, B12 is achieved.

Security and Performance Comparison

In this section, we compare the improved search protocol with search protocols proposed earlier.

In [Table 3](#), we compare the improved protocol with some other search protocol in terms of hash, HMAC, PUF, PRNG, LFSR, Encryption and Decryption functions, XOR, and concatenate used in them.

We observe that Tan et al.'s search protocol [3] uses hash functions, so it cannot be a lightweight protocol. Chun et al. use symmetric encryption functions when querying a specific tag [5], so their protocol cannot be EPC compatible. Sundaresan et al. use lightweight function within their protocols [8], [9], so their protocols

can be low cost and compatible with EPC standard.

Table 3: Complexity comparison

Protocol	[3]	[5]	[8]	[9]	[16]	Proposed Protocol
H1	2	0	0	0	0	0
H2	3	2	0	1	0	1
H3	2	2	6	8	4	12
H4	0	0	0	0	3	0
H5	0	0	2	3	0	5
H6	0	0	0	0	4	0
H7	0	2	0	0	0	0

H1: Number of Hash/ H2: Number of ||/H3: Number of \oplus /H4: Number of PUF /H5: Number of PRNG/ H6: Number of LFSR/ H7: Number of E(.) and D(.)

Kulseng et al. use PUF functions in their protocol [16], but it is hardly unlikely to implement PUF on a large scale [9]. In our improved protocol, the number of PRNG functions and XOR operators increased slightly, comparing to the Sundaresan et al. search protocol [9]. But since our protocol still uses lightweight functions, the cost of implementing it on RFID system will be low as well.

In Table 4, the improved search protocol is compared with some well-known search protocols based on various parameters such as tag and reader anonymity, tag and reader location privacy, EPC compliance and some attacks such as replay, tracing, DoS and de-synchronization.

Table 4: Security comparison

Protocol	[3]	[4]	[6]	[9]	[12]	[14]	IMPROVED PROTOCOL
Tag anonymity	S	S	S	S	S	S	S
Reader anonymity	NS	S	NS	S	S	S	S
Tag location privacy	NS	NS	NS	NS	S	S	S
Reader location privacy	NS	S	NS	S	S	S	S
Replay attack	NA	NA	NA	A	NA	NA	NA
Traceability attack	A	A	A	A	NA	NA	NA
DoS/ Desynchroniz ation attack	NA	NA	NA	NA	A	NA	NA
EPC compatible	N	N	N	Y	N	N	Y

Applicable (A)/ Not Applicable(NA)/ Satisfied(S)/ Not Satisfied (NS)/ Yes (Y)/ No (N)

Conclusion

In this paper, we probed the security of Sundaresan et al. search protocol [9] and discovered that an adversary could trace a tag. This weakness causes from the tag not being able to recognize an old query.

We showed that traceability attack applicable and explained it with two separate scenarios.

In the first scenario, we indicated that a traceability attack is executable, and the adversary could recognize whether the wanted tag is present with high probability by applying Chebyshev's inequality.

In the second scenario, we showed that a traceability attack is applicable with high probability if the adversary sends out a sufficient amount of queries to the tags in the area.

Following that, we proposed an improvement on Sundaresan et al. search protocol [9], that uses a counter when sending a query to ensure the freshness of the legitimate requests. Then, in the last section, security analysis of the improved protocol showed that it is immune against replay, traceability, Dos and de-synchronization attacks, and tag and reader anonymity and location privacy are maintained as well.

Author Contribution

S. Sadari, F. Moazami, G. Oudi proposed traceability attack on Sundaresan et al. search protocol. Sh. Sadari, F. Moazami proposed an improvement on the Sundaresan search protocol and analyzed security and performance of improved search protocol.

Acknowledgement

We would like to thank the anonymous reviewers for their valuable comments, which helped to increase the quality of our paper.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviation

R, S, T	Represents Server, Reader and Tag respectively
AL	Access List for the Reader
$TID, h(TID)$	Unique Tag ID and hash value of TID
$RID, h(RID)$	Unique Reader ID and hash value of RID
t_s	Secret key unique for each tag in the system, used to generate $id = h(TID, t_s)$; known only to the server
s, s^{-1}	Random number generated by server and previous value of s

p, q, g, h	Four large prime numbers generated by the server
m, n	$m=g.h$ stored in reader and $n=p.q$ stored in the tag
k	Number of the readers that can access a tag
l	Number of tags a reader is authorized to search
R_{TID}	Computed as $R_{TID}=h(TID)\oplus s$
R_{TID}^{-1}	Computed as $R_{TID}^{-1}=h(TID)\oplus s^{-1}$
rts, rts^{-1}	Shared secret between reader and a tag; previous value of rts
r_r, δ	Random numbers generated by the reader
t_r	Random number generated by the tag
$ctr, crmax$	Current and maximum value for counter
x, y	Queries sent from reader to tag in Sundaesan et al. search protocol
A, B, C	Queries sent from reader to tag in our improved search protocol
α, t_r	Wanted tag's respond to a legitimate query
$\tilde{\alpha}, \tilde{t}_r, \mu, \delta, r_r$	Legitimate reader's answers to server
ACK'	Server's acknowledgment sent to the reader
ACK	Reader's acknowledgment sent to the valid tag
λ	Probability that a tag replies if query is not for that tag
PRNG	Pseudorandom number generator
$\oplus, $	Exclusive-OR Function (XOR) and Concatenation of two values
X	Random variable
μ	Expected value of X
σ^2	Finite non-zero variance of X
k	A real number
N	Assumed number of tags in reader's searching area
m	Number of captured queries sent by an adversary
T1	Being told formula
P1	First possession rule
P2	Second possession rule
F1	Freshness rule
I1	Interpretation rule
J1	Jurisdiction rule
A	Assumptions used in security analysis
B	Security correctness goals
D	Security correctness proofs

References

- [1] S. Lahiri, RFID sourcebook, IBM press, 2005.
- [2] V. Chawla, D.S. Ha, "An overview of passive RFID," IEEE Commun. Mag., 45(9): 11–17, 2007.
- [3] C.C. Tan, B. Sheng, Q. Li, "Secure and serverless RFID authentication and search protocols," IEEE Trans. Wireless Commun., 7(4): 1400–1407, 2008.
- [4] T.Y. Won, J.Y. Chun, D.H. Lee, "Strong authentication protocol for secure RFID tag search without help of central database," in Proc. International Conference on Embedded and Ubiquitous Computing, 2: 153-158, 2008.
- [5] L. Chun, J. Hwang, D. Lee, "RFID tag search protocol preserving privacy of mobile reader holders," IEICE Electron. Express, 8(2): 50–56, 2011.
- [6] Z. Kim, J. Kim, K. Kim, I. Choi, T. Shon, "Untraceable and serverless RFID authentication and search protocols," in Proc. Ninth IEEE International Symposium on Parallel and Distributed Processing with Applications Workshops, : 278–283, 2011.
- [7] S. Sundaesan, R. Doss, W. Zhou, "A secure search protocol based on Quadratic Residues for EPC Class-1 Gen-2 UHF RFID tags," in Proc. 23rd International Symposium on Personal Indoor and Mobile Radio Communications, :30–35, 2012.
- [8] S. Sundaesan, R. Doss, S. Piramuthu, W. Zhou, "Secure tag search in RFID systems using mobile readers," IEEE Trans. Dependable Secure Comput., 12(2): 230–242, 2015.
- [9] S. Sundaesan, R. Doss, S. Piramuthu, W. Zhou, "A secure search protocol for low cost passive RFID tags," Computer Networks, 122: 70–82, 2017.
- [10] M. Safkhani, P. Peris-Lopez, N. Bagheri, M. Naderi, J. C. Hernandez-Castro, "On the security of Tan et al. serverless RFID authentication and search protocols," in Proc. International Workshop on Radio Frequency Identification: Security and Privacy Issues, 7739: 1–19, 2012.
- [11] L.C. Lin, S.C. Tsaur, S.-C. K.P. Chang, "Lightweight and serverless RFID authentication and search protocol," in Proc. Second Int. Conf. on Computer and Electrical Engineerin, 2: 95–99, 2009.
- [12] C. Mtitia, M. Laurent, J. Delort, "Efficient serverless radiofrequency identification mutual authentication and secure tag search protocols with untrusted readers," IET Inf. Secur., 10(5): 262–271, 2016.
- [13] S.I. Ahamed, F. Rahman, E. Hoque, F. Kawsar, T. Nakajima, "S3PR: secure serverless search protocols for RFID," in Proc. 2008 International Conference on Information Security and Assurance (isa 2008): 187–192, 2008.
- [14] Y. Zuo, "Secure and private search protocols for RFID systems," Inform. Syst. Front., 12(5): 507–519, 2009.
- [15] E.-J. Yoon, "Cryptanalysis of an RFID tag search protocol preserving privacy of mobile reader," in Proc. International Federation for Information Processing, : 575–580, 2012.
- [16] L. Kulseng, Z. Yu, Y. Wei, Y. Guan, "Lightweight secure search protocols for lowcost RFID systems," in Proc. 2009 29th IEEE International Conference on Distributed Computing Systems, : 40–48, 2009.
- [17] A. Falahati, H. Azizi, R.M. Edwards. "RFID light weight server-less search protocol based on nlfers," in Proc. 8th International Symposium on Telecommunications (IST), : 741-745, 2016.
- [18] C. Lv, H. Li, M. Jianfeng, B. Niu, "Vulnerability analysis of lightweight secure search protocols for low-cost RFID systems," Int. J. Radio Freq. Identif. Technol. Appl., 4(1): 3–12, 2012.
- [19] M. Eslamnezhad Namin, M. Hosseinzadeh, N. Bagheri, A. Khademzadeh, "RSPA: RFID search protocol based on authenticated encryption," J. Electr. Comput. Eng. Innovations, 6(2): 179-192, 2018.
- [20] A. Khattab, Z. Jeddi, E. Amini, E., M. Bayoumi, RFID security: a lightweight paradigm, Springer, 2016.

- [21] B. Gesuale, P. Agarwal, RFID: READ MY CHIPS!. Piper Jaffray Equity Research Report, 2004.
- [22] H. Jannati, B. Bahrak, "Security analysis of an RFID tag search protocol," *Inf. Process. Lett.*, 116(10): 618–622, 2016.
- [23] M. Eslamnezhad Namin, M. Hosseinzadeh, N. Bagheri, A. Khademzadeh, "A secure search protocol for lightweight and low-cost RFID systems," *Telecommunication Systems*, 67(4): 539–552, 2018.
- [24] L. Gong, R. Needham, R. Yahalom, "Reasoning about Belief in Cryptographic Protocols," in *Proc. 1990 IEEE Computer Society Symposium on Research in Security and Privacy*: 234–248, 1990.

Biographies



Shabnam Sadari Oskuiee received her M.Sc. degree in secure communication and cryptography from Shahid Beheshti University, Tehran, Iran, in 2018 and B.S. degree in communication engineering from Tabriz University, Tabriz, Iran, in 2015. Her research interests include RFID security and privacy.



Farokhlagha Moazami is an assistant professor at the Cyber Space Research center at Shahid Beheshti University, Tehran, Iran, since 2013. She received B.S. and Ph.D. degrees in mathematics from Zahra University, Tehran, Iran, in 2004 and 2012, respectively and M.S. degree in mathematics from Sharif University of Technology, Tehran, Iran, in 2006. She was a postdoctoral at Sharif University of Technology, Tehran, Iran, from 2012 to 2013. Her main research interest is theoretical and practical aspects of cryptography.



Gelare Oudi Ghadim received her M. Sc. degree in secure communication and cryptography from Shahid Beheshti University, Tehran, Iran, in 2018 and B.S. degree in communication engineering from Sistan and Baluchestan University, Zahedan, Iran, in 2015. Her research interests include RFID security and privacy.

Copyrights

©2020 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

S. Sadari Oskuiee, F. Moazami, G. Oudi Ghadim, "Improving the security of a low cost tag search protocol," *J. Electr. Comput. Eng. Innovations*, 9(1): 25-36, 2021.

DOI: [10.22061/JECEI.2020.7342.383](https://doi.org/10.22061/JECEI.2020.7342.383)

URL: http://jecei.sru.ac.ir/article_1481.html





Research paper

A Novel Method for Medical Image Segmentation Based on Convolutional Neural Networks with SGD Optimization

M. Taheri^{1,*}, M. Rastgarpour², A. Koochari¹

¹Department of Computer Engineering, Faculty of Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran.

²Department of Computer Engineering, Faculty of Engineering, Islamic Azad University, Saveh Branch, Saveh, Iran.

Article Info

Article History:

Received 10 May 2020
Reviewed 01 July 2020
Revised 08 September 2020
Accepted 12 November 2020

Keywords:

Deep learning
Convolutional neural networks
Medical image segmentation
Image processing
Computer vision

*Corresponding Author's Email Address:

mohammad-taheri@srbiau.ac.ir

Abstract

Background and Objectives: medical image Segmentation is a challenging task due to low contrast between Region of Interest and other textures, hair artifacts in dermoscopic medical images, illumination variations in images like Chest-Xray and various imaging acquisition conditions.

Methods: In this paper, we have utilized a novel method based on Convolutional Neural Networks (CNN) for medical image Segmentation and finally, compared our results with two famous architectures, include U-net and FCN neural networks. For loss functions, we have utilized both Jaccard distance and Binary-crossentropy and the optimization algorithm that has used in this method is SGD+Nestrov algorithm. In this method, we have used two preprocessing include resizing image's dimensions for increasing the speed of our process and Image augmentation for improving the results of our network. Finally, we have implemented threshold technique as postprocessing on the outputs of neural network to improve the contrast of images. We have implemented our model on the famous publicly, PH2 Database, toward Melanoma lesion segmentation and chest Xray images because as we have mentioned, these two types of medical images contain hair artifacts and illumination variations and we are going to show the robustness of our method for segmenting these images and compare it with the other methods.

Results: Experimental results showed that this method could outperformed two other famous architectures, include Unet and FCN convolutional neural networks. Additionally, we could improve the performance metrics that have used in dermoscopic and Chest-Xray segmentation which used before.

Conclusion: In this work, we have proposed an encoder-decoder framework based on deep convolutional neural networks for medical image segmentation on dermoscopic and Chest-Xray medical images. Two techniques of image augmentation, image rotation and horizontal flipping on the training dataset are performed before feeding it to the network for training. The predictions produced from the model on test images were postprocessed using the threshold technique to remove the blurry boundaries around the predicted lesions.

Introduction

Medical image segmentation is one of the most important and vital tasks that in recent years, many

computer-based procedures have developed to do this task precisely. Likewise, in this paper we have introduced

an algorithm for automatic medical image segmentation toward segmenting melanoma cancer images and Chest-Xray medical images.

Actually, this important and vital task contains the process of automatic or semi-automatic detection of boundaries within a 2D or 3D image. A major difficulty of this task is the high variability in medical images. First of all, human's anatomy and body's organs include lots of variations in each person. After that, different imaging modalities (X-ray, CT, MRI, Microscopy, PET, SPECT, Endoscopy, OCT, and etc.) that use to create medical images from different organs of patients. As we have mentioned before, there are some problems in medical images such as hair artifacts in dermoscopic medical images and the images which are related to the hairy parts of body.

Segmentation task in medical images resulted in separate the region of interest (ROI) from other regions that exist in different medical images. This task helps doctors for further diagnostic insights and cure the patients. Possible applications are automatic measurement of organs, cell counting, or simulations based on the extracted boundary information.

This paper is organized as following. first, we discuss about previous works in section II, then we introduce the details of the proposed segmentation method in Section III, then report the experimental design and results in Section IV. Finally, we conclude our study in Section V.

Previous Works

In the past years, researches introduced many computer-based algorithms to solve medical image segmentation problem. Some of these efforts include:

Kim.J.U in [1], proposed an iterative structure in deep neural networks for medical image segmentation. in this method an encoder-decoder structure mixed with iterative structure until they can improve the segmentation results that include complex shapes in medical images. they utilized two technics include transfer-learning and data augmentation for prevent overfitting.

Chang.Y in [2], proposed a method for segmenting cardiac medical images. first of all, ROI detects by YOLO neural network then it feeds to a convolutional neural network for segmentation. finally, by a fully connected neural network, the segmented images classify and then the doctor can detect the type of cardiac disease precisely.

Isin in [3], discussed about different methods for brain image segmentation. these methods divide to three categories include: non-automatic, semi-automatic and automatic methods. In automatic methods, there is no intervention from human and because of this the probability of error decreases.

Pim in [4], could train a CNN for segmenting six

texture in MRI brain images, Chest muscle in muscle MRI images and vessel in CT cardiac images. then this method is suitable for segmentation of different organs and in different imaging conditions.

Dina Abdelhafiz in [5], proposed a method by Unet neural networks for addressing the lesions that exist in mammogram images precisely. In this paper, the accuracy improved, by batch normalization and data augmentation.

Hauzhu Fu in [6], developed a method for eye vessels segmentation by fully connected neural networks. this network produces the feature map of eye vessels and feeds it to CRF neural network until produces the binary images of eye vessels.

Zhe Gou in [7], analysed medical images in different levels include: feature learning, classifier level and decision making. They developed a neural network until detect malignant tumors in different imaging conditions (computerized tomography, PET and MRI).

Celebi et al in [8], used a procedure for skin lesion detection by using four threshold methods.

Zhou et al in [9], optimized the automatic skin lesion segmentation by mean shift estimation. This method needs many computational sequences.

Xie et al in [10], combined a neural network method with genetic algorithm for skin lesion segmentation.

Recently, convolutional neural networks are one of the most important and practical algorithms in the image segmentation task, specifically for medical image segmentation [11]. This algorithm has been used for brain tumor segmentation in MR brain images [12], cardiac image segmentation [13], skin lesion segmentation by non dermoscopic images [14] and many other medical image segmentation tasks which are related to different parts of the body with different shapes of organs and various artifacts.

Proposed method

Stochastic Gradient Descent (SGD) is the optimization algorithm which has been used in this method. It is an iterative method for optimizing an objective function with suitable smoothness properties. It can be regarded as a stochastic approximation of gradient descent optimization, since it replaces the actual gradient by an estimate thereof. Especially in high dimensional optimization problems, it reduces the computational burden, achieving faster iterations in trade for a lower convergence rate. A good overview with convergence rates can be found in [15]. Given a set of training examples $(x_1, y_1), \dots, (x_n, y_n)$ where $x_i \in R^m$ and $y_i \in R (y_i \in -1, 1 \text{ for classification})$, our goal is to learn a linear scoring function $f(x) = w^T x + b$ with model parameters $w \in R^m$ and intercept $b \in R$. In order to make predictions for binary classification, we simply look at the sign of $f(x)$. To find the model parameters, we

minimize the regularized training error given by following formula:

$$E(w, b) = \frac{1}{n} \sum_{i=1}^n L(y_i, f(x_i)) + \alpha R(w) \quad (1)$$

In this method, there are two preprocess steps that have been done on medical images which feed to the network. These steps include resizing the images and image augmentation. After that, preprocessed images enter to the network for segmentation process. The outputs are segmented images which are blurred. We have solved this problem by a post processing step which uses threshold technique to make the segmented images sharper. We have shown the sequence of our algorithm in Fig. 1.

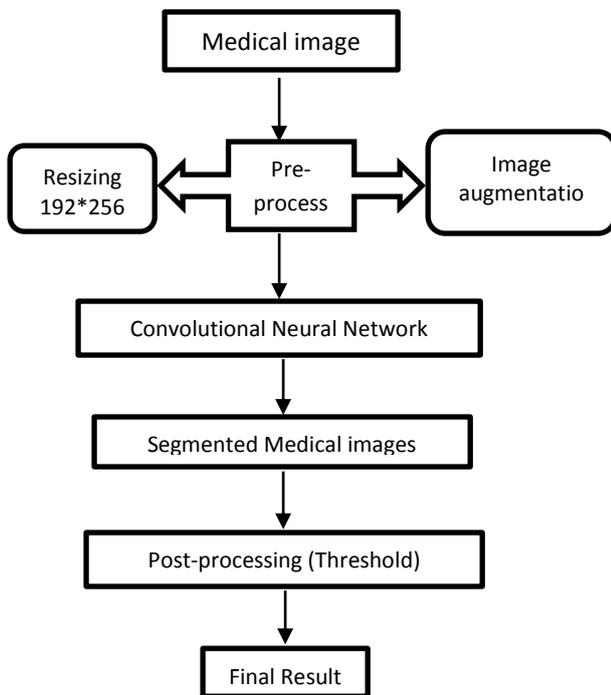


Fig. 1: The sequence of implemented algorithm.

A. Network Architecture

The proposed deep neural network contains an encoder decoder structure that includes 64-layered network excluding the final activation layer. Every sequence of encoder has multiple convolutional layers, batch normalized with ReLU nonlinearity which is followed by non-overlapping max pooling and sub-sampling. At the center of the network there are two dense layers present before the first up-sampling begins. The defining characteristic of proposed network is the use of max pooling indices in the decoders to perform up-sampling of low-resolution feature maps. This leads to keeping of the important detailed features in the image and non-useful features are erased. This neural

network provides smooth images without any post-processing technique involved. This method could outperform two FCN and Unet architectures.

FCN, despite up-convolutional layers and a few shortcut connections produces coarse segmentation maps. Therefore, more shortcut connections are introduced. However, instead of copying the encoder features as in FCN, indices from max-pooling are copied. This proposed method more memory-efficient than FCN. Also, in proposed method only the pooling indices are transferred to the expansion path from the compression path, using less memory. Whereas in the Unet, entire feature maps are transferred from compression path to expansion path making, using a lot of memory. Also, the outputs producing by this method are smooth and the borders between the region of interest and surrounded areas can't be detected precisely by the physicians, because of this we used a threshold technique for making the results sharper as post processing technique which is clarified in the next parts of this manuscript. The architecture of proposed neural network has been shown in Fig. 2.

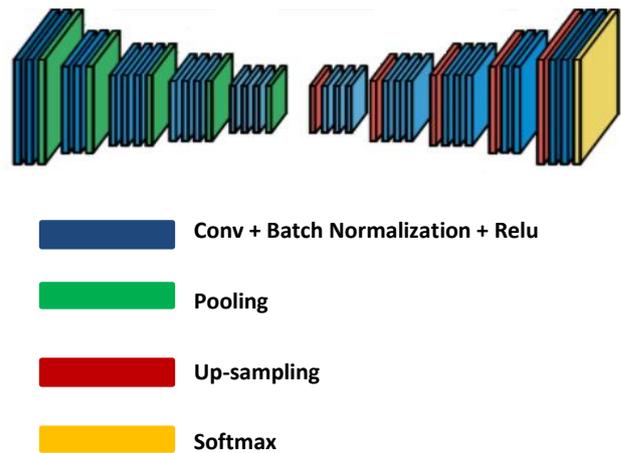


Fig. 2: Architecture of proposed neural network.

The architecture consists of a sequence of non-linear processing layers (encoders) and a corresponding set of decoders followed by a pixel-wise classifier. Typically, each encoder consists of one or more convolutional layers with batch normalization and a ReLU non-linearity, followed by non-overlapping max-pooling and sub-sampling.

The sparse encoding due to the pooling process is up-sampled in the decoder using the max-pooling indices in the encoding sequence. One key ingredient of this method is the use of max-pooling indices in the decoders to perform up-sampling of low resolution feature maps. This has the important advantages of retaining high frequency details in the segmented images and also

reducing the total number of trainable parameters in the decoders. The entire architecture can be trained end-to-end using stochastic gradient descent.

B. Loss Function

The loss functions used here are both Binary-crossentropy and Jaccard distance. The cross-entropy is a function which measures how far away from the true value the prediction is for each of the classes and then averages the errors class wise to obtain the final loss. In this problem, there lies only two classes for each pixel, either black or white (0 or 1) as per the mask. So, here binary cross-entropy is used as the loss function rather than the categorical cross-entropy originally proposed. The binary cross-entropy is in the below form:

$$l(y, \hat{y}) = -\frac{1}{N} \sum_{i=0}^N (y * \log(\hat{y}_i) + (1 - y) * (1 - (\hat{y}_i))) \quad (2)$$

Moreover, we have utilized Jaccard distance as loss function, but eventually we have reported our results based on binary cross entropy, because it had more accurate and higher evaluation metrics. Jaccard distance defines as following [16] :

$$d_j(M, C) = 1 - J(M, C) = 1 - \frac{|M \cap C|}{|M| + |C| - |M \cap C|} \quad (3)$$

M represents the ground truth of segmentation, which is normally a manually identified tumor region, and C represents a computer-generated mask. $d_j(M, C)$ itself is not differentiable, which makes it difficult to be directly applied into backpropagation.

C. Network Training

For the training process, we have utilized 75% of 200 images that are available in the PH2 dataset [17]. moreover, the actual number of images in both PH2 dataset and Chest-Xray images are more than 150, because after image augmentation technique as preprocessing, the number of images increases to 450 and then enter to the network. We have separated these images into training and validation dataset for improving the network performance. 20% of these 450 image associate to validation dataset and the rest of them associate for training dataset. As per the architecture of the network the total parameters to be trained are 33,377,795 out of 33,393,669 whereas the non-trainable parameters are 15,874. The implementations are in Keras and the environment used is the google colab, cloud space for python codes.

D. Image Augmentation

The procedure of image augmentation on training images has been used for increasing the robustness of the model and reducing the chances of overfitting. Also, it will increase the images that are available in the

dataset. The two simple techniques that have been used in this method are image rotation and horizontal flipping [18].

In the image rotation, the images are rotated around the [-40, +40] degrees and for horizontal flip, the images are reversed around horizontal axis. All the above transformations are exactly performed on the corresponding masks of the images as well to maintain the correct orientation of feature images with their truth masks. As we have mentioned in the last part, after the augmentation the transformed images are included in the training set which increases our training set from 150 to 450. Out of these 450 images, 90 (20%) have been excluded from training set for the formation of a validation set and the rest of images associate to training dataset. Results of our image augmentation technique have been displayed in Fig. 3.

In the image augmentation technique, both original medical images and their corresponding masks have been affected by this technique and because of this, the total number of training images that are going to feed to the neural network, increase to 450.

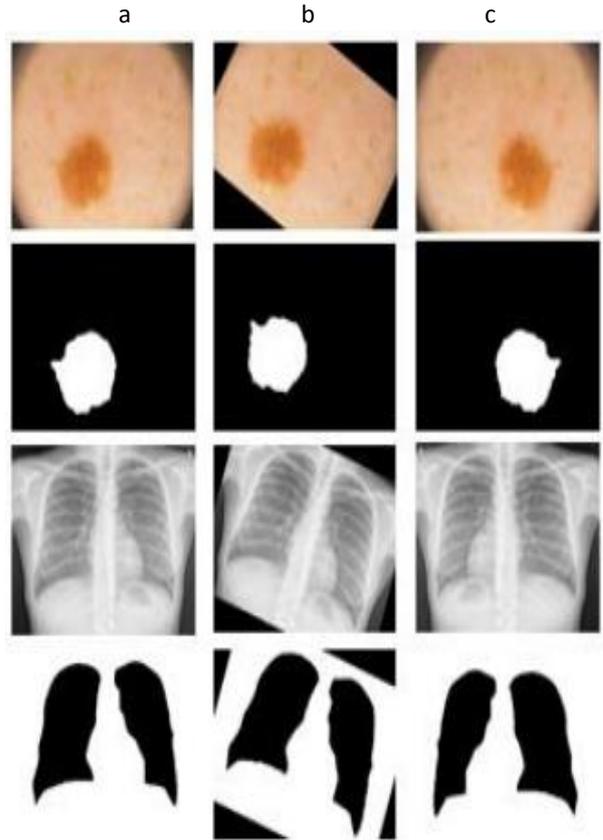


Fig. 3: image augmentation technique. a) original medical images with their corresponding masks. b) rotated images with their corresponding masks. c) horizontal flipped images with their corresponding masks.

E. Batch normalization

Batch normalization [14] is the technique of speeding

up the learning process of the neural network by normalizing the values in the hidden layers similar to the principle behind the normalization of the features in the data or activation values.

In the proposed network the batch normalization layer is presented after every convolution layer with a total of 25 batch normalization layers in the entire network architecture.

Experimental Results and Design

A. Datasets

The used medical datasets include PH2 dermoscopic dataset which contains 200 dermoscopic images and their label masks and Chest-Xray images that contains 200 RGB images and their corresponding label masks. The dimension of dermoscopic images is 572*765 and the dimension of Chest-Xray images is 512*512. The dataset has been provided publicly for experimental and studying purposes, to facilitate research on both segmentation and classification algorithms of medical images.

The PH2 dataset is acquired at the Dermatology Service of Hospital Pedro Hispano, Matosinhos, Portugal and Chest-Xray dataset has provided in Kaggle [19]. Before feeding the medical images to the network, we have two preprocessings that one of them is resizing the images for speeding up the neural network. In this work, we resize all of the medical images into 192*256 and make them smaller.

It largely reduces the parameters to be trained in the network as well as the training time and complexity without significantly affecting the results. We train and evaluate our model with RGB 8 bit images and test it by label masks. The size of evaluation set is 20% of train set and the size of test set is 25% of the label masks set (50 images).

B. Performance Evaluation

The outputs of the proposed neural network are binarized in to the lesion masks. The performance of the proposed algorithm has assessed by comparing computer generated masks as output with ground truths which are provided in the label masks dataset. We have assessed our model's performance by the following metrics:

C. Jaccard Index (JA)

This metric also has known as Intersection Over Union.

The Jaccard similarity coefficient is a statistical similarity measurement to check the diversity among the sample sets.

The IOU gives the similarity among sets and the formula is the size of the intersection over the size of the union of the sets.

we can define evaluation metrics based on confusion matrix components.

Let TP, TN, FP, FN refer to the number of true positives, true negatives, false positives, and false negatives. We can show Jaccard distance by JA and define it as following formula:

$$JA = \frac{TP}{TP + FN + FP} \quad (4)$$

TP is the number of pixels that have been predicted as ROI correctly, FN is the number of pixels that have not been predicted as ROI wrongly and FP is the number of pixels that have been predicted as ROI wrongly.

D. Dice Coefficient

The Dice score (F1 score) is like precision. It measures the positives as well as it applies penalty to the false positives given by the model.

It is more similar to precision than accuracy. We have shown dice coefficient by DI and have calculated it by the following formula:

$$DI = \frac{2 * TP}{(TP + FP) + (TP + FN)} \quad (5)$$

E. Recall

Recall is a measure which is targeted towards the actual or the true positives yielded by the model output. In the scenarios where the cost of the False Negatives is greater than recall is the better metric to choose the best model among the possible ones. Recall also known as sensitivity (SE). we can calculate Recall as following formula:

$$SE = \frac{TP}{TP + FN} \quad (6)$$

F. Accuracy

Accuracy approximates that how close our predicted output to the real ground truth. We have shown it by AC and have calculated it as following formula:

$$AC = \frac{TP + TN}{TP + FN + TN + FP} \quad (7)$$

Comparison with Two Famous Architectures

Moreover, we have compared our results with two famous architectures, include Unet and FCN networks and the methods that have implemented in [16], [20]. We have shown our results in Table 1 by using PH2 dataset.

Also, we have implemented our method on Chest-Xray dataset to develop the algorithm on different medical images. We have shown our results by using

Chest-Xray medical images and compare it with other methods in [Table 2](#).

Table 1: Comparison results with other methods, based on different optimization algorithms, using PH2 dataset

Method	AC	DI	JA	SE
Yuan [16]	96.3	92.2	86.1	92.6
FCN	91.7	64.8	87.8	82.2
Unet	90.9	85.3	92.1	93
Our method	97.2	76.5	94.5	96.3

Table 2: comparison results with other methods, based on different optimization algorithms, using Chest-Xray dataset

Method	AC	DI	JA	SE
Johnatan [20]	96.97	93.56	88.07	97.54
FCN	91.16	87.03	92.90	98.31
Unet	88.87	88.90	93.11	92.84
Our method	97.78	97.33	98.38	99.26

As shown in the [Table 1](#) and [Table 2](#), our method has better results with higher evaluation metrics than the other methods that we have mentioned. We have compared our method with two Unet and FCN methods, because they have robust results in medical image segmentation and many works have done by these two methods.

Results and Discussion

We have implemented our algorithm in the google colab space and finally, we have displayed our simulation results in [Fig. 5](#) and [Fig. 6](#). Our model is robust on blurry borders and segments the skin lesions clearly. We have implemented our final results in matlab2014 by implementing the segmentation results on their corresponding original medical images. The borders of region of interests (ROI) are clear in final results and doctors can recognize the ROI totally and cure the patients.

Moreover, we have implemented two FCN and Unet methods on our dataset, but as shown in [Fig. 4](#), the borders of ROI are blurry and not clear. Our method is a robust and precise algorithm toward medical image segmentation that can implement on other medical images like brain cancer dataset, bladder, Covid19-chest-xray medical images and etc.

we have shown each medical image with their segmentation result by this method and their final result. Firstly, we have displayed results for dermoscopic medical images in [Fig. 5](#). As mentioned before, we can observe in results that our method has more robust outputs than the other methods. The borders of ROI are clear and specialists can detect them totally. For dermoscopic images, there are lots of complex Region of Interests with illumination variations that our method could overcome these problems and segmented the legions in these images.

Also, for Chest-Xray images, there are illumination variations that our algorithm could overcome these problems too.

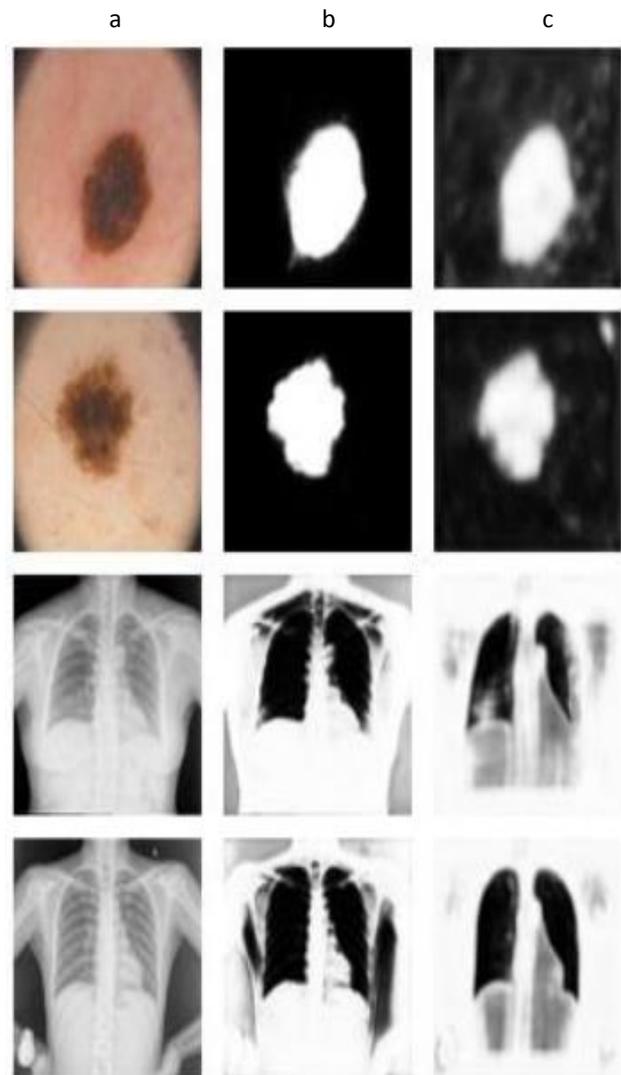


Fig. 4: Results of Unet and FCN. a) original medical images. b) Unet segmentation results. c) FCN segmentation results.

For Chest-Xray medical images, we have segmented and separated lung's texture from the other parts and also the Chest bone structures are clear in these images. By this work all of the problems including lung tumors,

lesions in lungs or Bone structure and also the infections in lungs and some diseases that happen by them like Covid19 can be detected completely. We have displayed the results for chest-Xray medical image segmentation in Fig. 6.

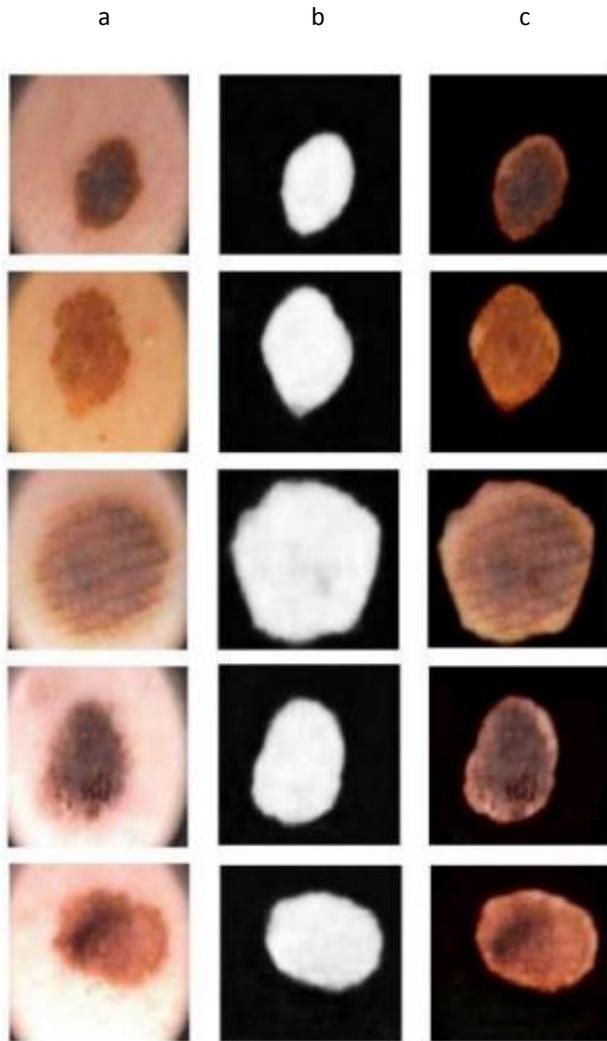


Fig. 5: Final simulation results. a) Original dermoscopic images. b) Segmentation outputs. c) Final results.

As results display, the proposed algorithm has better and more accurate results than Unet and FCN. The borders between lungs and other textures in segmented results are exactly as similar as original images. Also, in final results, we have implemented the segmentation of original images. We utilized a post-processing step in our algorithm that increases the resolution of segmented images and sharpens the borders between ROI and other regions by threshold technique. By this technique, we can compare the segmented outputs with their corresponding binary masks and evaluate the performance of our algorithm by observation too. We have shown the post-processed outputs in Fig. 7.

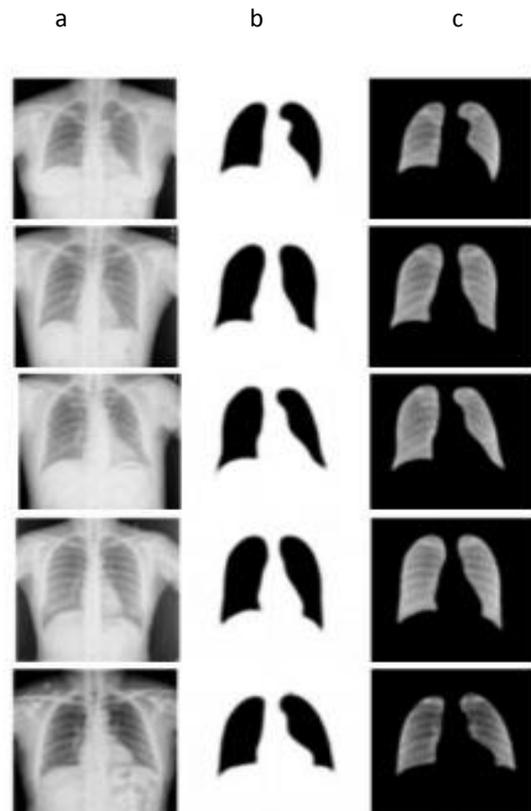


Fig. 6: Final simulation results. a) original chest-Xray images. b) segmentation outputs. c) final results.



Fig. 7: post-processed outputs. a) binary masks as ground truths. b) corresponding segmented outputs of each binary mask after post-processing.

Conclusion

In this work, we proposed an algorithm based on deep convolutional neural networks for medical image segmentation on dermoscopic and Chest-Xray medical images. Several effective training strategies were implemented to tackle the challenges that training a deep network may face when only limited training data is available.

We designed our loss function based on both Jaccard distance and binary-crossentropy. Two techniques of image augmentation, image rotation and horizontal flipping on the training dataset are performed before feeding it to the network for training. After the training process the model was evaluated on several measures for statistical values.

The predictions produced from the model on test images were postprocessed using the threshold technique to remove the blurry boundaries around the predicted lesions and make images sharper for better detection.

As shown in [Table 1](#) and [Table 2](#), Our approach could outperform the state-of-the-art methods like Unet and FCN or the algorithms that have been done in [\[16\]](#), [\[20\]](#) when evaluating on an open challenge dataset of Skin Lesion Analysis Towards Melanoma Detection in PH2 and Chest-Xray medical images. We have displayed our final results in [Fig. 5](#), [Fig. 6](#) and it has demonstrated clearly that the proposed method is robust to various image artifacts and imaging acquisition conditions while using minimum pre- and post-processings. We believe this method can generalize well to other medical image segmentation tasks.

We hope future works will develop our method for dice coefficient improvement and implement it on other famous datasets, like: ISIC2017 [\[21\]](#) and ISBI2016 [\[22\]](#) and also the other medical datasets for different parts of the body.

Author contributions

M.Taheri and M.Rastgarpour conceived of the presented idea. M.Taheri developed the theory and performed the computations. M.Rastgarpour and A.Koochari verified the analytical methods. M.Rastgarpour encouraged M.Taheri to investigate and supervised the findings of this work. All authors discussed the results and contributed to the final manuscript.

Acknowledgment

The authors are thankful to the dermatology service of Hospital Pedro Hispano in Portugal, who makes the 11 dermoscopic image data available. They also thank the anonymous reviewers whose comments and suggestions helped improve this manuscript. Also, we are thankful to

Kaggle website that has provided so many different datasets that are available for researches and programmers.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

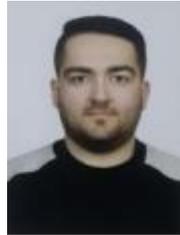
<i>CNN</i>	Convolutional Neural Network
<i>SGD</i>	Stochastic Gradient Descent
<i>FCN</i>	Fully Convolutional Neural Network
<i>ROI</i>	Region Of Interest
<i>JA</i>	Jaccard Index
<i>DI</i>	Dice Coefficient
<i>SE</i>	Sensitivity
<i>AC</i>	Accuracy
<i>IOU</i>	Intersection Over Union
<i>TP</i>	True Positive
<i>TN</i>	True Negative
<i>FP</i>	False Positive
<i>FN</i>	False Negative

References

- [1] J.U. Kim, H.G. Kim, Y.M. Ro, "Iterative deep convolutional encoder-decoder network for medical image segmentation," presented at the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Seogwipo, South Korea, 2017.
- [2] Y. Chang et al., "Automatic segmentation and cardiopathy classification in cardiac MRI images based on deep neural networks," presented at the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 2018.
- [3] A. Işın, C. Direkoğlu, M. Şah, "Review of MRI-based brain tumor image segmentation using deep learning methods," *Procedia Comput. Sci.*, 102: 317-324, 2016.
- [4] P. Moeskops et al., "Deep learning for multi-task medical image segmentation in multiple modalities," in *Proc. International*

- Conference on Medical Image Computing and Computer-Assisted Intervention,: 468-486, 2016.
- [5] D. Abdelhafiz et al., "Convolutional neural network for automated mass segmentation in mammography," presented at the 2018 IEEE 8th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS), Las Vegas, NV, USA, 2018.
- [6] H. Fu et al., "Retinal vessel segmentation via deep learning network and fully-connected conditional random fields," Presented at the 2016 IEEE 13th international symposium on biomedical imaging (ISBI), Prague, Czech Republic, 2016.
- [7] Z. Guo et al., "Deep learning-based image segmentation on multimodal medical imaging," *IEEE Trans. Radiat. Plasma Med. Sci.*, 3(2): 162-169, 2019.
- [8] M. Emre Celebi, Q. Wen, S. Hwang, H. Iyatomi, G. Schaefer, "Lesion border detection in dermoscopy images usingensembles of thresholding methods," *Skin Res. Technol.*, 19(1): e252– e258, 2013.
- [9] H. Zhou, X. Li, G. Schaefer, M.E. Celebi, P. Miller, "Meanshift based gradient vector flow for image segmentation," *Comput.Vis. Image Understand.*, 117(9): 1004–1016, 2013.
- [10] F. Xie, A.C. Bovik, "Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm," *Pattern Recognit.*, 46(3): 1012–1019, 2013.
- [11] T. Mendonça et al., "PH 2-A dermoscopic image database for research and benchmarking," presented at the 2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC), Osaka, Japan 2013.
- [12] S. Pereira, A. Pinto, V. Alves, C.A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Trans. Med. Imag.*, 35(5): 1240–1251, 2016.
- [13] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, "Gradientbased learning applied to document recognition," in *Proc. the IEEE*, 6(11): 2278–2324, 1998.
- [14] S. Ioffe, C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *Proc. the 32nd International Conference on Machine Learning, PMLR*, 37:448-456, 2015.
- [15] T. Zhang "Solving large scale linear prediction problems using stochastic gradient descent algorithms" presented at the 21th International Conference on Machine Learning, Banff, Canada, 2004.
- [16] Y. Yuan, M. Chao, Y.C. Lo, "Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance," *IEEE Trans. Med. Imaging*, 36(9): 1876-2017.
- [17] M.H. Jafari, E. Nasr-Esfahani, N. Karimi, S. Soroushmehr, S. Samavi, K. Najarian, "Extraction of skin lesions from nondermoscopic images using deep learning," *arXiv preprint arXiv: 1609.02374*, 2016.
- [18] M. Avendi, A. Kheradvar, H. Jafarkhani, "A combined deeplearning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI," *Med. Image Anal.*, 30: 108–119, 2016.
- [19] <https://www.kaggle.com/search?q=Chest+xray+segmentation>
- [20] J. Carvalho Souza, J.O.B. Diniz, J.L. Ferreira, G.L. Franc, a da Silva, A. Corr'ea Silva, A. Cardoso de Paiva, "An automatic method for lung segmentation and reconstruction in chest X-Ray using deep neural networks," *Comput. Methods Programs Biomed.*, 177: 285-296, 2019.
- [21] D. Gutman, N.C.F. Codella, E. Celebi, B. Helba, M. Marchetti, N. Mishra, A. Halpern, "skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC)," presented at the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI), Washington, DC, USA, 2018.
- [22] D. Gutman, N.C.F. Codella, E. Celebi, B. Helba, M. Marchetti, N. Mishra, A. Halpern, "Skin lesion analysis toward melanoma detection: A challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC)," May 2016.

Biographies



Mohammad Taheri was born in 1994 in the Damghan city of Semnan province, Iran. He received his B.Sc. degree in information technology engineering from Semnan university, Semnan, Iran, M.Sc degree in Artificial intelligence from Islamic Azad university, Science & Research branch, Tehran, Iran in 2018 and 2020, respectively. His major interests include image processing, computer vision, medical image segmentation, deep neural networks, convolutional neural networks and machine learning.



Maryam Rastgarpour received her B.Sc. in software engineering from Kharazmi university, Tehran, Iran, in 2003, M.Sc. and Phd degrees in computer engineering: AI from Islamic Azad University (IAU), Science & Research branch, Tehran, Iran, in 2014. She is currently an assistant professor in the department of Computer engineering at the Islamic Azad University, Saveh branch, Saveh, Iran. Her research interests include in the areas of Machine Vision, Optimization, Data Mining and expert systems specifically in medical image segmentation. She has supervised or advised over 60 MS theses or PhD dissertations and contributed or coauthored more than 60 papers published in journals or conference records or as book chapters.



Abbas Koochari received his B.Sc. in Computer Engineering from Shahed University, Tehran, Iran, M.Sc. degree in Computer Engineering: AI from the Amirkabir University of technology, Tehran, Iran and Ph.D. degree in Computer Engineering: AI from the Iran University of Science and technology, Tehran, Iran. He is currently an assistant professor in the department of computer engineering at the Islamic Azad University, Science & Research Branch, Tehran, Iran. His research interests include Signal Processing, Speech Recognition, Image and Video Processing.

Copyrights

©2020 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

M. Taheri, M. Rastgarpour, A. Koochari, "A novel method for medical image segmentation based on convolutional neural networks with SGD optimization," *J. Electr. Comput. Eng. Innovations*, 9(1): 37-46, 2021.

DOI: [10.22061/JECEI.2020.7404.390](https://doi.org/10.22061/JECEI.2020.7404.390)

URL: http://jecei.sru.ac.ir/article_1482.html





Research paper

Deep Neural Network with Extracted Features for Social Group Detection

A. Akbari, H. Farsi* , S. Mohamadzadeh

Department of Communication Engineering, Faculty of Electrical and Computer Engineering, University of Birjand, Birjand, Iran.

Article Info

Article History:

Received 24 May 2020
Reviewed 24 July 2020
Revised 09 September 2020
Accepted 12 November 2020

Keywords:

Social group detection
Deep neural network
Feature extraction
Video processing

*Corresponding Author's Email
Address:

h.farsi@birjand.ac.ir

Abstract

Background and Objectives: Video processing is one of the essential concerns generally regarded over the last few years. Social group detection is one of the most necessary issues in crowd. For human-like robots, detecting groups and the relationship between members in groups are important. Moving in a group, consisting of two or more people, means moving the members of the group in the same direction and speed.

Methods: Deep neural network (DNN) is applied for detecting social groups in the proposed method using the parameters including Euclidean distance, Proximity distance, Motion causality, Trajectory shape, and Heat-maps. First, features between pairs of all people in the video are extracted, and then the matrix of features is made. Next, the DNN learns social groups by the matrix of features.

Results: The goal is to detect two or more individuals in social groups. The proposed method with DNN and extracted features detect social groups. Finally, the proposed method's output is compared with different methods.

Conclusion: In latest years, the use of deep neural networks (DNNs) for learning and detecting has been increased. In this work, we used DNNs for detecting social groups with extracted features. The indexing consequences and the outputs of movies characterize the utility of DNNs with extracted features.

Introduction

The importance of video processing has increased over time and the expanded use of a camera to detect suspicious activity in a crowd and social anomalies [1]-[9]. Seeing social groups is concerned by governments for detecting dangerous situations. To identify abnormal behavior, recognizing social groups is a prerequisite. Contextual abnormal human behavior detection is presented in [10], [11]. People's interest in moving in groups and tracking in the groups are detected by multiple cameras [12].

To walk in the group means to be a subsystem in the group; in other words, a group of two or more individuals may be considered in the same direction of motion. Walking in a group is known to be moving

through crowds through the personal control of someone or other men. An individual joining a group is affected by the group, so that the group suits the person's pace and direction.

Social signal processing examines social relationships, conversational relationships, and even the position of people during the conversation. In addition, it shows the importance of identifying a social group by a robot. Here, it is essential to identify social groups to match these relations. Detection of social groups to aid the behavior of the robot in teamwork with humans was reported in [13], where linear extrapolation of inter-event intervals implemented an anticipation method. Skeletal data from participants detected social groups, and the method of

anticipating events was used to transfer robots among the human group [14], [15]. The technique described in [16], [17] detects the robot's conversation and social classes, using people's direction, and lower-body orientation. Clustering games are applied for detecting conversation groups in image and video [8], [18]. Deep matching-based pairwise potential with a conventional spatiotemporal relation-based pairwise potential was used to track many people in the video [19]. In this article, social group is detected by deep neural network (DNN).

The rest of this paper is structured as follows: A summary of the literature is discussed in 2th Section. The fundamentals of a DNN are outlined in 3th Section. The proposed method is defined in 4th Section. The experimental findings are shown in 5th Section. The premise is eventually set out in 6th Section.

Related Work

Human analysis applications in a crowded scene are divided into four categories: visual surveillance, crowd management, public space and entertainment design [20]. Visual surveillance is a system of monitoring activity in a building or area. Crowd management is important to manage crowd areas like stadiums. Public space is generally accessible to people. Human analysis is important for Public space and Entertainment design.

In this article, social group detection is classified into three categories: Support Vector Machine (SVM)-based, clustering-based, and deep neural network-based detection.

In SVM-based methods, sociological features between pairs of trajectories and supervised learning are applied for group detection. Body and head orientation features extracted from the social scene are used for detecting the social groups. An electric dipole and each person's eyesight are used for this purpose. If a connection is found between the eyes of the peoples, this group of people is put within a social group [21]. The relationship between people is identified using graph-based clustering in the method reported in [22], and the SVM based classification detects further social group activity. Spatial proximity prior, similarity properties prior, and spatiotemporal closure prior are used to detect social group in the crowd. Features are used to train SVM for human group detection [23].

Hierarchical clustering is used to classify groups within society. Tracking of salient points and adaptive clustering is used to identify hierarchical social classes. To detecting social groups, agglomerative hierarchical clusters with pair proximity and speed are applied [24].

The Density-Based clustering algorithm is used to detect the pedestrian groups. Spatiotemporal-oriented energy, slight motion energies, inter-group flow-field distance, and bottom-up hierarchical clustering are

calculated for this purpose. If the inter-group flow-field distance between two persons is less than a threshold, then two persons are considered as a social group [25]. Generating coherent filtering clusters, anchor tracklet, seeding tracklets are applied for group detection and crowd understanding [26].

In DNN Based method, data-driven Generation of Socially Acceptable Trajectories (App-LSTM) is employed for detecting small groups. The App-LSTM consists of position LSTM encoder, orientation LSTM encoder, and Group Interaction Module (GIM). A deep affinity network is used based on position and orientation to detect conversational groups. For this purpose, an interaction graph is made based on location and orientation. Then pairwise affinities are computed, and the affinity matrix is made. Finally, the conversational group is detected [27]. Generative adversarial networks and LSTM encoder are used in crowds for the estimation of trajectories and group identification [28].

In recent years, DNNs have been used in many fields of science and proved their applications [29]-[30]. The Relu and Softmax functions are the most commonly used activation functions. In the ReLU activator function, if the value of the input is greater than 0, the output is equal to the input.

On the other hand, if the value of the input is less than 0, the output is 0. The main advantage of using the ReLU activator function is that it has a constant derivative for all inputs greater than 0. Better and faster network learning is the result of this constant derivative. In [31], it is proven that supervised coaching of DNNs is a good deal quicker, if the hidden layers are composed of the ReLU.

Softmax activation functions are usually used in the output layer for classification purposes. The outputs of this function are normalized to a range of 0 to 1 [32]. With the softmax activation function, the categorization becomes very simple and the output is finally probabilistic.

The Proposed Method

The Architecture of the proposed method is shown in Fig. 1.

The DNN in the Proposed Method is designed with input layers, three intermediate layers with 200 nodes, and one output layer. The output of each input layer is the input of the next layer. The structure of the DNN is shown in Fig. 1. The DNN used in the proposed method has five inputs and one output. The number of entries is equal to the number of features between pairs. If the output is equal to 1, two persons are in the same social group, and if the output is 0, two persons are not in the same social group. DNN training is performed on training data. In the training phase, weights and bias of all layers of the DNN are computed.

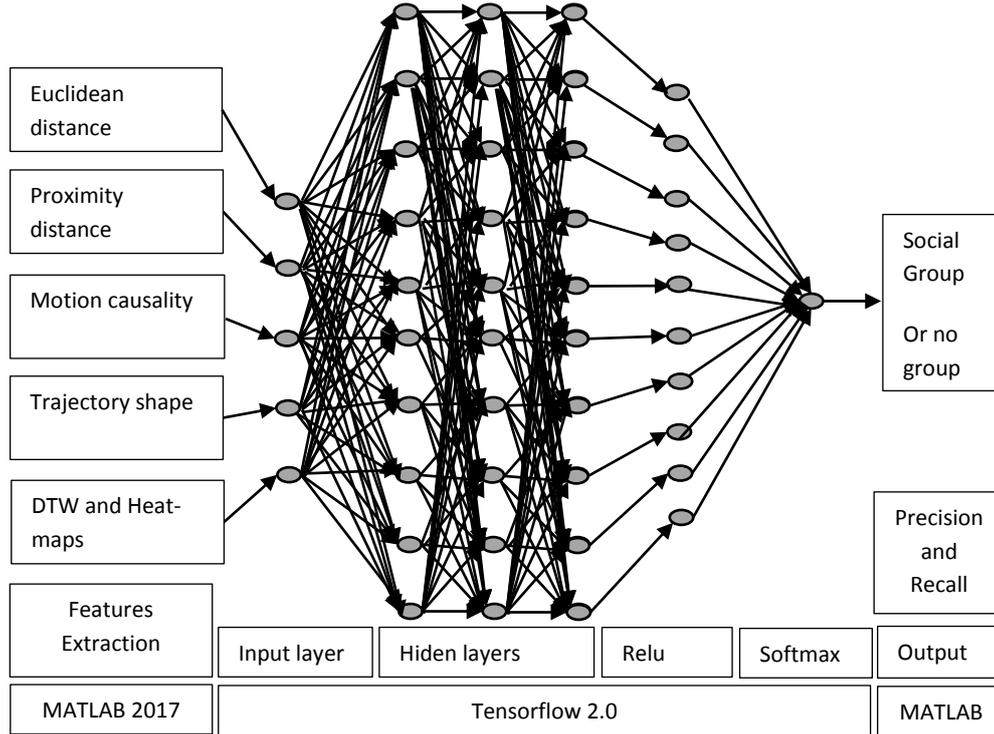


Fig. 1: The Architecture of the proposed method.

Training and experimentation in identifying social groups require extracting attributes between two or more persons. In the following, the features used in the proposed method are defined. In this method, a row of feature matrices are defined for persons 'a' and 'b' as:

$$d = [a, b, d_{eu}, d_{ph}, d_{sh}, d_{ca}, d_{he}, i, g]_{a,b} \quad (1)$$

The feature matrix consists of the label number of individuals, five features between two persons, the number of videos in the dataset, and the parameter represent two persons 'a' and 'b' in one group. Here, d_{eu} is Euclidean distance, d_{ph} is the Proximity distance, d_{sh} is motion causality, d_{ca} is trajectory shape, and d_{he} is the heat map between persons 'a' and 'b'. In (1), the number of video database is 'i'. The parameter 'g' represents whether persons 'a' and 'b' are in one group. If 'g' is equal to one, persons 'a' and 'b' are in the same group, and if 'g' is equal to 0, it means that two persons are not in the same social group.

All five features are calculated between the two presented people in the video. In the matrix, the first and second columns of each row are the label number of persons, and the third column to the seventh column shows the features.

There are a large number of people in the video, and finding five features between them cause high computational costs.

For example, in the feature matrix in the first movie of the dataset, 'i', the number of videos is one, and the number of people in the first movie is 48.

Then, 1128 rows of feature matrix represent the features between them. In the second movie of the dataset, 'i' is equal to two, the number of people in the second movie is 49, and then the next 1,176 rows of feature matrix represent the features between them. In the following, the features used in the proposed method are introduced.

Euclidean distance between all pairs is computed for each frame of videos. The mean of Euclidean distance in all frames is defined as the Euclidean distance feature for each pair.

Investigating the proximity distance of individuals in a group is one of the most critical issues of Social Group Representation. Figure 2 illustrates the importance of a Proximity distance feature. Proximity theory describes the theory of proximity based on the physical distance property [21].

According to proximity theory, there are four classes for separating the type of relationship between individuals based on the distance of two individuals. Each class is separated by the boundary shown in Fig. 2. Intimate class is lower than 0.5 meter, between 0.5 and 1.2 meter is personal class, between 1.2 and 3.7 meter is social class and between 3.7 and 7.6 meter is public class.

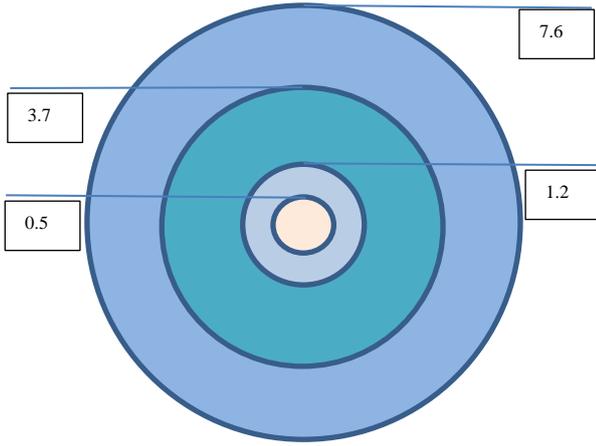


Fig. 2: Proximity theory and the importance of physical distance in identifying social groups.

The proposed method uses a Gaussian mixture model inspired by the proximity theory. The Gaussian mixture model [23] is expressed by (2):

$$GMM(P_a^t - P_b^t) = \frac{1}{4} \sum_{i=1}^4 N(P_a^t - P_b^t | 0, \sigma_i) \quad (2)$$

The Gaussian mixture model between two persons ‘a’ and ‘b’ is obtained at moment t, where P_a^t is the position of the person ‘a’ at moment t. In this model, N is a simple Gaussian model with zero mean. The variances are the boundaries of the proximity theory, namely 0.5, 1.2, 3.7, and 7.6. Consequently, between two persons ‘a’ and ‘b’, at moment t, the four variances are calculated, and the final output is the mean of four Gaussians with different variances. Next, for the entire consecutive frame in the video between the two people, it is necessary to calculate the average Gaussian mixture model based on the next relation [23].

$$d_{ph} = \frac{1}{T} \sum_{t \in T} GMM(P_a^t - P_b^t) \quad (3)$$

Equation 3 extracts the proximity distance feature for identifying the social groups, where T is the total video time. The output of the physical distance attribute between the two people will be a numerical value.

Motion causality is extracted from the movement of people. All datasets provide people’s location in each frame. Based on this information, $\bar{V}_a(n)$ is the stored information about the path traveled from (t-n to t-1). $P(V_a(t)|\bar{V}_a(n))$ is the trajectory predictor of a person where is calculated based on the ‘n’ previous frames. Figure 3 illustrates the similarities of the movement of individuals in successive frames.

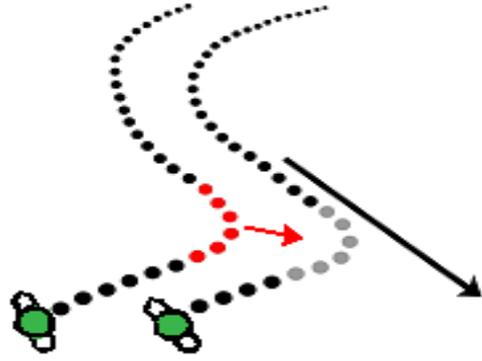


Fig. 3: Motion causality feature.

This problem is formulated as follows [23]:

$$RSS_c = \sum_{t=1}^T (V_a(t) - P(V_a(t)|\bar{V}_a(n)))^2 \quad (4)$$

where $\bar{V}_a(n)$ is the stored information about the path traveled by the person ‘a’ from (t-n to t-1). Trajectory predictor of the person ‘a’ with the stored information of two persons ($\bar{V}_a(n)$ for the person ‘a’ and $\bar{V}_b(n)$ for the person ‘b’) is $P(V_a(t)|\bar{V}_a(n), \bar{V}_b(n))$. This trajectory predictor with information of two persons is formulated as follows [23]:

$$RSS_u = \sum_{t=1}^T (V_a(t) - P(V_a(t)|\bar{V}_a(n), \bar{V}_b(n)))^2 \quad (5)$$

Where trajectory predictor is computed with information of one person in the (5) and trajectory predictor is computed with information of two-person in the (6). In (6), the similarity of these trajectory predictors is computed [23].

$$S_{b \rightarrow a} = \frac{(RSS_c - RSS_u)/n}{RSS_u/(T - 2n - 1)} \quad (6)$$

In the (6), T is the time of each video. In the (7), Fisher-Snedecor probability function is used to extract the motion causality feature [23].

$$d_{ca}(a, b) = \max_{S \in \{S_{b \rightarrow a}, S_{a \rightarrow b}\}} \int_0^S F(x|n, K - 2n - 1) dx \quad (7)$$

In the (7), the similarity characteristic of the changes of movement of individuals ‘a’ and ‘b’ in successive frames is computed.

Trajectory shape is extracted from the Dynamic Time Warping algorithm between two-time sequences. Figure 4 presents the importance of resembling the shape of the path taken by individuals in identifying social groups.

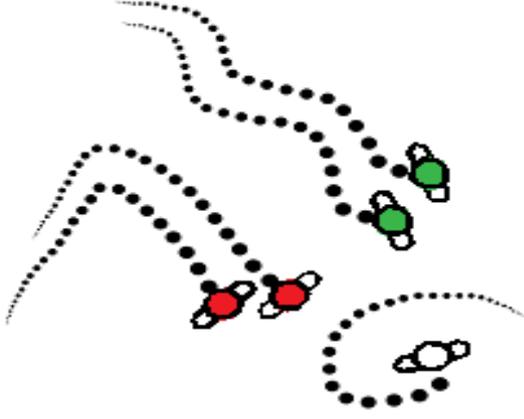


Fig. 4: Trajectory shape feature.

To compute the cumulative cost function $\gamma_{ab}(i, j)$, it is necessary to calculate the Euclidean distance matrix of the path of the person 'a' at the first moment with the path of the person 'b' at the first moment (D_{ab}^{11}). Here, D_{ab}^{ij} is the Euclidean distance matrix of the path of the person 'a' at the moment i , with the path of the person 'b' at moment j .

Dimensions of the Euclidean distance matrix are f and g , which represent the path traveled by the person 'a' and 'b', respectively.

The cumulative cost function is defined as follows [23]:

$$\gamma_{ab}(i, j) = D_{ab}^{ij} + \min\{\gamma_{ab}(i-1, j) + \gamma_{ab}(i-1, j-1) + \gamma_{ab}(i, j-1)\} \quad (8)$$

To calculate the cumulative cost function $\gamma_{ab}(i, j)$, it is necessary to calculate D_{ab}^{11} , which is equal to $\gamma_{ab}(1, 1)$, based on (8).

Then, $\gamma_{ab}(i, j)$ is obtained based on the location of two individuals in the time sequence. Finally, the similarity property of the shape of the path followed by individuals is obtained from (9).

$$d_{sh}(a, b) = \gamma_{ab}(f, g) / \max(f, g) \quad (9)$$

Since the two numbers of f and g can be different, trajectory shape is divided into the maximum of these two numbers.

Heat map parameter for person 'a' is obtained using (10) based on the average heat map in a small rectangular plot of p and q .

The rectangular segments R and C are derived from the path information obtained by individuals 'a' and 'b', which are equal to the rectangle that two people traveled in and out of the time window [23].

$$H_a(i, j) = \sum_{p=1}^R \sum_{q=1}^C \bar{E}(p, q) e^{-\tau \|(p-i, q-j)\|} \quad (10)$$

where τ in (10) is the time window of the video. Then, the features of the DTW and Heat-maps between individuals 'a' and 'b' are estimated by [23]:

$$d_{he}(a, b) = \sum_{i=1}^R \sum_{j=1}^C H_a(i, j) H_b(i, j) \quad (11)$$

The output of (11) is the similarity feature of the DTW and heat map between individuals 'a' and 'b', indicating the sharing of heat maps of individuals 'a' and 'b' in the rectangular segment R and C .

Evaluation

MATLAB 2017 software was used to extract the feature and TensorFlow 2.0 was used for training and testing of deep learning. The output of the TensorFlow is a matrix, and the first and second columns of this matrix are people's number.

If the third column of this matrix is one, two people in the corresponding row are in the same group, and if the third column is zero, the two people in the relevant row are not in the same social group. MATLAB 2017 has been used to identify groups more than two person and to display the output on video.

If the person with number of 40 is in the same group with the person with number of 43, the person with number of 43 and the person with number of 44 are in the same group, and the persons with numbers of 40 and 44 are in the same group, a three-person group consists of people with numbers 40, 43 and 44 are made.

Due to a large number of people in the video as well as people crossing different paths, it is difficult to identify groups. Crossing groups together also complicates the problem. Student003 [8], GVEII [22], ETH [28], Hotel [28], and MPT 20x100 [23] datasets were used to evaluate the efficiency of the proposed method. Table 1 lists the features of these five databases.

In Table 1, the variable of v is the number of videos, p is the number of participants, and g is the number of groups. The variables of $d1$ and $d2$ are the minimum and the maximum distances to a person (in meters), respectively. These datasets include people's route data and the number of people. The third column of the table shows the number of people in the five databases. The average distance between groups in the five datasets varies, indicating the need for training.

Table 1: Comparison of five datasets

	v	P	G	d1	d2
ETH	1	117	18	0.99	2.79
Hotel	1	107	11	0.75	2.0
Student 003	20	406	108	0.41	0.70
GVEII	30	630	207	0.77	1.66
MPT 20x100	20	82	10	0.63	1.45

The results are compared using the precision and recall parameters.

The precision is the ratio of the quantity of organizations successfully recognized to the variety of all the agencies recognized.

The recall is the ratio of the variety of agencies efficaciously listed in the database to the number of faulty companies. The standard F-score, F1, is set out as follows:

$$F1 = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (12)$$

Because ETH, Hotel, student003 (CBE), GVEII, and MPT-20x100 datasets have unique properties, actual situations are evaluated with them. We current the effects output from the proposed technique in this section, and look at them with distinct methods. The results of the proposed method in each database are compared to existing methods.

In [23], SVM based classification was used. Hierarchical clustering was reported in [24]. In [26], generating coherent filtering clusters was used. In [28], Generative adversarial networks and LSTM encoder were used.

Table 2 shows the results of the proposed method in the ETH and Hotel datasets.

In Table 2, the letter 'P' indicates precision, and the letter 'R' indicates recall. ETH and Hotel datasets consist of one video and have less information than student003 (CBE), GVEII, and MPT-20x100 datasets. Deep neural network output in ETH and Hotel datasets is much weaker than existing methods.

This means that little information for the deep neural network will not lead to the desired result.

Table 3 shows the results of the proposed method in the CBE dataset.

Table 2: Compare precision and recall for the ETH and Hotel datasets

	ETH		Hotel	
	P	R	P	R
Proposed method	62.5	59.7	65.2	55.9
[23]	91.1	83.4	89.1	91.3
[24]	80.7	80.8	88.9	89.3
[26]	69.3	68.2	67.3	64.1
[28]	91.3	83.5	90.2	93.1

Table 3. Compare precision and recall for the CBE dataset

	Precision	Recall
Proposed method	81.8	80.6
[23]	82.3	74.1
[24]	72.2	65.1
[26]	10.6	76.0
[28]	82.1	63.4

As shown in Table 3, the recall of the technique has an exceptional application, and the precision of the [23] method has the best application. The execution of the proposed method in the CBE dataset is shown in Fig. 5.

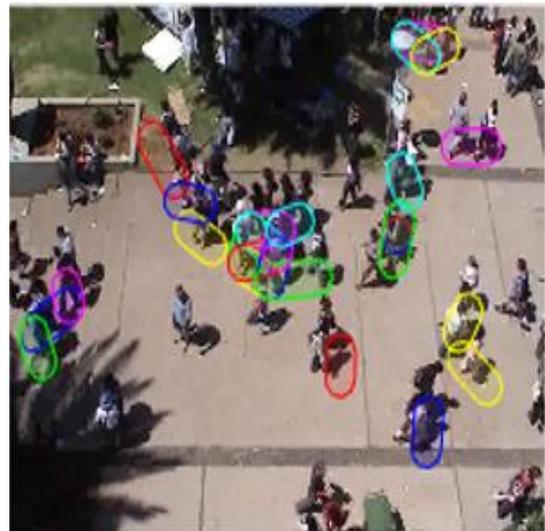


Fig. 5: The output of the social groups detected by the proposed method in the CBE dataset.

Table 4 shows the results of the proposed method in the GVEII dataset. GVEII dataset has 30 video and the most data for training and test of the deep neural network.

Table 4. Compare precision and recall for the GVEII dataset

	Precision	Recall
Proposed method	80.9	81.5
[23]	79.7	77.5
[28]	77.6	73.1

As shown in Table 4, the precision and recall of the proposed method are the best among all methods. The execution of the proposed method in the GVEII database is presented in Fig. 6.



Fig. 6: The output of the social group detection by the proposed method in the GVEII dataset.

In Fig. 7 present the results of the proposed method in the MPT-20x100 dataset.

As shown in Fig. 8, the F-score of the proposed method has the best performance in airport1, 1chinacross4, 1dawei5, 1grand1, 1japanacross2, 1manko3, 1manko29, 1shatian3, 2dawei1, 2jiansha5, 2manko2, and 2niurunning2. In Fig. 8, SCSL is [23] and VASG is [24].

The results of the proposed method in the ETH, Hotel, student003 (CBE), GVEII, and MPT-20x100 datasets showed that small datasets like the ETH, and Hotel datasets have little data for deep neural network training.

Large databases like the CBE, and GVEII databases have enough data to train deep neural networks, and the performance of this system in databases with high information is acceptable.

MPT-20x100 dataset has one video in some scenes and two different videos in some scenes. As a result, in videos containing only one video in scene, MPT-20x100 dataset does not perform well in the deep neural network.



Fig. 7: The output of social group detection by the proposed method in the MPT-20x100 dataset.

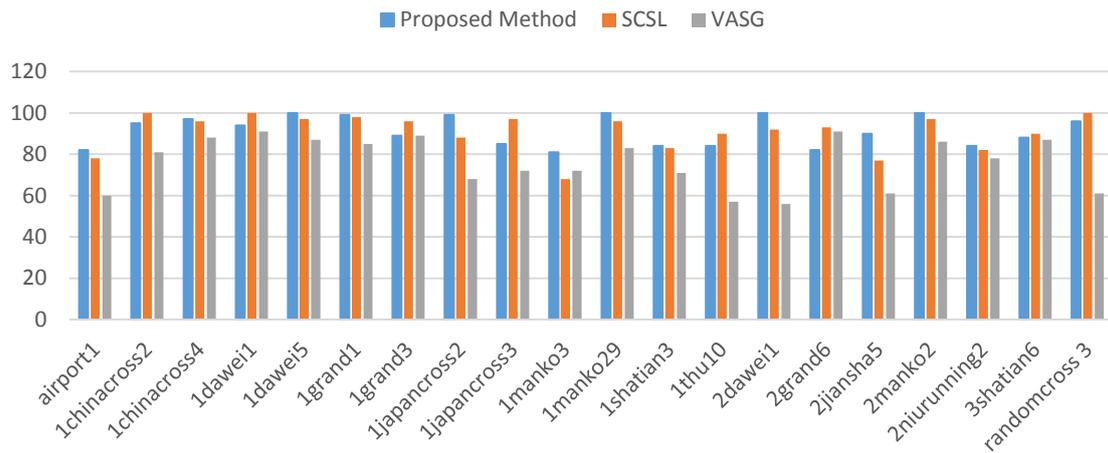


Fig. 8: Compare F-score for the MPT-20x100 dataset.

Conclusion

In this work, we used DNNs for detecting social groups with extracted features. Social group detection is one of the most necessary issues involved these days in the evaluation of interpersonal members in groups. In latest years, the use of deep neural networks (DNNs) for learning and detecting has been increased. The indexing consequences and the outputs of movies characterize the utility of DNNs with extracted features.

Author Contributions

A. Akbari, H. Farsi, and S. Mohamadzadeh designed the experiments. A. Akbari collected the data. A. Akbari carried out the data analysis. A. Akbari, H. Farsi, and S. Mohamadzadeh interpreted the results and wrote the manuscript.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Acknowledgement

The authors gratefully acknowledges the supports provided by Electrical and Computer Engineering department from University of Birjand, Birjand, Iran.

Abbreviations

<i>DNN</i>	Deep neural network
<i>GIM</i>	Group Interaction Module
<i>SVM</i>	Support Vector Machine

LSTM

Long short-term memory

References

- [1] S. Guo, Q. Bai, S. Gao, Y. Zhang, A. Li, "An analysis method of crowd abnormal behavior for video service robot," *IEEE Access*, 7: 169577 – 169585, 2019.
- [2] S.M. Hosseini, H. Farsi, H.S. Yazdi, "Best clustering around the color images," *Int. J. Comput. Electr. Eng.*, 1(1): 20-24, 2009.
- [3] P. Etezadifar, H. Farsi, "A new sample consensus based on sparse coding for improved matching of SIFT features on remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, 58(8): 5254-5263, 2020.
- [4] A. Jalali, H. Farsi, "A new steganography algorithm based on video sparse representation," *Multimed. Tool. Appl.*, 79(3): 1821-1846, 2020.
- [5] R. Nasiripour, H. Farsi, S. Mohamadzadeh, "Visual saliency object detection using sparse learning," *IET Image Proc.*, 13(13): 2436-2447, 2019.
- [6] M. Hasheminejad, H. Farsi, "Sample-specific late classifier fusion for speaker verification," *Multimed. Tool. Appl.*, 77(12): 15273-15289, 2018.
- [7] H. Farsi, R. Nasiripour, S. Mohammadzadeh, "Eye gaze detection based on learning automata by using SURF descriptor," *J. Inf. Syst. Telecommun.*, 6(1): 41-49, 2018.
- [8] M. Hasheminejad, H. Farsi, "Frame level sparse representation classification for speaker verification," *Multimed. Tool. Appl.*, 76(20): 21211-21224, 2017.
- [9] H. Farsi, R. Nasiripour, S. Mohammadzadeh, "Improved generic object retrieval in large scale databases by SURF descriptor," *J. Inf. Syst. Telecommun.*, 5(2): 128-137, 2017.
- [10] O.P. Popoola, K. Wang, "Video-based abnormal human behavior recognition—A review," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, 42(6): 865-878, 2012.

- [11] H. Farsi, "Improvement of minimum tracking in minimum statistics noise estimation method," *Signal Process. Int. J. (SPIJ)*, 4(1): 17-22, 2010.
- [12] F. Solera, S. Calderara, E. Ristani, C. Tomasi, R. Cucchiara, "Tracking social groups within and across cameras," *IEEE Trans. Circuits Syst. Video Technol.*, 27(3): 441-453, 2017.
- [13] T. Iqbal, M. Moosaei, L. D. Riek, "Tempo adaptation and anticipation methods for human-robot teams," in *Proc. RSS, Planning HRI: Shared Autonomy Collab. Robot. Workshop: 1-3*, 2016.
- [14] T. Iqbal, S. Rack, L.D. Riek, "Movement coordination in human-robot teams: a dynamical systems approach," *IEEE Trans. Rob.*, 32(4): 909-919, 2016.
- [15] X.-T. Truong, T.-D. Ngo, "'To approach humans?': A unified framework for approaching pose prediction and socially aware robot navigation," *IEEE Trans. Cognit. Dev. Syst.*, 10(3): 557-572, 2017.
- [16] M. Vázquez, A. Steinfeld, S. E. Hudson, "Parallel detection of conversational groups of free-standing people and tracking of their lower-body orientation," in *Proc. Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference.:* 3010-3017, 2015.
- [17] H. Farsi, M. Mozaffarian, H. Rahmani, "Improving voice activity detection used in ITU-T G. 729. B," presented at the 3rd WSEAS International Conference on Circuits, Systems, Signal and Telecommunications, Ningbo, China, 2009.
- [18] S. Vascon, M. Pelillo, "Detecting conversational groups in images using clustering games," in *Multimodal Behavior Analysis in the Wild: Elsevier*, 12: 247-267, 2019.
- [19] S. Tang, B. Andres, M. Andriluka, B. Schiele, "Multi-person tracking by multicut and deep matching," in *Proc. European Conference on Computer Vision: 100-111*, 2016.
- [20] T. Li, H. Chang, M. Wang, B. Ni, R. Hong, S. Yan, "Crowded scene analysis: A survey," *IEEE Trans. Circuits Syst. Video Technol.*, 25(3): 367-386, 2015.
- [21] H.S. Park, J. Shi, "Social saliency prediction," in *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE Conference: 4777-4785*, 2015.
- [22] K.N. Tran, A. Bedagkar-Gala, I.A. Kakadiaris, S. K. Shah, "Social cues in group formation and local interactions for collective activity analysis," in *Proc. International Conference on Computer Vision Theory and Applications (VISAPP): 539-548*, 2013.
- [23] F. Solera, S. Calderara, R. Cucchiara, "Socially constrained structural learning for groups detection in crowd," *IEEE Trans. Pattern Anal. Mach. Intell.*, 38(5): 995-1008, 2016.
- [24] W. Ge, R.T. Collins, R.B. Ruback, "Vision-based analysis of small groups in pedestrian crowds," *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(5): 1003-1016, 2012.
- [25] S. Huang, D. Huang, M.A. Khuhroa, "Social pedestrian group detection based on spatiotemporal-oriented energy for crowd video understanding," *KSII Trans. Internet Inf. Syst.*, 12(8): 3769-3789, 2018.
- [26] J. Shao, C.C. Loy, X. Wang, "Learning scene-independent group descriptors for crowd understanding," *IEEE Trans. Circuits Syst. Video Technol.*, 27(6): 1290-1303, 2017.
- [27] M. Swofford, J.C. Peruzzi, M. Vázquez, R. Martín-Martín, S. Savarese, "DANTE: deep affinity network for clustering conversational interactants," *arXiv preprint arXiv:1907.12910*, 2019.
- [28] T. Fernando, S. Denman, S. Sridharan, C. Fookes, "GD-GAN: generative adversarial networks for trajectory prediction and group detection in crowds," in *Proc. Asian Conference on Computer Vision.:* 314-330, 2018.
- [29] A. Sezavar, H. Farsi, S. Mohamadzadeh, "A modified grasshopper optimization algorithm combined with cnn for content based image retrieval," *Int. J. Eng.*, 32(7): 924-930, 2019.
- [30] A. Sezavar, H. Farsi, S. Mohamadzadeh, "Content-based image retrieval by combining convolutional neural networks and sparse representation," *Multimed. Tool. Appl.*, 78: 20895-20912, 2019.
- [31] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning," *Nature*, 521: 436-444, 2015.
- [32] B. Zhao, J. Feng, X. Wu, S. Yan, "A survey on deep learning-based fine-grained object classification and semantic segmentation," *Int. J. Autom. Comput.*, 14(2): 119-135, 2017.

Biographies



Ali Akbari received his B.S., M.S. and PhD degree in Communication Engineering from University of Birjand, Birjand, Iran, in 2020. His research interests are the fields of computer vision, image processing, video processing, machine learning and deep learning.



Hassan Farsi received the B.Sc. and M.Sc. degree from Sharif University of Technology, Tehran, Iran, in 1992 and 1995, respectively. Since 2000, he started his Ph.D. in the Center of Communications Systems Research (CCSR), University of Surrey, Guildford, UK, and received the Ph.D. degree in 2004. He is interested in speech, image and video processing on wireless communications. Now, he works as professor in communication engineering in department of Electrical and Computer engineering, university of Birjand, Birjand, Iran.



Sajad Mohamadzadeh received the B.Sc. degree in electrical engineering from Sistan & Baloochestan, University of Zahedan, Iran, in 2010. He received the M.Sc. degree in telecommunication engineering from university of Birjand, Birjand, Iran, in 2012. He received the Ph.D. degree in telecommunication engineering from university of Birjand, Birjand, Iran, in 2016. He is currently an academic staff in Department of electrical and computer engineering, university of Birjand, Birjand, Iran. His area research includes image processing and retrieval, pattern recognition, digital signal processing and sparse representation.

Copyrights

©2020 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

A. Akbari, H. Farsi, S. Mohamadzadeh, "Deep neural network with extracted features for social group detection," J. Electr. Comput. Eng. Innovations, 9(1): 47-56, 2021.

DOI: [10.22061/JECEI.2020.7451.392](https://doi.org/10.22061/JECEI.2020.7451.392)

URL: http://jecei.sru.ac.ir/article_1483.html





Research paper

Fault Detection in Thermoelectric Energy Harvesting of Human Body

H. Yektamoghadam, A. Nikoofard*

System and Control Department, Faculty of Electrical Engineering, K.N. Toosi University of Technology, Tehran, Iran.

Article Info

Article History:

Received 07 May 2020
Reviewed 03 July 2020
Revised 13 September 2020
Accepted 15 November 2020

Keywords:

Wireless sensor network
Wireless body area network
Thermal electric generator
Principle component analysis
Linear discriminant analysis
Neural network

*Corresponding Author's Email
Address:
A.Nikoofard@kntu.ac.ir

Abstract

Background and Objective: Human health is an issue that always been a priority for scientists, doctors, medical engineers, and others. A Wireless Body Area Network (WBAN) connects independent nodes (e.g. sensors and actuators) that are situated in the clothes, on the body, or under the skin of a person. In the 21st century, advent the technology in different aspects of human life caused WBAN has a special value in future medical technology. Energy harvesting from the ambient or human body for self-independent from the battery or power supply is an important issue in WBAN. Photovoltaic energy harvesting (PVEH), piezoelectric energy harvesting (PEH), RF energy harvesting (RFEH), and thermal electric energy harvesting (TEH) are some techniques used for energy harvesting in WBAN. Fault detection and diagnosis is an important problem in engineering. Engineers and researchers are always trying to find better ways to identify, detection, and control the fault in different systems.

Methods: We consider a thermal electric generator (TEG) for measurement energy harvested from the human body and power generation on people at different ambient conditions. Also, we used data reduction methods including principle component analysis (PCA), linear discriminant analysis (LDA), and neural network methods including PCA and MLP, LDA and MLP, Dynamic PCA and MLP, Dynamic LDA and MLP to fault detection for thermal electric generator (TEG).

Results: This study shows different data reduction algorithm, in the case studied in this paper, can detect well and nonlinear methods have a more accurate answer than linear methods but implementing the linear methods are easier.

Conclusion: According to simulation results, all the methods discussed in this paper are acceptable for fault detection. In this paper, we introduce data reduction linear and nonlinear algorithm as new methods for fault detection in WBAN.

Introduction

Fault detection and diagnosis means is a deviation from the acceptable range of the variable or parameter to be measured. In other words, a fault is one of the physical characteristics of the system compared to normal conditions. When safety and reliability are the main parameters of system quality, fault detection and

diagnosis are a way to improve these indicators [2], [15], [22], [23].

Manage and predicting the fault in different systems is a crucial target in fault detection and diagnosis. Identifying the fault location cause to improve system safety and reliability [4]. Also, it can reduce system maintenance costs.

An accurate and fast fault identification method can prevent the fault from spreading in the system and its consequences. Using fault detection methods causes the performance of the monitored system and any faults in the system are notified to the control unit before it leads to irreparable accidents [6]-[9].

Wireless sensor networks (WSN) have important applications in the scientific, medical, commercial, and military fields. Its application in medicine is also significant. Transferring medical care from a hospital setting to a home environment is an opportunity for patients to make optimal use of hospital resources, detect medical symptoms earlier, and reduce care costs ultimately. Wireless body area network (WBAN) devices have a wide range of applications and most of them use batteries to supply power [16], [17]. Researchers and engineers are trying to use ambient energy and human body energy to eliminate the batteries [1], [3], [5].

A. Related Works

The recent years, there has been a literature review that has attempted to highlight the key finding regarding fault detection issues of WBAN. Among these, some well-cited papers [2], [6], [8], [15] fault detection on wireless sensor networks(WSN) and wireless body area networks (WBAN). Data reduction methods are based on classification [3], [4], [9]. The fault detection in WSN and WBAN is a challenging problem due to sensor resources limitation and the variety of deployment fields. Energy harvesting is a process of extracting various types of energy from the environment or the human body through various methods. Lack of need to replace worn-out batteries, reducing maintenance costs due to checking and replacing batteries for wireless networks and the use of medical devices that can be installed inside the body are important reasons for using energy harvesters. So, the development of low-consumption wearable medical devices provides the necessary energy for them from the body itself and the replacement of this method instead of using batteries [11]- [13].

In this paper, we consider a thermal electric generator and use energy harvesting data in four cases, including 1) Normal data, 2) Thermal harvester module fault data, 3) thermal electric generator (TEG) fault data and 4) DC-DC conversion fault data. Then, we use data reduction methods to separate the fault from the normal data, [11], [17], [22].

Most machine learning and data mining techniques may not be effective for high-dimensional data in fault detection. So, we use a feature reduction algorithm to improve performance. In this paper, we have some supervised data and expect the linear discriminant analysis (LDA) [10] to be a better result than other methods for supervised data. So, LDA is more suitable for detection. A multilayer perceptron (MLP) [14] is a

class of feed-forward artificial neural network (ANN). We use PCA-MLP, LDA-MLP, Dynamic PCA-MLP, and Dynamic LDA-MLP as other methods [18]-[21].

TEG is a new technology and the fabricated samples did not have the desire performance. So, fault detection results not very acceptable. Fault detection methods may have different results for different devices. We use the linear and nonlinear algorithms and compare them to choose the best solution for this type of device. The purpose of this paper is to use some new solutions and algorithms based on feature extraction and neural networks for fault detection in TEG.

B. Contribution

This review paper highlights the overall fault detection of WBAN. In more concrete points, the contributions of this study are classification, data reduction, and wireless sensor networks.

C. Structure of the Paper

In this paper, we proposed new methods for fault detection in WBAN using data reduction algorithms. Moreover, we considered a TEG as a case study and show simulation results with a different algorithm. Finally, we will discuss the results and compare them. Also, the ability of these methods for fault detection was examined. This paper is organized as follows: The background of WBAN, fault detection, TEH, and control strategy, introduces the case study and their properties, explain the theory of PCA and LDA, simulation, results, conclusion and comparison between methods.

Background

The number of patients affected with chronic diseases is increasing day by day. Intermittent and transient faults are the largest source of failure for body sensor networks. A detailed overview of WBAN is presented in the next sub-section.

A. Wireless Body Area Network (WBAN)

WBAN is a collection of multiple sensors attached to or in the body, which are used to receive different physical parameters, such as body temperature, blood sugar level, heart rate, pulse rate, respiratory measurement, and even the amount of calories, burnt after exercise, etc. WBAN is not only used in medical applications but is also in multi-media and gaming applications. Several sources of non-electric renewable energy exist all around us. Power from these sources can be harnessed using appropriate hardware and converted to electrical form to fulfill energy requirements, referred to as energy harvesting. In general, energy harvesting in WBAN is classified into two-source, the human body, and the ambient. Energy source from the human body consists of a biochemical type such as Glucose, Lactate, and endocochlear potential, and biomechanical type

such as blood pressure, heartbeat, breathing, and locomotion. Energy source from ambient include sun, RF, heat, motion, and other sources can be further minimized energy constraints.

B. Fault Detection

The fault is an unpermitted deviation of at least one characteristic property or variable of the system from standard behavior. Fault detection is a determination of faults present in the system and the time of detection. Fault can be also classified taking into account the time-variant behavior of a fault. Three classes include abrupt, incipient, and intermittent can be distinguished.

C. Thermoelectric Energy Harvesting (TEH)

Body energy harvesting is the primary alternative for batteries to enhance the functionality of wearable and wireless devices and has been the subject of many recent investigations. A large amount of human energy is released in the form of heat. Therefore, technologies for body heat harvesting using thermoelectric devices have been central for many investigations. The amount of generated power from thermoelectric energy harvesters depends on the size, position, type, and efficiency of the harvesters. The thermoelectric phenomenon has been known since the discoveries made by Seebeck in 1851, followed by Peltier in 1834 and Thomson in 1851. Thermoelectric materials provide reliable conversion of heat to electricity and vice versa.

D. Control Strategy

Generally, we have two main types of control for fault-tolerant, robust control, and adaptive control. Robust control is a fixed controller designed that tolerates changes of the plant dynamics. The controlled system satisfies its goals under all faulty conditions. Fault tolerance is obtained without changing the controller parameters.

It is, therefore, called passive fault tolerance. However, the theory of robust control has shown that robust controllers exist only for a restricted class of changes of the plant behavior that may be caused by faults. Further, a robust controller works suboptimally for the nominal plant because its parameters are fixed to get a trade-off between performance and robustness. Adaptive control is controller parameters adapted to changes of the plant parameters. If these changes are caused by some fault, the adaptive control may provide active fault tolerance.

However, the theory of adaptive control shows that this principle is particularly efficient only for plants that are described by linear models with slowly varying parameters. These restrictions are usually not met by systems under the influence of faults, which typically have a nonlinear behavior with sudden parameter changes.

Case Study

Thermoelectric energy harvesting of human body heat represents a promising alternative candidate for energy harvesting. This method is independent of external factors. One of the most common methods used in thermoelectric energy harvesting is a wearable thermoelectric generator (TEG) heating system. A thermoelectric generator is a device that includes two aluminum oxide ceramic headers and is surrounded by Polydimethylsiloxane (PDMS) to help insulate and reduce the amount of heat lost when transferred from the heat spreader to the TEG. TEG can be generally classified in m-TEG that have macroscopic thermolegs and are manufactured using classic fabrication technology, and μ -TEG that have a high number and density of TC and are produced with microfabrication techniques. In this paper, we consider a wristband that includes seven thermal harvester modules that their collection makes TEG. Another device we use in the wristband is DC-DC conversion. Each part thermal harvester module, thermal electric generator (TEG), and DC-DC conversion may have a problem and cause make fault for the system. Fault detection for energy harvesting from the body has some advantages such as:

- Since, energy harvesting from the body provides power supply, fault in the system cause power off the device.
- The cost of thermal harvester module, TEG, and DC-Dc conversion is high and it is not economical to replace them.
- Proper fault detection improves system performance.
- Fault detection can improve the WBAN device in the next versions.

Seven harvester modules can be attached to the wrist while maintaining high wearability and are connected electrically in series and thermally in parallel to maximize the produced open-circuit voltage. The thermal harvester is connected to a DC-DC conversion and energy storage circuit, an application circuit completes the device. All element integrates into an elastic band that wraps comfortably around the wrist. [Figure 1](#) Shows TEG arrangement on the wrist.

Theory

We considered two types of the method include linear such as PCA, LDA, and nonlinear with help of the MLP algorithm. The base of nonlinear methods is usually neural networks. The expansion of fault detection in theory in various fields cause better efficiency of methods in practice. In this section, we will discuss the theory of linear methods.

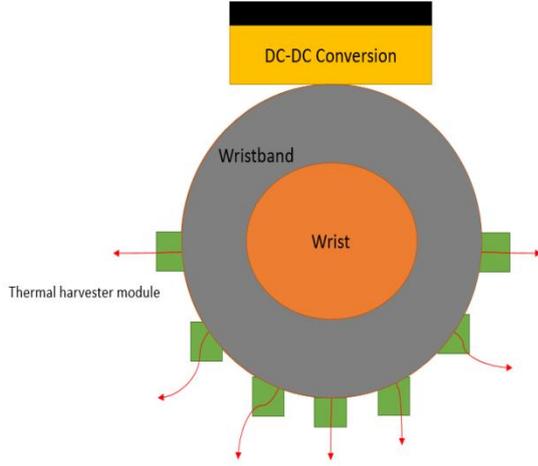


Fig. 1: TEG arrangement on the wrist.

A. Principle Component analysis theory

Principle component analysis consists of n sampled data that each of them containing m of the selected variable.

The data set can be displayed as: $X \in R^{n \times m}$. First, the mean of the data is calculated by using (1) and the variance values of the data are calculated using Equations (2), (3), and (4). The Table 1 Shows the frequency-time indicators.

We will use a table 1 Properties in PCA and LDA algorithms.

$$\bar{X}(m) = \sum_{j=1}^m \frac{\sum_{i=1}^n X(i, j)}{n} \quad (1) \quad (1)$$

$$X(n, m) = \sum_{j=1}^m \sum_{i=1}^n \{X(i, j) - \bar{X}(j)\} \quad (2) \quad (2)$$

$$X(m) = \sum_{j=1}^m \sqrt{\frac{1}{n} \sum_{i=1}^n \{X(i, j) - \bar{X}(j)\}^2} \quad (3) \quad (3)$$

$$X(n, m) = \sum_{j=1}^m \sum_{i=1}^n \{X(i, j) / \sigma_x(j)\} \quad (4) \quad (4)$$

then, the covariance matrix of the data is calculated using (5).

$$C = \frac{1}{n-1} X^T X \quad (5) \quad (5)$$

The singular values of the above covariance matrix are:

$$C = V \Lambda V^T \quad (6) \quad (6)$$

In (6), Λ is a diagonal matrix that contains the eigenvalues of the covariance matrix. V is a matrix of eigenvectors for Covariance matrix C . In this solution, by arranging the matrix of eigenvalues from larger to smaller, the main components are identified.

In fact, the eigenvalue in each row is related to the number of that component.

Table 1: frequency-time indicators

Feature	Equation
peak to peak	$ \max(X_i) - \min(X_i) $
RMS	$\sqrt{\frac{1}{N} \sum_{N=1}^N (X_i)^2}$
Kurtosis	$\frac{\frac{1}{N} \sum_{N=1}^N (X_i - \bar{X})^4}{(\frac{1}{N} \sum_{N=1}^N (X_i - \bar{X})^2)^2}$
Crest factor	$\frac{X_{peak}}{X_{RMS}}$
Skewness	$\frac{\frac{1}{N} \sum_{N=1}^N (X_i - \bar{X})^3}{(\frac{1}{N} \sum_{N=1}^N (X_i - \bar{X})^2)^{3/2}}$
Impulse factor	$\frac{X_{peak}}{\frac{1}{N} \sum_{N=1}^N X_i }$
Average power	$\int_0^{f_{max}} S_x(f) df$
Mean Frequency	$\frac{\int_0^{f_{max}} f * S_x(f) df}{P_x}$
Median frequency	$\int_0^{f_{max}} S_x(f) df = \int_{f_{med}}^{f_{max}} S_x(f) df$

B. Linear discriminant analysis theory

The objective of LDA is to perform dimensionality reduction while preserving as much of the class discriminatory information as possible. LDA tries to make the classes very crowded in the destination space and classed be scattered.

In other words, maximize the following cost function:

$$J(w) = \text{Max} \frac{W^T S_B W}{W^T S_W W} \quad (7) \quad (7)$$

S_B is a scattering matrix between classes and defined as follows:

$$S_B = \sum_{i=1}^C N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (8) \quad (8)$$

$$\mu = \frac{1}{N} \sum_{\forall x} x \quad (9) \quad (9)$$

$$\mu = \frac{1}{N_i} \sum_{\forall x \in \omega_i} x \quad (10) \quad (10)$$

S_W is the scattering matrix in the classes and defined as follows:

$$S_W = \sum_{i=1}^C S_i \tag{11}$$

$$S_i = \sum_{x \in \omega_i} (x - \mu_i)(x - \mu_i)^T \tag{12}$$

Prove the optimal W for the optimization problem posed in (7) is equal to the vector of the properties corresponding to the largest value properties of the matrix is $S_W^{-1} S_B$. In the next part, we use linear and nonlinear methods in simulation, show results, and discuss them. Fig. 2 shows the simulation methods diagram include PCA, LDA, MLP-PCA, MLP-LDA, MLP-DPCA, and MLP-DLDA.

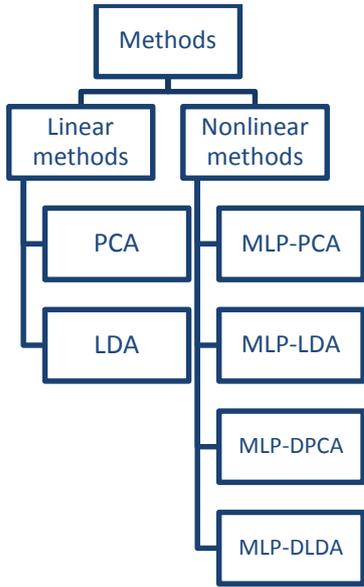


Fig. 2: Simulation methods diagram.

Simulation and Results

The purpose of using fault detection methods is advanced monitoring, fault management, improve reliability, accessibility, and optimal maintenance. Researchers use a variety of methods and algorithms for fault detection in different systems but in general, fault detection methods are divided into two main groups based on signal analysis and based on the process model. In methods based on signal analysis, one or more measurable signals in the system are analyzed using various algorithms such as PCA or LDA and any unauthorized change in signal properties indicates a fault in the system. Signal-based methods are the most common methods of fault detection. In this paper, we used signal-based methods with the linear algorithm include PCA and LDA and nonlinear algorithms include PCA-MLP, LDA-MLP, Dynamic PCA-MLP, and Dynamic LDA-MLP.

In the following, we analyze each of this algorithm:

A. Principle Component Analysis (PCA)

Principle component analysis is a simple way to extract important variables from a large set of variables in a data set. PCA method extracts a low-dimensional set of features from a high-dimensional set to analyze more information with fewer variables. In the PCA algorithm, we must first obtain the covariance matrix and then calculate the eigenvalues and eigenvectors. If the scale of the properties is very different, we use the time characteristics of the signal to synchronize. In our case study, scale properties not very different. So, we consider linear PCA without the time characteristics of the signal. The results are shown in Fig. 3 and Fig. 4.



Fig. 3: Train data after linear PCA.

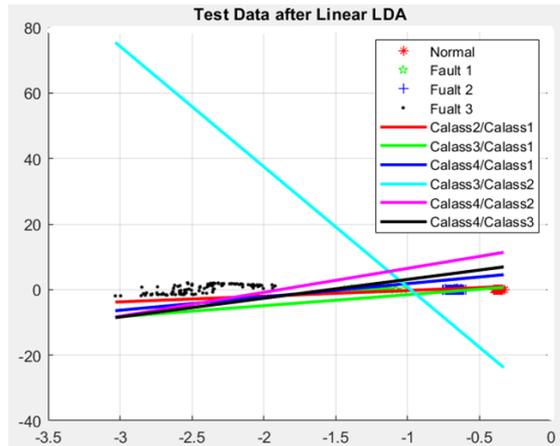


Fig. 4: Test data after linear PCA.

The results show that data have good separable and fault detection with linear PCA algorithm is a good solution for this problem.

In our case study, results show good detection and identification. Scattering between DC-DC conversion is more than the other class.

B. Linear Discriminant Analysis (LDA)

LDA is a statistical method to reduce the size of data and identify classes by maximizing the ratio of scatters between groups to scatters within groups. LDA method is similar to the method used by Ronald Fisher to determine the degree of differentiation between groups and became the basis for the analysis of variance. The objective of LDA is to perform dimensionality reduction while preserving as much of the class discriminatory information as possible. The objective of LDA is to perform dimensionality reduction while preserving as much of the class discriminatory information as possible. The solution proposed by Fisher is to maximize a function that represents the difference between the means, normalized by a measure of the within-class scatter. Table 2 shows the pseudocode of the LDA algorithm.

LDA has some limitations such as:

- LDA produces at most C-1 feature projections.
- LDA is a parametric method since it assumes unimodal Gaussian likelihoods.
- LDA will fail when the discriminatory information is not in the mean but rather in the variance of the data.

Variants of LDA include non-parametric LDA, orthonormal LDA, generalized LDA, and multilayer perceptrons. One of the disadvantages of the LDA method is that the maximum number of dimensions creates C-1, which C is the number of classes and if the number of classes is small, but the number of features is large, it will not be a good result.

Algorithm 1. LDA

Input: The Total data in N classes

Output: Classified data

1. Compute the average of each class and the average of total data from (9) and (10).
2. Normalize the data.
3. Compute between class scatter and in the class scatter by (8) and (11).
4. Compute eigenvalue and eigenvector of $S_w^{-1}S_b$.
5. Consider the eigenvectors corresponding to the eigenvalues that are bigger than others as optimal W.
6. Solve the optimization problem by a linear classifier.

Table 2. Pseudocode of LDA.

The second limitation for the LDA is a response to be optimal, the data must be Gaussian, otherwise, a good response may not be obtained or even a very bad response may be obtained. Third, if the information is not averaged but is scattered, then the LDA would be wrong. In supervised data, the LDA method is a proper

way to feature reduction and fault detection. Train data and test data after the LDA result shown in Fig. 5 and Fig. 6. The Linear LDA result shows better classification than the linear PCA. Although, both methods show good results and fault detection has been acceptable in both methods.

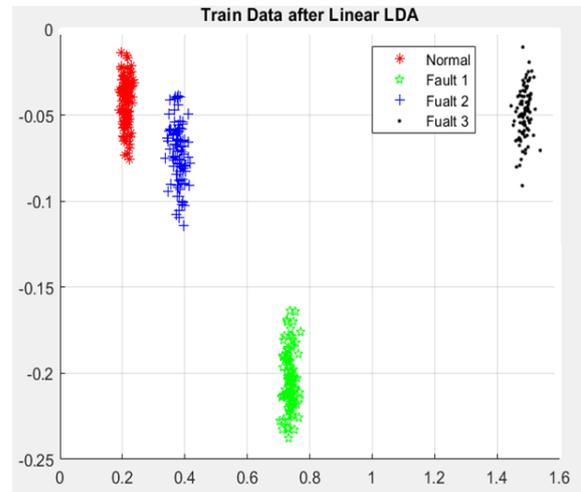


Fig. 5: Train data after linear LDA.

C. PCA and MLP

Multilayer perceptron (MLP) is a class of feedforward artificial neural networks (ANN). In an MLP, there will be at least three-node layers:

- Input Layer
- Hidden Layer
- Output layer

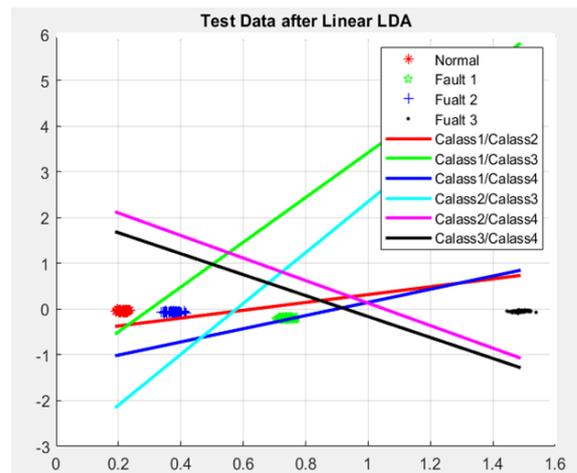


Fig. 6: Test data after linear LDA.

Neural network nodes, called neurons, are the computational units in a neural network. In the neural networks, the outputs of the first layer (input) are used as the inputs of the next layer (hidden). After a certain number of layers, the output of the last hidden layer is

used as the input of the output layer. We obtain the optimal number of neurons in the middle layer. The method of network training is the Levenberg-Marquardt algorithm. Figure 7 is shown the neural network schematic used in this problem.

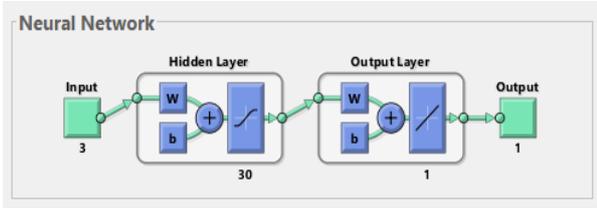


Fig. 7: Neural network schematic.

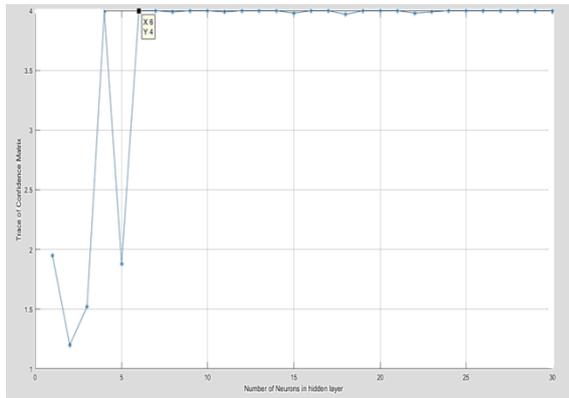


Fig. 8: Trace of confidence matrix for different number of middle layer neurons.

Table 3: Confidence matrix

	Normal	Fault1	Fault2	Fault3
Normal	1	0	0	0
Fault1	0	1	0	0
Fault2	0	0	1	0
Fault3	0	0	0	1

As shown in Fig. 8, the optimal value for middle layer neurons is 6. The classification results for 6 neurons in the middle layer are shown in Table 3. It can be seen due to the sum of the elements on the original diameter, the nonlinear MLP method gives a good answer like the linear data methods. The reason for this increase in diagnostic accuracy is the nonlinear activator function used in the neural networks.

D. LDA and MLP

In this section, we use the neural network as a classifier. Moreover, feature vectors are reduced with the LDA method that is considered as the input of the neural network. Train and find optimal neurons are like

part C. The method of network training is the Levenberg-Marquardt algorithm.

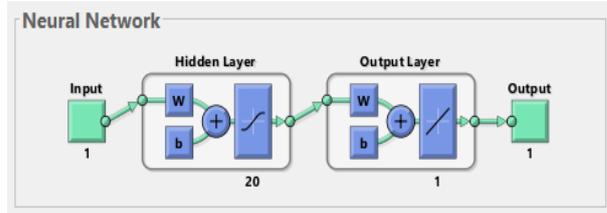


Fig. 9: Neural network schematic

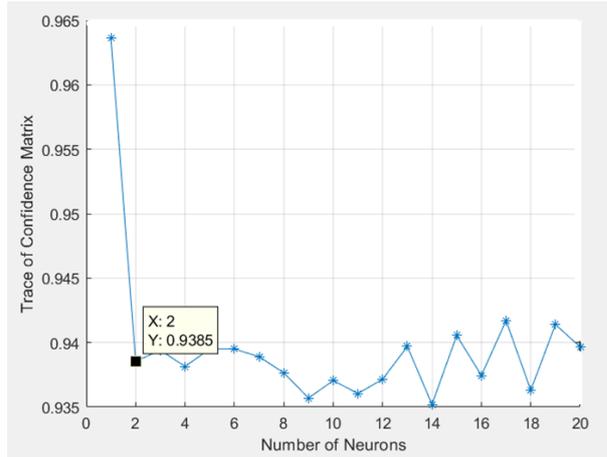


Fig. 10: Trace of confidence matrix for different number of middle layer neurons

Figure 9 is shown the neural network schematic used in this problem. According to Fig. 10, the number of optimal neurons in the middle layer is 2. The confidence matrix is shown in Table 4.

Table 4: Confidence matrix

	Normal	Fault1	Fault2	Fault3
Normal	1	0	0	0
Fault1	0	1	0	0
Fault2	0	0	1	0
Fault3	0	0	0	1

According to Fig. 10, the number of optimal neurons in the middle layer is 2. The confidence matrix is shown in Table 3.

E. Dynamic PCA (DPCA) and MLP

A novel dynamic PCA (DPCA) algorithm is proposed to extract explicitly a set of dynamic latent variables with which to capture the most dynamic variations in the data. Dynamic Principal Component Analysis (DPCA) and Artificial Neural Networks (ANN) are compared in the fault diagnosis task. Both approaches are processed

history-based methods, which do not assume any form of model structure and rely only on process historical data. Furthermore, the neural networks classifier is trained by a Multilayer perceptron (MLP).

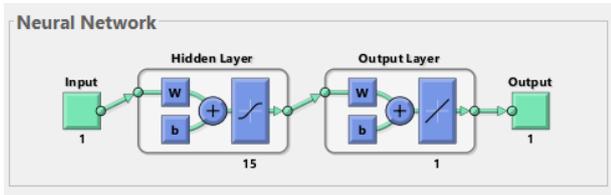


Fig. 11: Neural network schematic.

Table 5: Confidence matrix

	Normal	Fault1	Fault2	Fault3
Normal	1	0	0	0
Fault1	0	1	0	0
Fault2	0	0	1	0
Fault3	0	0	0	1

Figure 11 is shown the neural network schematic. As shown in Fig. 12, the optimal value for middle layer neurons is 2. The classification results for 2 neurons in the middle layer are shown in Table 5. The sum of the elements on the original diameter shows the nonlinear MLP method is a good answer like the linear method.

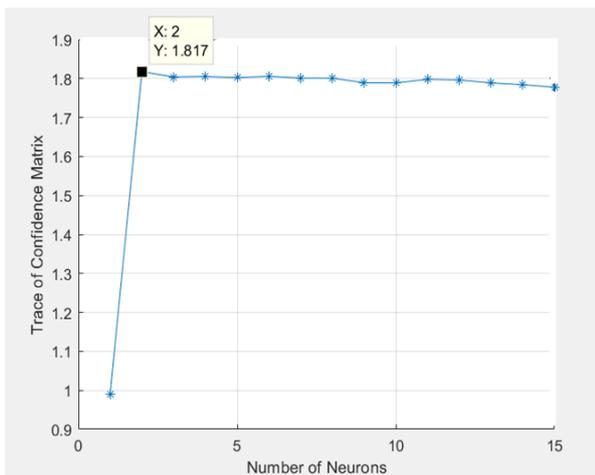


Fig. 12: Trace of confidence matrix for different number of middle layer neurons.

F. Dynamic LDA (DLDA) and MLP

Dynamic networks are networks that contain delays or integrators for continuous-time networks and that operate on a sequence of inputs. In this section before reducing the dimension, we add the delay of the feature

to findings the feature space and then apply the LDA dimension reduction algorithm to large data.

The goal of the Dynamic LDA and MLP methods is to reduce dimension and classification. So, we have three optimal value parameters; include:

- Number of middle layer neurons for the MLP neural network
- Number of eigenvectors corresponding to large eigenvalues in Dynamic LDA
- The optimal amount of delay for the Dynamic LDA method

The method of network training is the Levenberg-Marquardt algorithm. Figure 13 is shown the neural network schematic used in this problem.

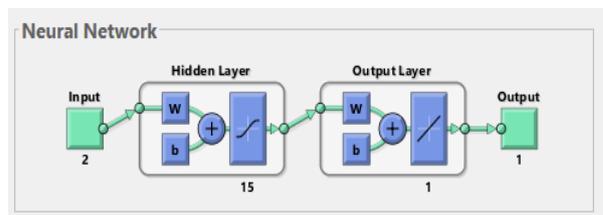


Fig. 13: Neural network schematic.

According to Fig. 14, the number of optimal neurons in the middle layer is 6. The confidence matrix is shown in Table 6. The ideal state of separation in MLP methods is to maximize the diagonal of the confidence matrix. The diagonal matrix is the best result of the separation. So, fault detection is well detected in all methods. If data are separable, use the linear and nonlinear methods to give us the desired answer. Therefore, the optimal design and fabrication can improve the efficiency of the devices, and fault detection is more easily.

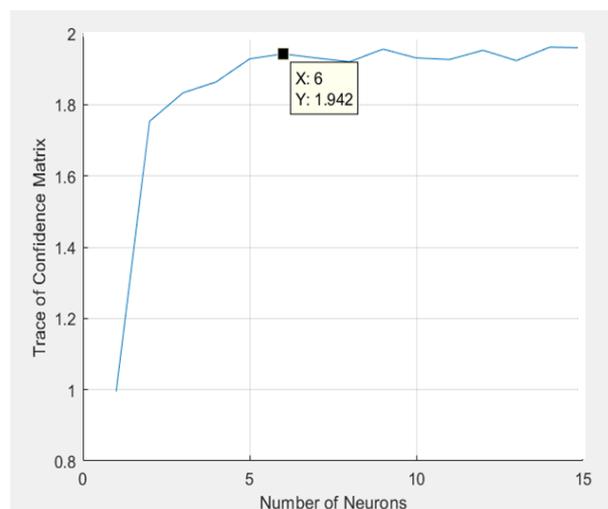


Fig. 14: Trace of confidence matrix for different number of middle layer neurons.

Table 6: Confidence matrix

	Normal	Fault1	Fault2	Fault3
Normal	1	0	0	0
Fault1	0	1	0	0
Fault2	0	0	1	0
Fault3	0	0	0	1

Conclusion

In modern industry, fault detection and isolation (FDI) is very important to enhance the system reliability, prevent serious system performance deterioration, and to ensure optimal process operation. In this paper, some classification approach has been proposed for fault detection in a case study of WBAN. Our proposed solution is based on PCA, LDA, and MLP techniques. Due to the high importance of medical technologies and their relationship with human health, rapid fault detection in these systems is vital.

We considered supervised data from a wristband in modes normal, thermal harvester module fault, TEG fault, and DC-DC conversion fault. Then analyzed this data with help of the linear methods include linear PCA and linear LDA and nonlinear methods with help of neural networks include PCA-MLP, LDA-MLP, Dynamic PCA-MLP, and Dynamic LDA-MLP algorithms. Results show fault in all methods were well detected. We expected the LDA method has the best solution because this method can detect supervised data well. Although, data were good separable and faults were well detected in all methods. In the nonlinear method is difficult to predict the best algorithm between different methods because the neural network follows a nonlinear function and forecasting nonlinear function is too difficult. Results are shown linear LDA is better detection than linear PCA but in general, the nonlinear methods show better results than linear methods. However, we choose the ideal state of confidence matrix with random parameters in the nonlinear methods. The data used were reasonably separable and this is a good reason for the accuracy of the results.

Table 7: Compare between methods

	Fault detection	Linear method	Nonlinear method	Optimal result	Fixed result	Random result	good efficiency
PCA	✓	✓			✓		✓
LDA	✓	✓			✓		✓
PCA-MLP	✓		✓	✓		✓	✓
LDA-MLP	✓		✓	✓		✓	✓
DPCA-MLP	✓		✓	✓		✓	✓
DLDA-MLP	✓		✓	✓		✓	✓

The comparison between methods is shown in the [Table 7](#).

The mentioned results can be developed and used in WBAN.

Author Contributions

Every author has equally contributed to accomplish the targeted results.

Acknowledgment

The authors wish to thank Dr. Mahdi Aliyari shooredeli, for providing the necessary assistance to pursue the research. Also, we would like to thank the anonymous referees for expressing relevant literature and helping to improve the paper.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work.

In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

References

- [1] A. Liberale, E. Dallago, A. Lazzarini Barnabei, "Energy harvesting system for wireless body sensor nodes," presented at the 2014 IEEE Biomedical Circuits and Systems Conference (BioCAS), Lausanne, Switzerland, 2014.
- [2] A. Mahapatro, P.M. Khilar, "Fault diagnosis in body sensor networks," *Int. J. Comput. Inf. Syst. Ind. Manage. Appl.*, 5: 252-259, 2012.
- [3] L. Xiao, R. Li, J. Luo, Z. Xiao, " Energy-efficient recognition of human activity in body sensor networks via compressed classification," *Int. J. Distrib. Sens. Netw.*, 12(12): 1-8, 2016.
- [4] P. Sarcevic, Z. Kincses, S. Pletl, "Online human movement classification using wrist-worn wireless sensors," *J. Ambient Intell. Hum. Comput.*, 10: 89–106, 2019.

- [5] R. Gravina, P. Alinia, H. Ghasemzadeh, G. Fortino, "Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges," *Inform. Fusion*, 35: 68-80, 2017.
- [6] S. Zidi, T. Moulahi, B. Alaya, "Fault detection in wireless sensor networks through SVM classifier," *IEEE Sens. J.*, 18(1): 340-347, 2017.
- [7] Z. He et al., "An incipient fault detection approach via detrending and denoising," *Control Eng. Pract.*, 74: 1-12, 2018.
- [8] Z. Noshad et al., "Fault detection in wireless sensor networks through the random forest classifier," *MDPI*, 19: 1-21, 2019.
- [9] B. Saneja, R. Rani. 2019. "A scalable correlation-based approach for outlier detection in wireless body sensor networks," *Int. J. Commun. Syst.*, 32(7): e3918, 2019.
- [10] L. Ciabattini et al., "A novel LDA-based approach for motor bearing fault detection," in *Proc. 2015 IEEE 13th International Conference on Industrial Informatics (INDIN)*: 771-776, 2015.
- [11] V. Leonov, "Thermoelectric energy harvesting of human body heat for wearable sensors," *IEEE Sens. J.*, 13(6): 2284-2291, 2013.
- [12] Y. Wang, Y. Shi, D. Meia, Z. Chen, "Wearable thermoelectric generator to harvest body heat for powering a miniaturized accelerometer," *Appl. Energy*, 215: 690-698, 2018.
- [13] A. Sultana et al., "A pyroelectric generator as a self-powered temperature sensor for sustainable thermal energy harvesting from waste heat and human body heat," *Appl. Energy*, 221: 299-307, 2018.
- [14] J.C. Tudón-Martínez et al., "Comparison of artificial neural networks and dynamic principal component analysis for fault diagnosis," in *Proc. International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*: 10-18, 2011.
- [15] H. Yuan, Z. Xiaoxia, Y. Liyang, "A distributed bayesian algorithm for data fault detection in wireless sensor networks," presented at the *IEEE International Conference on Information Networking (ICOIN)*, Cambodia, Cambodia, 2015.
- [16] K. Hasan, K. Biswas, K. Ahmed, N.S. Nafi, M. Saiful Islam, "A comprehensive review of wireless body area network," *J. Netw. Comput. Appl.*, 143: 178-198, 2019.
- [17] M. Thielen et al., "Human body heat for powering wearable devices: From thermal energy to application," *Energy Convers. Manage.*, 131: 44-54, 2016.
- [18] C. Wu et al., "A signal-based fault detection and tolerance control method of current sensor for PMSM drive," *IEEE Trans. Ind. Electron.*, 65(12): 9646 – 9657, 2018.
- [19] J. Huang, X. Yan, "Dynamic process fault detection and diagnosis based on dynamic principal component analysis, dynamic independent component analysis, and Bayesian inference," *Chemometr. Intell. Lab. Syst.*, 148: 115-127, 2015.
- [20] S. Zhoua, J. SinghDhupia, "Online adaptive water management fault diagnosis of PEMFC based on orthogonal linear discriminant analysis and relevance vector machine," *Int. J. Hydrogen Energy*, 45(11): 7005-7014, 2020.
- [21] B. Raheli et al., "Uncertainty assessment of the multilayer perceptron (MLP) neural network model with implementation of the novel hybrid MLP-FFA method for prediction of biochemical oxygen demand and dissolved oxygen: a case study of langat river," *Environ. Earth Sci.*, 76(14): 1-16, 2017.
- [22] M. Saida, G. Zaibi, M. Samet, A. Kachouri, "Improvement of energy harvested from the heat of the human body," presented at the 17th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), Sousse, Tunisia, 2016.
- [23] J. Faiz, E. Mazaheri-Tehrani, "Demagnetization modeling and fault diagnosing techniques in permanent magnet machines under stationary and nonstationary conditions: An overview," *IEEE Trans. Ind. Appl.*, 53(3): 2772 – 2785, 2017.
- [24] S. Moon, H. Jeong, H. Lee, S.W. Kim, "Detection and classification of demagnetization and interturn short faults of IPMSMs," *IEEE Trans. Ind. Electron.*, 64(12): 9433 - 9441, 2017.

Biographies



Hossein Yektamoghadam received the B.Sc degree in electrical engineering from the K.N.Toosi University of Technology, Tehran, Iran, in 2019. He is currently a graduate student at the K.N.Toosi University of Technology. His current research interests include wireless body Area network (WBAN), Fault diagnosis, optimization, game theory, and model predictive control.



Amirhossein Nikofard received the B.Sc. and M.Sc. degrees in electrical and computer engineering from the University of Tehran, Tehran, Iran, in 2008 and 2011, respectively, and the Ph.D. degree in electrical and computer engineering from the Norwegian University of Science and Technology, Trondheim, Norway, in 2016. He is currently an Assistant Professor of Electrical Engineering with the K. N. Toosi University of Technology, Tehran. His current research interests include nonlinear and adaptive estimation, optimization, adaptive control, game theory, soft computing, such as fuzzy logic, neural networks, and evolutionary algorithms, and MPC.

Copyrights

©2020 The auor(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

H. Yektamoghadam, A. Nikoofard, "Fault detection in thermoelectric energy harvesting of human body," *J. Electr. Comput. Eng. Innovations*, 9(1): 57-66, 2021.

DOI: 10.22061/JECEI.2020.7392.388

URL: http://jecei.sru.ac.ir/article_1485.html





Research paper

NodeFetch: High Performance Graph Processing Using Processing in Memory

M.A. Mosayebi, M. Dehyadegari*

Department of Computer Systems Architecture, Faculty of Computer Engineering, K. N. Toosi University of Technology, Tehran, Iran.

Article Info	Abstract
<p>Article History: Received 24 April 2020 Reviewed 24 June 2020 Revised 14 September 2020 Accepted 15 November 2020</p> <hr/> <p>Keywords: Graph processing Hybrid Memory Cube (HMC) Processing in memory Hardware Accelerator</p> <hr/> <p>*Corresponding Author's Email Address: dehyadegari@kntu.ac.ir</p>	<p>Background and Objectives: Graph processing is increasingly gaining attention during era of big data. However graph processing applications are highly memory intensive due to nature of graphs. Processing-in-memory (PIM) is an old idea which revisited recently with the advent of technology specifically the ability to manufacture 3D stacked chipsets. PIM puts forward to enrich memory units with computational capabilities to reduce the cost of data movement between processor and memory system. This approach seems to be a way of dealing with large-scale graph processing, considering recent advances in the field.</p> <p>Methods: This paper explores real-world PIM technology to improve graph processing efficiency by reducing irregular access patterns and improving temporal locality using HMC.</p> <p>We propose NodeFetch, a new method to access nodes and their neighbors while processing a graph by adding a new command to HMC system.</p> <p>Results: Results of our simulation on a set of real-world graphs point out that the proposed idea can achieve 3.3x speed up in average and 69% reduction of energy consumption over the baseline PIM architecture which is HMC.</p> <p>Conclusion: Most of the techniques in the field of processing-in-memory, hire methods to reduce movement of data between processor and memory. This paper proposes a method to reduce graph processing execution time and energy consumption by reducing cache misses while processing a graph.</p>

Introduction

Graph processing is increasingly gaining attention during era of big data. While graph algorithms are varied, but most of them have expensive memory accesses. There are an important reason which makes graph algorithms memory intensive. Graph structure is irregular and therefore access to nodes are irregular and difficult to predict which leads to poor temporal locality [1]. So Numerous techniques for large-scale graph processing have been proposed in the literature that address the memory bandwidth and data movement problems [2]-[4]. Hardware accelerators have proven successful in achieving significant speedup and energy efficiency in comparison to general purpose processors [5]-[6].

Many graph accelerators hire different techniques to decrease movement of data between memory and the host processor to achieve noticeable performance improvements [5]-[10].

Processing in memory (PIM) is an old idea which introduced few decades ago by academia. Back then it could not gain enough attention due to several issues such as complexity and fabrication technology. Recently PIM has been revisited by both industry and academia. Micron proposed Hybrid memory cube (HMC) as one solution to hire PIM in real word.

HMC is one of the most promising DRAM systems which is a true 3D stacked DRAM. HMC contains of

multiple DRAM dies on the top of a logic die. Several reaches have been made based on HMC to improve graph computation performance.

For example, HMC-MAC proposed a PIM architecture based on the hybrid memory cube (HMC) that adds a MAC operation to HMC to accelerate graph and NN applications. Enhanced Tesseract is another example which propose an idea for large scale graph processing based on HMC. In this paper, we explore real-world PIM technology to improve graph processing efficiency by reducing irregular access patterns and improving temporal locality. We are proposing NodeFetch, a new method to access node neighbours while processing a graph. Our research follows the HMC 2.1 specification [11].

Background

This section contains the background knowledge on PIM accelerators and graph processing.

A. Graph Processing

Based on nature of a graph, graph processing applications suffer from several issues such as random access patterns, poor locality and unbalanced workloads [1]. Two main issues of graph processing are:

- Neighbours of a given node might be somewhere else in memory causing cache miss.
- Accessing nodes are unpredictable and depends of the shape of input graph and graph algorithm.

In this Paper we are aiming to clearly address these issues and propose a novel architecture to properly fix them.

B. Processing in Memory

Processing in memory (PIM) is an old idea. With the advent of big data computing and also recent advances in memory technologies, such as emerging nonvolatile memories, die stacking, and high-bandwidth memory interfacing, PIM has been revisited recently by both academia and industry.

Processing-in-memory proposes to move computational components to the memory units to alleviate the high cost of data movement in big data processing.

Hybrid memory cube (HMC) proposed and manufactured by Mircron company in 2011 trying to reach high memory bandwidth using PIM idea. They put several 3D-stacked DRAM dies on a logic layer to increase available bandwidth while providing high performance near memory processing. Using trough silicon Vias (TSVs) inside a memory cube, several DRAM layers are connected to the logic layer at the bottom of the cube.

A single memory cube consists of 32 vertical slices Based on Hybrid Memory Cube Specification 2.1 [11]. Each of these slices called a vault. Each vault benefits

from 10GBps of memory bandwidth, therefore a single cube has total of 320GBps of bandwidth.

Motivation and Innovation

Graph processing suffers from random access patterns, poor locality and unbalanced workloads [1]. Therefore accessing each neighbour may lead to a cache miss. Upon a cache miss, processor should bring necessary data block from memory into each level of the cache and then use that data to continue the application process, which degrades system performance. This can happen to each and every neighbor of each node of graph without a proper prediction. Repeating cache misses makes the cache useless. Several techniques has been hired to solve this problem such as graph mapping and custom prefetchers.

These methods try to solve proposed issue indirectly and could gain noticeable performance increase in some cases. But if one solution can resolve the problem directly, huge performance increase can be achieved. One solution that may come to mind is to use GPUs and many cores such as [12], to tackles these issues. GPUs are being used to accelerate various applications through parallelism such as Neural network algorithms. As said earlier, the main challenge of graph processing is their random memory accesses and irregularity of their algorithms.

One possible way to accelerate graph processing is to hire GPUs.

GPU is a highly structured SIMT architecture and it is not suitable for graph applications [1].

The performance of graph processing on GPU is still limited by memory latency despite of many efforts spent of accelerating graph applications using them [1]. The whole concept of processing in memory (PIM) is to overcome an important issue which is memory bandwidth wall. On one hand, GPU can't be used properly to generally accelerate graph processing due to memory latency.

On the other processing in memory is a solution to overcome memory latency and memory bandwidth wall. Therefore processing in memory is selected as a baseline technique to directly solve issues tied with graph processing.

Several other researches in the field of processing in memory has been done to use internal bandwidth of memory to help with processing graphs. Some of them added logic into or near conventional memory units such as GraphR [13] and which used ReRAM or Graphicionado [15], But others used HMC to achieve their desired goal. Graph processing acceleration can be done by moving computations into the logic layer of HMC to exploit High in-memory bandwidth. Among those who used HMC as their baseline, there are two main categories.

- Hiring Network of several memory cubes with a specific topology [14]-[16]. Tesseract [17] achieved a significant performance improvement. Several researches tried to improve performance in a Tesseract-based system such as GraphH [9], GraphP [8] and Enhanced Tesseract.
- Implementing on a single cube which is extendable to be used in any network of multiple cubes. For example HMC-MAC [18] tried to add a MAC operation to a HMC device.

In-memory graph processing have to address several issues such as random access patterns, poor locality and unbalanced workloads. This paper proposes a novel PIM accelerator called NodeFetch based on a single HMC to accelerate graph processing by reducing random accesses and poor locality.

Related Work

Several recent related works and ideas in graph processing acceleration with the help of the HMC have been reviewed in this section. Various techniques have been proposed to accelerate graph processing such as Graphicionado [5], Tesseract [17], GraphH [9], GraphP [8], Enhanced Tesseract [7] and Centaur [19].

Graphicionado [5] accelerates graph processing by the use of parallelism and. They proposed a domain-specific hardware accelerator. HMC is not being used in their sub-system. They could achieve a better performance than a state-of-the-art software graph processing framework being executed on a 16-core Haswell Xeon processor. Tesseract [17] is a large-scale graph processing architecture which uses a network of modified HMCs towards graph processing acceleration. Although They could gain a remarkable performance, but the main problem is very long waiting times in processors. Apparently Tesseract spends 59% of execution time waiting for synchronization barriers [17]. Several researches tried to improve performance in a Tesseract-based system such as GraphH [9], GraphP [8] and Enhanced Tesseract [7].

GraphP considers data organization as a first-order design consideration to improve Tesseract-base system. Therefore they could provide a better performance in comparison to Tesseract by designing a hardware/software co-designed graph processing. GraphH on the other hand is a PIM architecture for graph processing on the Hybrid Memory Cube array. It integrates SRAM-based on-chip vertex buffers to eliminate local bandwidth degradation.

Enhanced Tesseract [7] targets the main problem of Tesseract which is low utilization due to synchronization barriers. They modified each HMC device in a way to manage and accelerate message queues and could reduce execution time by 40% in average.

Centaur [19] tries to divide graph processes into two parts. One part that can be processed in off-chip memory and the other part which should be processed in on-chip memory to accelerate graph processing. Processes related to each vertex can be done in an on-chip or off-chip memory based on the intensity of process related to that particular vertex.

Architecture

To solve the irregular data access pattern while processing a graph, we are proposing NodeFetch. NodeFetch consists of a hardware and software co-design.

From processor perspective, NodeFetch is a new command which is supported by memory subsystem. Processor can use this command to bring neighbors of a given node from memory into host processor cache. Therefore reducing cache miss rate while access to irregular neighbors of that node during executing a graph application.

Presented hardware is able to collect neighbors of a given node, inside memory and send them back to host processor as a response.

The process of finding a node and its neighbors inside memory and putting them together as a block, happens inside memory.

Therefore this is a case of using processing in memory.

To avoid building from ground up, Presented hardware placed inside logic layer of a HMC device by providing a new command inside HMC device. Figure 1 shows the flow of data between software and hardware in presented architecture.

1. Processor sends a NodeFetch command to memory, requesting to fetch a node and its neighbours using NodeFetch hardware inside the memory.
2. NodeFetch hardware receives the request and starts to gather requested node itself, and neighbors of the node, inside a buffer. The buffer size equals to a normal memory response.
3. After collecting node neighbors and the node itself, memory returns node data and its neighbors to processor. As of now, software running on processor knows that all neighbors are in adjacent addresses in memory. Therefore Upon each software request to access any of neighbors, neighbor data is already inside cache.

NodeFetch only finds level one neighbors which means the response does not contain neighbors of neighbors and so on. Figure 2 shows logic layer of a HMC device. NodeFetch hardware placed between crossbar switch and vault controllers. NodeFetch hardware consists of several components. Figure 3 shows block diagram of a NodeFetch unit.

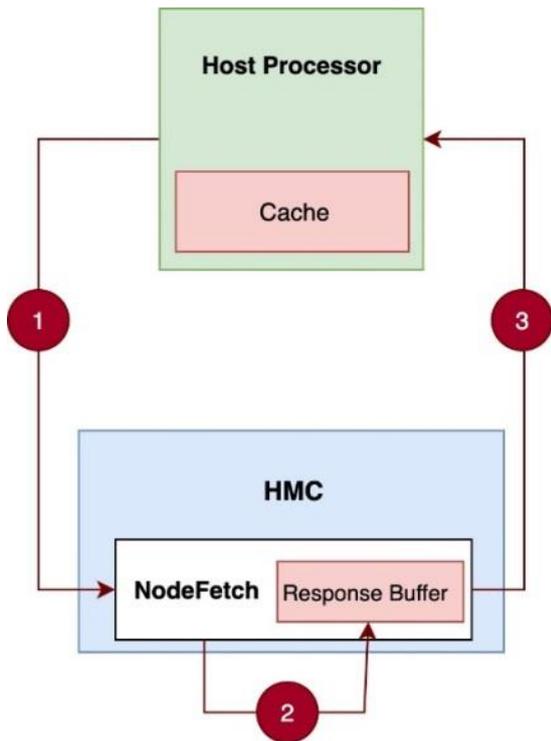


Fig. 1: Flow of data between software and hardware.

Activation Register is used to enable or disable the whole unit. If NodeFetch unit is not activated, then memory packets will pass around the unit.

Vertex Address Register keeps data address of requested node.

Offset Register holds the offset of neighbors to fetch them. Processor sends this variable with the memory request. When fetching neighbours of a node, there might not be enough space in a memory response to fit all neighbors. Therefore processor can request for the remaining neighbors of a node by properly setting the Offset Register.

Neighbor Prefetcher checks Activation register to determine whether to start the process or not. First, Finds address of node data and sets the Vertex Address Register with that address. Then fetch node neighbors and put them inside Block Buffer.

Neighbor Address Buffer keeps address of neighbors during the actual process.

Block Buffer keeps final memory response. Figure 4 shows more details of a NodeFetch unit. After activation, NP checks the request type to find out whether it's a read request or a write request.

If the request type is of type read, Finds address of given node and write it into vertex address register (VAR).

Also write given offset from request into offset register (OR).

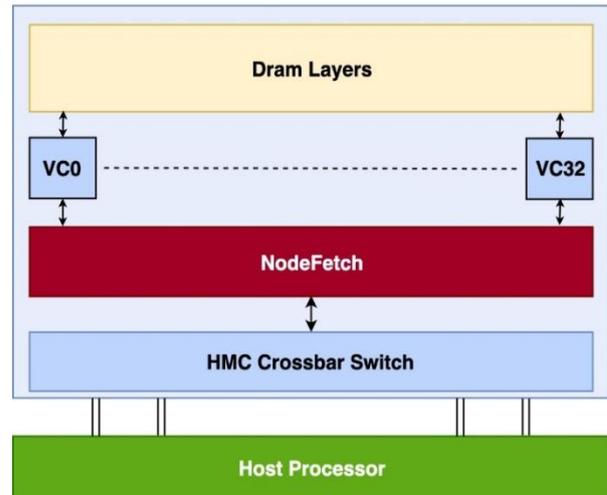


Fig. 2: HMC Logic layer including NodeFetch hardware.

After that, fetches address of neighbors from memory and writes them in neighbor address buffer (NAB). Offset Flag (OF) become activated if there isn't enough space to fit all neighbors inside NAB. Following that, iterates through NAB and fetch neighbor data into block buffer (BB).

After all a memory response emerges from BB and which goes to processor. This response contains node itself, neighbors and offset flag. On the other hand if the request is of type write, repeats previous steps, only instead of fetching neighbors data, update them inside memory.

NodeFetch is not a programmable unit and therefore only can work with a standard form graph storage inside memory. In this standard, there is an array for graph nodes.

Each key in array, refers to a node and the value for that key, contains a pointer to node data and a list of neighbors connected to that node. As a result finding address of node data and neighbors only takes up to $O(1)$.

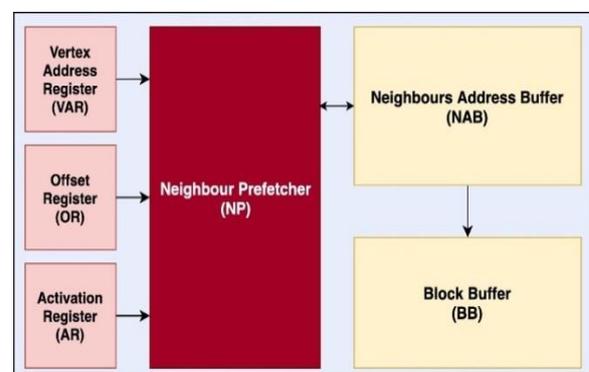


Fig. 3: NodeFetch block diagram.

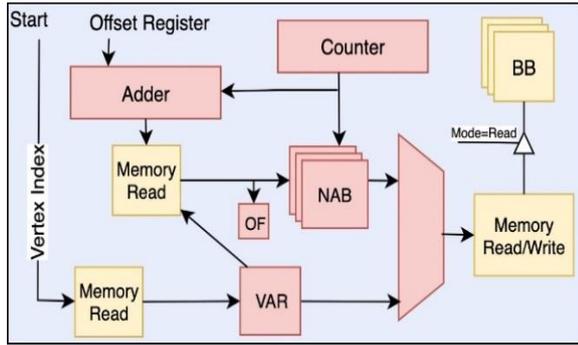


Fig. 4: Nodefetch detail.

Evaluation Methodology

A. Simulation Configuration

We evaluate the proposed system using an in-house cycle accurate simulator. Table 1 shows simulation configuration:

Table 1: Simulation configuration of Nodefetch

Memory	1 HMC module with 8GB of memory capacity and a NodeFetch unit
Processor	2 cores of 2GHz ARM Cortex A15 with 64KB of level 1 cache for data and instruction

Figure 5 illustrates block diagram of the simulator. Input graph and graph algorithm are inputs of the simulator. Dispatcher stores input graph of HMC memory.

Programmer puts graph algorithm in processor. Processor starts working on input graph based on given algorithm. After all, generates several reports such as timings and power consumption.

This simulator developed in a way to use the help of modified HMCsim.

B. Workloads

For evaluation, we implement five different graph algorithms, namely Page Rank (PR), Single-Source Shortest Paths (SSSP), Connected Components (CC), Triangle Counting (TC) and Betweenness Centrality (BC). PR computes importance nodes in graph. This is commonly used in search engines.

PR algorithm assign a number to each node of graph which indicates importance of that node. SSSP computes shortest path between two nodes of graph. SSSP has various applications in networks and also used to find critical path.

SP used in results as a shorter form for SSSP. CC finds connected components in a graph which has several applications in image processing. TC counts triangles in a graph. TC is used in social networks. BC finds the most important node between two given nodes.

BC has various applications in social networks and computer networks. There are a few benchmark suites such as GAP [20] or CRONO [21].

These suits are known to researchers and are being used to evaluate their architectures and ideas.

These benchmark suits include similar applications such as SSSP, BC and PR to process graphs.

We chose GAP [20] benchmark suite as a baseline for graph algorithms. We simulate well-known real-world workloads from Stanford large network dataset (SNAP) [22].

Results and Discussion

This section provides the results of simulations.

A. Execution Time

Figure 6 shows speedup of the chosen workloads normalized to the baseline HMC. Due to better management of poor graph locality, the proposed architecture could reach a better execution time for all

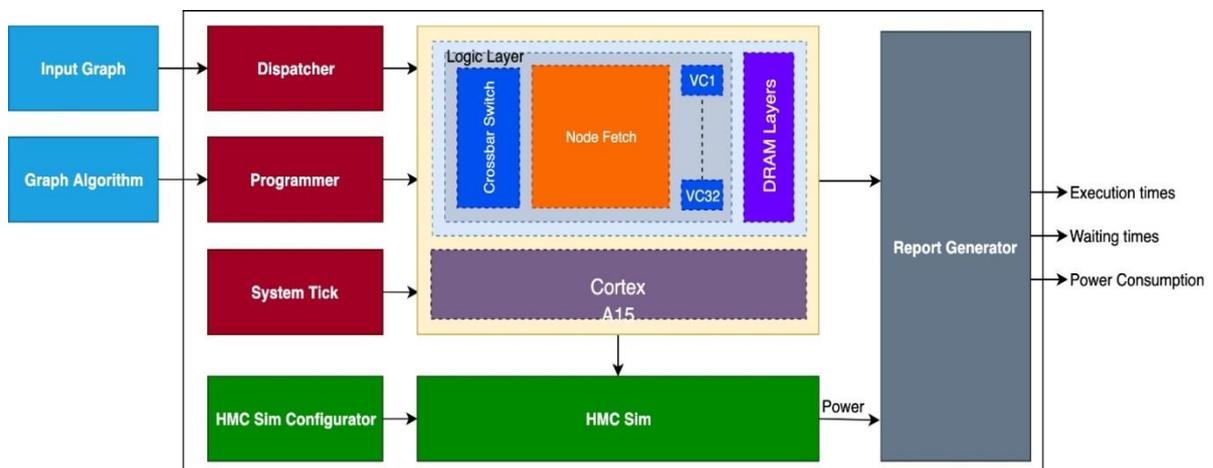


Fig. 5: Block diagram of the simulator.

of the benchmark graph applications. The simulation results indicate an average speed up of 3.3x in comparison to the baseline.

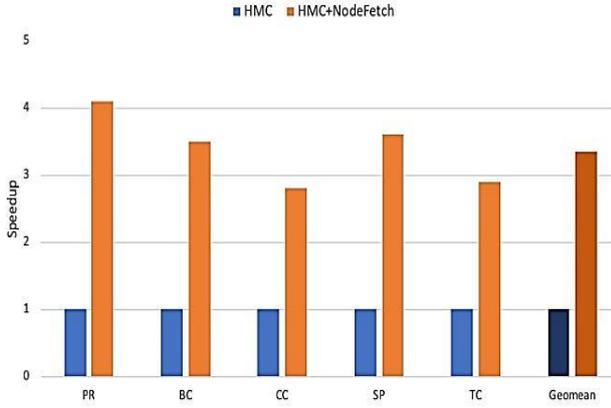


Fig. 6: NodeFetch Speedup in comparison to baseline.

B. Power and Energy Consumption

Improving the execution time and offloading parts of computation from the processor to HMC, results in reduction of energy consumption, indeed at the cost of energy overheads caused by the additional hardware. Simulation results shows that the system energy overheads are significantly less than energy savings. As a result, the overall system energy is decreased. Figure 7 shows the system energy of the NodeFetch normalized to the baseline HMC.

An average of 69% energy reduction is obtained for the evaluated workloads.

Table 2 shows area and power consumption overheads of NodeFetch, Tesseract [17] and Enhanced Tesseract [7] relative to one HMC device. It demonstrates the proposed idea leads to a very low power and area overhead.

Table 2: Area and power overhead NodeFetch, Tesseract and Enhanced Tesseract relative to HMC

Relative to Logic layer of One HMC	Tesseract [17]	Enhanced Tesseract [7]	NodeFetch
Area Overhead	9.6%	9.73%	0.1%
Power Overhead	40%	42%	4.5%

Table 3 shows area power density of NodeFetch, Tesseract and Enhanced Tesseract.

The highest power density of the logic die across all workloads in our design is 14mW/mm² which is by far below the maximum power density that does not require faster DRAM refresh using a passive heat sink (i.e. 133mW/mm² [23]).

Table 3: Power density and area comparison

	Tesseract [17]	Enhanced Tesseract [7]	NodeFetch
Max Power Density	94 mW/mm ²	96 mW/mm ²	14 mW/mm ²
Area Overhead per HMC	21.75 mm ²	22 mm ²	0.07 mm ²

The total area of a NodeFetch unit is 0.07mm² which solely account for 0.1% area overhead. Our approach increases the average power consumption by 4.5% in comparison to HMC, which may lead a negative impact on device temperature. However according to recent measurements in industrial research on thermal feasibility of 3D-stacked PIM [23], the power consumption should be within the power budget. Therefore, proposed idea is thermally feasible.

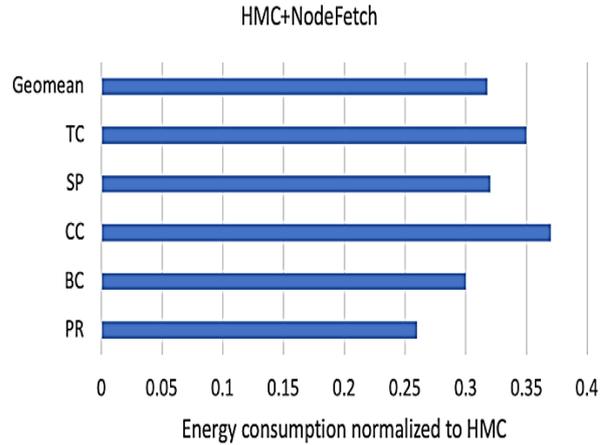


Fig. 7: Relative system energy consumption of the proposed architecture normalized to the baseline architecture.

Conclusion

This paper proposed an optimization to PIM-based graph processing with the help of HMC. Most of the techniques in the field of processing-in-memory, hire methods to reduce movement of data between processor and memory.

This paper proposed a method to reduce graph processing execution time and energy consumption by reducing cache misses while processing a graph. Proposed idea which named NodeFetch, adds a command to HMC for that purpose.

NodeFetch helps graph processing to have a better performance by increasing locality and decreasing irregularity.

Simulation results shows that NodeFetch in average is 3.3x faster than HMC itself, and reduces energy consumption by 69% in average.

Author Contributions

M.A. Mosayebi, and M. Dehyadegari contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript

Acknowledgment

We thank the editor and all anonymous reviewers.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

References

- [1] X. Chen, "GraphCage: Cache Aware Graph Processing on GPUs," arXiv preprint arXiv:1904.02241, 2019.
- [2] J.E. Gonzalez, Y. Low, H. Gu, D. Bickson, C. Guestrin, "Powergraph: Distributed graph-parallel computation on natural graphs," in Proc. the 10th Symposium on Operating Systems Design and Implementation (OSDI): 17-30, 2012.
- [3] A. Fidel, N.M. Amato, L. Rauchwerger, "Kla: A new algorithmic paradigm for parallel graph computations," in Proc. 23rd International Conference on Parallel Architecture and Compilation Techniques (PACT): 27-38, 2014.
- [4] S. Hong, H. Chafi, E. Sedlar, K. Olukotun, "Green-Marl: a DSL for easy and efficient graph analysis," in Proc. Seventeenth International Conference on Architectural Support for Programming Languages and Operating Systems: 349-362, 2012.
- [5] T.J. Ham, L. Wu, N. Sundaram, N. Satish, M. Martonosi, "Graphicionado: A high-performance and energy-efficient accelerator for graph analytics," in Proc. 2016 49th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO): 1-13, 2016.
- [6] S. Ghose, K. Hsieh, A. Boroumand, R. Ausavarungnirun, O. Mutlu, "Enabling the adoption of processing-in-memory: Challenges, mechanisms, future research directions," arXiv preprint arXiv:1802.00320, 2018.
- [7] M.A. Mosayebi, A.M. Hasani, M. Dehyadegari, "Enhanced graph processing in PIM accelerators with improved queue management," *Microelectron. J.*, 94: 104637, 2019.
- [8] M. Zhang et al., "GraphP: Reducing communication for PIM-based graph processing with efficient data partition," in Proc. 2018 IEEE International Symposium on High Performance Computer Architecture (HPCA): 544-557, 2018.
- [9] G. Dai et al., "Graphh: A processing-in-memory architecture for large-scale graph processing," *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.*, 38(4): 640-653, 2018.
- [10] L. Nai, R. Hadidi, J. Sim, H. Kim, P. Kumar, H. Kim, "Graphpim: enabling instruction-level pim offloading in graph computing frameworks," in Proc. 2017 IEEE International symposium on high performance computer architecture (HPCA): 457-468, 2017.
- [11] H.M.C. Specification, "2.1, Nov. 2015, Hybrid Memory Cube Consortium," Tech. Rep.
- [12] B. Soltani Farani, H. Dorosti, M. Salehi, S. M. Fakhraie, "Ultra-low-energy dsp processor design for many-core parallel applications," *JECEI*, " J. Electr. Comput. Eng. Innovations (JECEI), 8(1): 71-84, 2019.
- [13] L. Song, Y. Zhuo, X. Qian, H. Li, Y. Chen, "GraphR: Accelerating graph processing using ReRAM," in Proc. 2018 IEEE International Symposium on High Performance Computer Architecture (HPCA): 531-543, 2018.
- [14] G. Kim, J. Kim, J. H. Ahn, J. Kim, "Memory-centric system interconnect design with hybrid memory cubes," in Proc. the 22nd international conference on Parallel architectures and compilation techniques: 145-155, 2013.
- [15] J. Kim, W. Dally, S. Scott, D. Abts, "Cost-efficient dragonfly topology for large-scale systems," *IEEE micro*, 29(1): 33-40, 2009.
- [16] J. Kim, W. J. Dally, D. Abts, "Flattened butterfly: a cost-efficient topology for high-radix networks," in Proc. 34th Annual International Symposium on Computer Architecture: 126-137, 2007.
- [17] J. Ahn, S. Hong, S. Yoo, O. Mutlu, K. Choi, "A scalable processing-in-memory accelerator for parallel graph processing," in Proc. 42nd Annual International Symposium on Computer Architecture: 105-117, 2015.
- [18] D.-I. Jeon, K.-B. Park, K.-S. Chung, "HMC-MAC: Processing-in-memory architecture for multiply-accumulate operations with hybrid memory cube," *IEEE Comput. Archit. Lett.*, 17(1): 5-8, 2017.
- [19] A. Addisie, V. Bertacco, "Centaur: Hybrid processing in on/off-chip memory architecture for graph analytics," in Proc. 57th ACM/IEEE Design Automation Conference (DAC): 1-6, 2020.
- [20] S. Beamer, K. Asanović, D. Patterson, "The GAP benchmark suite," arXiv preprint arXiv:1508.03619, 2015.
- [21] M. Ahmad, F. Hijaz, Q. Shi, O. Khan, "Crono: A benchmark suite for multithreaded graph algorithms executing on futuristic multicores," in Proc. EEE International Symposium on Workload Characterization: 44-55, 2015.
- [22] J. Leskovec, A. Krevl, "SNAP Datasets: Stanford large network dataset collection," ed, 2014.
- [23] Y. Eckert, N. Jayasena, and G. H. Loh, "Thermal feasibility of die-stacked processing in memory," 2014.

Biographies



Mohammad Amin Mosayebi received his MSc. Degree from K. N. Toosi University in 2019 in computer engineering. He is currently a Ph.D. student at Shahid Beheshti University researching on bio inspired hardware designs and bio-medical signal processing.



Masoud Dehyadegari received his Ph.D. degree from University of Tehran, Tehran, IRAN, in 2013 in computer engineering. He is currently an Assistant Professor of school of computer engineering with the K. N. Toosi University of Technology. His research interests include Low-power system design, Network-on-chips, and Multi-Processor System-on-chip.

Copyrights

©2020 The auor(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

M.A. Mosayebi, M. Dehyadegari, "NodeFetch: High performance graph processing using processing in memory," J. Electr. Comput. Eng. Innovations, 9(1): 67-74, 2021.

DOI: [10.22061/JECEI.2020.7453.393](https://doi.org/10.22061/JECEI.2020.7453.393)

URL: http://jecei.sru.ac.ir/article_1486.html





Research paper

A New Routing Protocol in MANET using Cuckoo Optimization Algorithm

*S. Tabatabaei**, *H. Nosrati Nahook*

Department of Computer Engineering, Faculty of Engineering, Higher Educational Complex of Saravan, Saravan, Iran.

Article Info

Article History:

Received 14 April 2020
Reviewed 15 June 2020
Revised 14 September 2020
Accepted 15 November 2020

Keywords:

MANET
Routing algorithm
Cuckoo Optimization Algorithm (COA)
AODV

*Corresponding Author's Email
Address:
shtabatabaey@yahoo.com

Abstract

Background and Objectives: With the recent progressions in wireless communication technology, powerful and costless wireless receivers are used in a variety of mobile applications. Mobile networks are a self-arranged network, which is including of mobile nodes that communicate with each other without a central control. Mobile networks gained considerable attention due to the adaptability, scalability, and costs reduction. Routing and power consumption is a major problem in mobile networks because the network topology changes frequently. Mobile wireless networks suffer from high error rates, power constraints, and limited bandwidth. Due to the high importance of routing protocols in dynamic multi-hop networks, many researchers have paid attention to the routing problem in Mobile Ad hoc Networks (MANET). This paper proposes a new routing algorithm in MANETs which is based upon the Cuckoo optimization algorithm (COA).

Methods: COA is inspired by the lifestyle of a family of birds called cuckoo. These birds' lifestyle, egg-laying features, and breeding are the basis of the development of this optimization algorithm. COA is started by an initial population. There are two types of population of cuckoos in different societies: mature cuckoos and eggs. This algorithm tries to find more stable links for routing.

Results: Simulation results prove the high performance of proposed work in terms of throughput, delay, hop count, and discovery time.

Conclusion: The cuckoo search convergence is based on the establishment of the Markov chain model to prove that it satisfies the two conditions of the global convergence in a random search algorithm. Also, the cuckoo search that suitable for solving continuous problems and multi-objective problems. We have done a lot of experiments to verify the performance of the Cuckoo algorithm for routing in MANETs. The result of experiments shows the superiority of the proposed method against a well-known AODV algorithm.

Introduction

MANET is a collection of mobile wireless nodes that are organized in a temporary network without any infrastructure and central management. Routers in these networks can move freely and organized themselves. Therefore, maybe the network topology changes frequently without any prior knowledge [1]. This kind of networks can be used individually, or they can be

connected to the Internet. Features such as multi-hop, mobility, supporting heterogeneity, limited power, and scarce bandwidth, make the designing of routing protocols a challenging task.

Due to the high dynamicity of these networks, maybe some routing protocols, require frequent control packet exchange to be informed about connection losing [2]. Mobile users want to make a connection in an

environment that there is not any prior infrastructure. MANETs do not have any fixed communication infrastructure. Due to the lack of communication infrastructures such as access points (Aps) and a base station (BS), communication in such a network depends on forwarding nodes. Therefore, each node in the network except the role of the client could take the role of router and exchange data packets for other nodes [3].

In this kind of network, the network topology changes frequently because of node mobility. Besides, some nodes may join or leave the network, or maybe they alter to the sleep mode. Because of these characteristics, the main problem in MANETs is to route data packets efficiently. Energy consumption is another problem. Since most of the mobile hosts are equipped with batteries and have limited energy, energy consumption should be minimized to a lower extent. Many scenarios such as emergency operations, military operations, conferences, etc. can benefit from these kinds of networks because networks with fixed infrastructure cannot fulfill these requirements [3].

The dynamic nature of MANETs provided a very convenient area for applying heuristic and meta-heuristic approaches. Many routing protocols have been proposed for these networks, but most of them make simple assumptions about the amount of mobility in these networks. Since, in different situations automaticity and informed interactions between nodes avoids link failure and cause the path reliability till the end of the transmission phase, investigating the possibility of using heuristic and meta-heuristic approaches in MANET routing could be an attractive research topic [5].

In MANETs, network nodes do not have any prior information about network topology, and because of that, they are forced to discover the location of a destination node to communicate with other nodes. In other words, before transmitting data packets to the destination, the sender node has to perform some operations to discover the right location of the destination node.

In [6] suggested an Adaptive Neuro-Fuzzy Inference System (ANFIS) and make use of some fixed and mobile agents. The proposed algorithm uses three kinds of agents: (a) fixed any-cast management agent, (b) fixed optimization agent, and (c) mobile any-cast route construction agent.

In [6] introduced a new algorithm called Multi-Channel On-Demand Routing with Coordinate Awareness (MCORCA). This algorithm uses several wireless channels to enhance network performance in MANETs. The proposed cross-layer approach uses channel assignment and collision avoidance strategies. This approach uses two kinds of channels one control

channel for scheduling and another for data transmitting. MCORCA is an extended version of a request-based routing algorithm named ORCA for single-channel networks.

In [8] proposed a multipath routing protocol for OSDM-TDMA mesh networks, which uses several paths to reach the destination. When these paths face a problem, they can replace each other immediately. In [9] had made some modifications on MMQR in which alternative paths make some changes in the network, and they will be ordered adaptively. They suggest a heuristic model for selecting attractors by which biological entities would be identified for dynamic environment changes.

With today's wireless technologies, most of the applications, need Quality of Service (QoS), self-organization, and self-healing services. Authors [10], consider the above mentioned problems, and propose a new cross-layer method that improves existing routing protocols. These improvements include adding new decision metrics to all network layers and using a fuzzy method with [11] proposed a new algorithm named Energy-Aware Span Routing Protocol (EASRP) which utilizes energy-saving approaches such as Span and Adaptive Fidelity Energy Conservation Algorithm (AFECA) approaches. Also, energy consumption could be reduced to a large extent because of a hardware circuit named Remote Activated Switch (RAS) for awaking nodes that are in sleep mode. These kinds of energy-saving protocols can work better than passive routing protocols. Even though, there are lots of problems that should be solved when using the EASRP algorithm when a hybrid protocol.

In [12] presented a novel AODV based routing algorithm called QoS Routing Protocol Based on AODV (AQA-AODV). This algorithm established routes based on the QoS requirements of a specific application. A mechanism for bandwidth estimation of a path and an adaptive scheme that could provide helpful feedback about the current status of the network is utilized. Therefore, special applications could optimally tune their transmission rate. Also, a route recovery (retrieving) approach has been suggested, which provides an error detection mechanism in path links and communication reestablishment.

Authors [13] proposed a reliable and Markov Chain based multi-path algorithm, which has a great performance in terms of energy consumption. The proposed protocol computes multiple paths from source to destination and among them, a path will be chosen in a random manner which consumes a smaller amount of energy. Also, due to the transmission of data packets using random paths, this protocol makes the data flow safer.

Authors [14] proposed a fuzzy logic-based routing protocol for mobile ad hoc network MANETs. The proposed method uses the fuzzy logic strategy to select suitable routes based on the energy level of the battery and speed of mobile nodes. The proposed method with the AODV protocol was simulated in OPNET simulator 10.5 and simulation results revealed that the proposed protocol outperformed the AODV protocol considering the throughput, data dropped, packet delivery rate, and hop count. Authors [15] proposed an intelligent algorithm to find the feasible route in an ad-hoc network. The proposed protocol uses the Cuckoo Search (CS) algorithm, to satisfy the constraint of QoS in MANET and select the QoS path based on a computation of the highest fitness value with three parameters such as routing load, remaining energy, and hop count. The proposed method was simulated with the Ant Colony Optimization protocol, Particle Swarm Optimization, and AODV protocol was simulated and their evaluation was accomplished based on important metrics, including mobility, scalability, and congestion. The Simulation result analysis shows that the proposed protocol is outperformed for all the above mentioned Metrics. Authors [16] proposed a new routing protocol in Manet networks to increase the efficiency of the network with a zone-Based method, which has been created by modifying the multi-path directing protocol AOMDV protocol, by improving power usage. This protocol uses metrics such as the power of battery consumption and network lifetime to selecting the best path from the multipath based on labels, node tracking, and power analysis. Proposed routing protocol uses an energy sensitive system. The Simulation result shows that the proposed method more reliable than is with AOMDV and AODV protocols.

Authors [17] proposed a method in Optimized Link State Routing (OLSR) protocol that select the best multi-point relay node for forwarding packet control to increase the QoS of the link while making a routing decision. The proposed protocol was simulated with an NS-2 simulator by varying the pause time of nodes, simulation time of nodes, and speed of nodes. Many simulations have been performed for evaluating the performance of the proposed protocols regarding throughput, end to end delay, energy cost per packet, and node's average remaining energy. Experimental results reveal that the protocol has a more efficient mechanism for routing and a more satisfactory solution for maximizing the throughput in MANET.

Cuckoo Algorithm

Cuckoo optimization algorithm [18] is a novel and powerful optimization method, which is suitable for continuous and nonlinear optimization problems. Like other evolutionary algorithms, cuckoo starts with an

initial generation. The population of cuckoos has some amount of eggs that lay them in some host birds. Some eggs that are similar to the host bird eggs have more chances to be born and become a mature cuckoo. Other eggs will be recognized by the host cuckoo, and they will be destroyed. The amount of grown eggs determines the suitability of the nest of that region. More grownup eggs in a region result in more profit for that region. Thus, a region where the largest amount of eggs are saved is a parameter that cuckoo intended to optimize. The cuckoo optimization algorithm can be applied to purposes such as (a) districted and continuous optimization problems, (b) layout designing problems, (c) scheduling and sequence operation problems, (d) designing intelligent networks, and I problems that have many parameters for optimization, and their global solution could not be achieved by these methods. The superiority of the cuckoo algorithm in comparison to other algorithms is latent by multiple operations of cuckoo optimization operators, like lying eggs and migration. Other evolutionary optimization algorithms use operations that have a single purpose. However, in the cuckoo optimization algorithm, the defined operators have several purposes simultaneously.

Operation of Proposed Protocol

When a source node attends to send data to an away node, following step should be performed:

1. Source node 'S' checks its routing table to find a path to the destination. If the available path has the required amount of energy to send packets, then data will be sent through that path.

2. If there is no path toward the destination, a route discovery process will be started and a Route Request message (RREQ) is created continuing destination address and itself parameters. This message will be broadcasted on the network. It is worth noting, the proposed protocol before sending the message computed the reminded energy after sending the packet using (1) and will be added to the RREQ message, and the value of the fitness function is set to zero.

$$PR = PL - ((PS)/BW) * TP \quad (1)$$

where PR is the remaining energy, PL energy level of the battery, PS data packet size, BW available bandwidth, and YP is the required energy for data transmitting.

3. Through the path, each intermediate node that receives the RREQ message extracts the information and determines the suitability of the link by which the packet is received using the cuckoo optimization algorithm. In other words, it checks whether this node has a sufficient amount of energy for data transiting or not? The node has a sufficient amount of energy, and the value of the fitness function is greater than the threshold value does the same operation for the available bandwidth

parameter. For example, if there is enough bandwidth for packet transmitting, it means that the fitness function has a greater value in comparison to the threshold value. It means that the node has a good potential for data transmission. It should be noted, the value of fitness function for remained energy and bandwidth computed using (2) and (3), respectively. Furthermore, the cumulative value of the fitness function can be computed using (4).

$$FITP = PR/MAX PL \tag{2}$$

$$FITBW = ABW/MAX BWL \tag{3}$$

$$FIT = ((\alpha * FIT) + (FITBW + FITP)) \tag{4}$$

Where in (2), FITP is the fitness function for energy, PR is reminded energy, and MAX PL is the maximum level of the battery.

In (3), FITBW is the fitness function for bandwidth, ABW is the available bandwidth, and MAX BWL is the maximum level of bandwidth.

In (4), FIT is the general faintness function, α is the reduction confection for function convergence which is equal to 0.9. Also, the intermediate node checks to find a path to the destination in the routing table. If the path has existed, creates a Route Reply (RREP) packet, which includes the fitness value of the node and itself parameters, then sending them back to the previous node. If there is no path to the destination, the RREQ message will be broadcasted again.

When the RREQ message reaches the destination node, like intermediate nodes, it extracts itself parameters and fitness value from the packet and checks the suitability of the node using the cuckoo optimization algorithm. Furthermore, determines those paths that have higher fitness value using extracted information from the RREQ message. Four paths are selected as best paths, then one that has a higher fitness value will be chosen for data transmitting, and the other three paths are maintained as alternate paths for link failures.

The destination node creates an RREP message and copies the source address and destination address from the RREQ message and also adds its parameters and cumulative fitness function to the RREP message and sends it back to the source node.

Each node in the reverse path extracts parameters and adds its parameters to the RREP message to determine the suitability of the link. The source node uses the discovered path to send data to the destination.

The source code of proposed method show in below:

Proposed Cuckoo Search Algorithm

Procedure routing function ($f(p)$, $v=(P1, \dots, Pn)N$)

Get available node of N Paths $p_i \quad i=1,2, \dots, n$ for source and destination

While $i \leq N$ or(stop criterion)

Get intermediate node n_i

Evaluate its fitness function $F(n_i)$ with (4).

If intermediate node to find a path to the destination in the routing table

Creates a Route Reply (RREP) packet and RREP message sends to the source node.

Else

RREQ message will be broadcasted again.

End if

Increment i

End While

If RREQ message reaches the destination node

Choose an optimized path with high fitness value.

Creates a Route Reply (RREP) packet and RREP message sends to the source node.

End if

The best solution is chosen or kept(Quality solutions nests);

The current best is rank to find solutions

The source node uses the discovered path to send data to the destination.

Post-process results and visualization

End

Simulating the Proposed Method

S. Simulation Environment

This paper uses the OPNET Modeler 10.5 [19] to implemented and verify the effectiveness of the proposed method and comparing it to the well-known AODV (Ad hoc On-Demand Distance Vector Routing) [20] algorithm and fuzzy logic-based reliable routing protocol (FRRP) [20] As shown in Fig. 1 network topology contains 30 nodes that are distributed in an 1127 * 1127 meter area. Furthermore, it is assumed that each node can turn in 0 to 359 degrees.

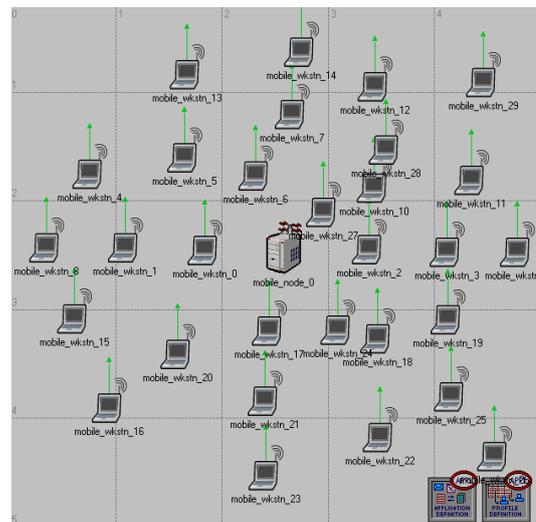


Fig. 1: Network topology for simulating network in an error-free state.

The rotating time is between 0 to 200 seconds and the movement of each node is 1 to 10 meters. Figure 2 shows a view of the node editor for the network model. Transmission range of each node equals 250 meters and the bandwidth is considered about 1 to 10 megabytes randomly.

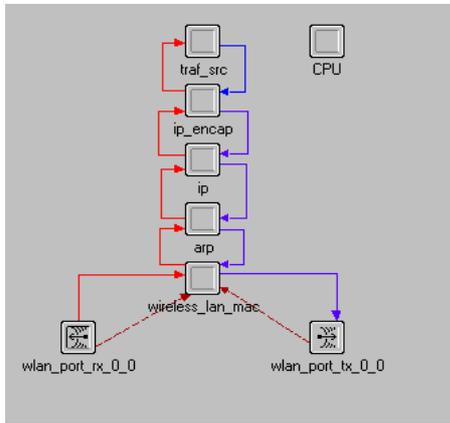


Fig. 2: Node editor for simulated model.

We use the following metrics to evaluate the performance of our algorithm. We want to analyze the performance of the proposed algorithm in two error-free and failure states.

Throughput: the amount of data that successfully passed through a network in a given time.

Received packets: number of successfully received packets to the destination.

Delay: the average time between the sending time of a packet and its receiving time of a packet which is calculated for all packets.

The number of hops: the distance between source and destination in hops.

Route discovery time: the required time to find a path to the destination.

In the first scenario, we want to evaluate the performance of the proposed algorithm in terms of throughput for error-free and failure states. In the error-free state, all nodes that are participating in the network work without any specific problem. However, in a failure state may be some nodes get out from the network and encounter some serious problems. Now we want to evaluate the performance of our proposed algorithm against the AODV algorithm and FRRP to show the superiority of the proposed method. As it can be seen from Fig. 3 and Fig. 4, the cuckoo based algorithm has a higher throughput in both stats than AODV algorithm and after learning with fuzzy logic, FRRP protocol can determine appropriate links and in case the link is suitable, it can transmit data through that link. Hence, the throughput rate for FRRP scenario is more than the proposed algorithm. Also, in the mode with errors, after

learning, FRRP protocol can select appropriate links for transmitting data and it can have more throughput rate.

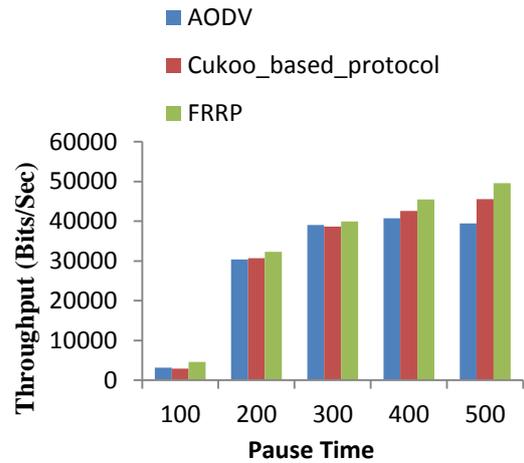


Fig. 3: Throughput of Cuckoo against AODV in error-free state.

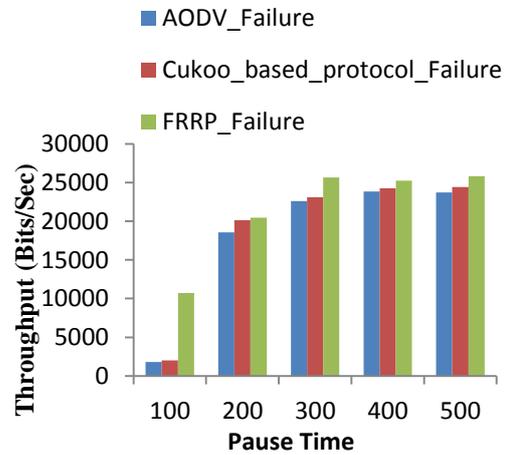


Fig. 4: Throughput of Cuckoo against AODV in failure state.

In the second scenario number of successfully received packets are counted to show the great performance of our proposed algorithm in comparison with the AODV algorithm and FRRP algorithm. Based on detailed information about the cuckoo algorithm which selects more stable paths to the destination, this algorithm can deliver a greater number of packets to the destination than AODV algorithm also the FRRP protocol can determine suitable links for data transmission than proposed algorithm. Figure 5 and Fig. 6 illustrate the performance evaluation of the proposed algorithm in error-free and failure states.

Considering Fig. 7 and Fig. 8, it can be concluded that with the progress of simulation time, network delay will be decreased by the Cuckoo algorithm than AODV algorithm, and the reason is that suitable paths are chosen for data transmitting. In other words, the routes that deliver packets with high possibility are established and there is no need to discover routes again. Therefore, network delay decreases significantly.

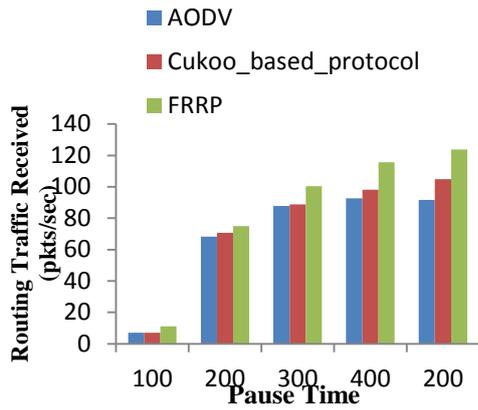


Fig. 5: Number of received packet in cuckoo algorithm and AODV in error-free state.

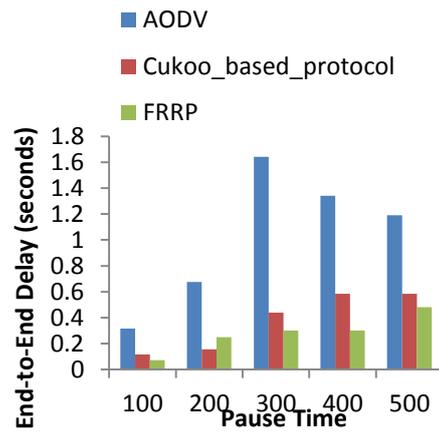


Fig. 7: Network delay for Cuckoo against AODV in error-free state.

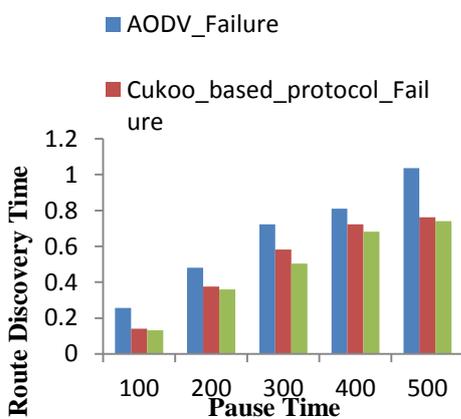


Fig. 6: Number of received packet on cuckoo algorithm and AODV in failure state.

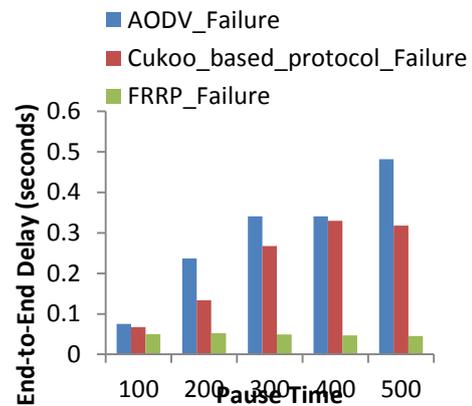


Fig. 8: Network delay for Cuckoo against AODV in failure state.

Considering Fig. 7 and Fig. 8, it can be concluded that with the progress of simulation time, network delay will be decreased by the Cuckoo algorithm than AODV algorithm, and the reason is that suitable paths are chosen for data transmitting. In other words, the routes that deliver packets with high possibility are established and there is no need to discover routes again. Therefore, network delay decreases significantly. Computing available bandwidth and energy level increase the delay but in our algorithm, the delay has a great reduction which shows that the required time for computing and comparison is low and does not impact the performance of the network. In the first seconds, the suitable links are not specified yet and because of that network delay is high. Also FRRP protocol, after learning, can consider stability parameters and determine appropriate links from transmitting data. Moreover, in the failure mode, FRRP protocol selects appropriate links for transmitting data.

It is noted that cuckoo uses those paths they have a lower hop count. In other words, cuckoo makes paths that have a lower hop count in comparison to AODV and tries to choose the shortest path to decrease the consumed energy for data transmitting.

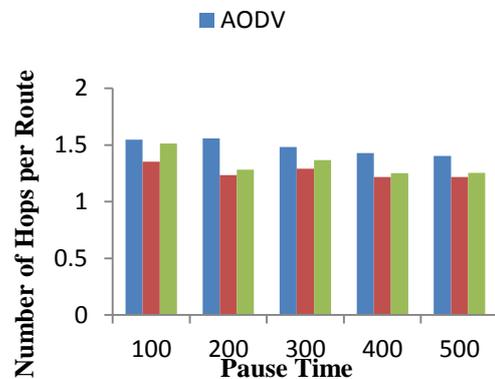


Fig. 9: Number of hop counts for each route in Cuckoo and AODV algorithm in error-free state.

Figure 9 and Fig. 10 show the achieved result from comparing cuckoo, FRRP protocol, and AODV algorithms in terms of hop count. In Fig. 11 and Fig. 12 network discovery time for Cuckoo, FRRP protocol, and AODV are computed and this parameter has a great reduction for the Cuckoo algorithm than the AODV algorithm.

This reduction arises from the fact that suitable links are recorded in the neighbor routing table of the nodes. Initially, network discovery time is high, and it is because the network is not fully trained with the Cuckoo algorithm, and suitable links are not determined yet.

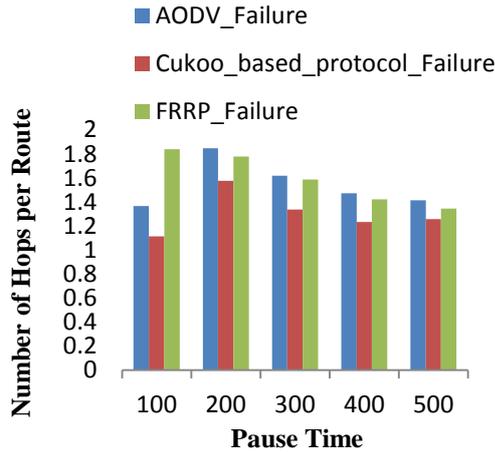


Fig. 10: Number of hop counts for each route in Cuckoo and AODV algorithm in failure state.

Also as shown in this figure, as the simulation time increases, route discovery time for FRRP protocol decreases in comparison with the proposed method. This is attributed to the FRRP protocol can determine suitable links for data transmission.

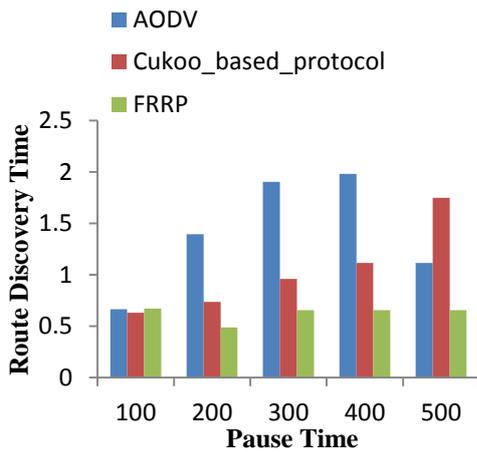


Fig. 11: Discovery time for Cuckoo and AODV in error-free stat.

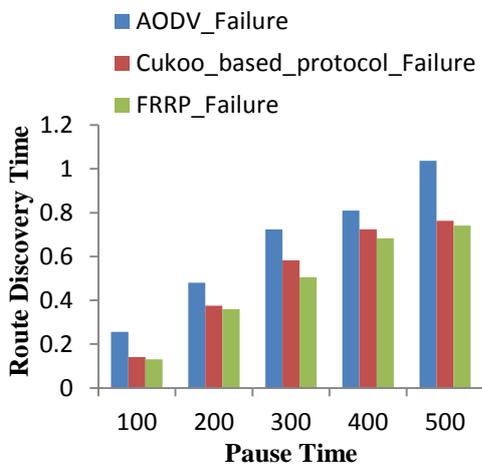


Fig. 12: Discovery time for Cuckoo and AODV in failure stat

Conclusion

MANETs have gained a significant attraction in the last decades. Routing in MANTS became a very challenging research area due to its special characteristics. In this paper, a new algorithm for routing in MANETs is proposed, which is based on the Cuckoo optimization algorithm. Cuckoo optimization algorithm [18] is a novel and powerful optimization method that is suitable for continuous and nonlinear optimization problems. Cuckoo search is a Metaheuristic algorithm that continuously iterating the optimal solution until the globally optimal solution is obtained, which helps to provides finding the optimal path. The cuckoo search convergence is based on the establishment of the Markov chain model to prove that it satisfies the two conditions of the global convergence in a random search algorithm. Also, the powerful attribute such as portability and platform independence of the cuckoo search is of profound significance it has strong global search ability. That suitable for solving continuous problems and multi-objective problems [22]. We have done a lot of experiments to verify the performance of the Cuckoo algorithm for routing in MANETs. The result of experiments shows the superiority of the proposed method against a well-known AODV algorithm.

Author Contributions

S. Tabatabaei designed the experiments, analysis the data and wrote the manuscript. H. Nosrati Nahook carried interpreted the results and revised the manuscript.

Acknowledgment

This work is completely self-supporting, thereby no any financial agency’s role is available.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

References

- [1] S. Taneja, A. Kush, "A survey of routing protocols in mobile ad hoc networks," *Int. J. Innovation Manage. Technol*, 1(3): 279-285, 2010.
- [2] S.K. Sarkar, T.G. Basavaraju, C. Puttamadappa, "Ad hoc mobile wireless networks: principles, protocols and applications," CRC Press, 2007.
- [3] T. Camp, J. Boleng, V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Commun. Mobile Comput.*, 2(5): 483-502, 2002.

- [4] J. Hoebeke et al., "An overview of mobile ad hoc networks: applications and challenges," *J Comm Netw*, 3(3): 60-66, 2004.
- [5] I. Almomani et al, "FEAR: Fuzzy-based energy aware routing protocol for wireless sensor networks," *Int'l J. Commun. Netw. Sys. Sci.*, 4(6): 403-415, 2011.
- [6] V.R. Budyal, S.S. Manvi, "ANFIS and agent based bandwidth and delay aware anycast routing in mobile ad hoc networks," *J. Netw. Comput. Appl.*, 39: 140-151, 2014.
- [7] Y. Wang, J.J. GARCIA-LUNA-ACEVES, "A distributed cross-layer routing protocol with channel assignment in multi-channel MANET," in *Proc. 2015 International Conference on Computing, Networking and Communications (ICNC): 1050-1054, 2015.*
- [8] Y. Yuan, H. Chen, M. Jia, "An optimized ad-hoc on-demand multipath distance vector (AOMDV) routing protocol," in *Proc. 2005 Asia-Pacific Conference on Communications: 569-573, 2005.*
- [9] H. Lee, D. Jeon, "A mobile ad-hoc network multi-path routing protocol based on biological attractor selection for disaster recovery communication," *ICT Express*, 1(2): 86-89, 2015.
- [10] T. Carvalho, J.J. Júnior, R. Frances, "A new cross-layer routing with energy awareness in hybrid mobile ad hoc networks: A fuzzy-based mechanism," *Simul. Modell. Pract. Theory*, 63: 1-22, 2016.
- [11] S.K. Das, S. Tripathi, "Energy efficient routing formation algorithm for hybrid ad-hoc network: A geometric programming approach," *Peer-to-Peer Netw. Appl.*, 12: 102-128, 2019.
- [12] W.E. Castellanos, J.C. Guerri, P. Arce, "A QoS-aware routing protocol with adaptive feedback scheme for video streaming for mobile networks," *Comput. Commun.*, 77: 10-25, 2016.
- [13] S. Sarkar, R. Datta, "A secure and energy-efficient stochastic multipath routing for self-organized mobile ad hoc networks," *Ad Hoc Networks*, 37: 209-227, 2016.
- [14] S. Tabatabaei, F. Hosseini, "A fuzzy logic-based fault tolerance new routing protocol in mobile ad hoc networks," *Int. J. Fuzzy Syst.*, 18(5): 883-893, 2016.
- [15] V.V. Mandhare, V.R. Thool, R.R. Manthalkar, "QoS Routing enhancement using metaheuristic approach in mobile ad-hoc network," *Comput. Networks*, 110: 180-191, 2016.
- [16] R. Sahu, S. Sharma, M.A. Rizvi, "ZBLE: Zone based efficient energy multipath protocol for routing in mobile ad hoc networks," *Wireless Pers. Comm.*, 113: 2641-2659, 2020.
- [17] R. Jain, I. Kashyap, "An QoS aware link defined OLSR (LD-OLSR) routing protocol for MANETs," *Wireless Pers. Comm.*, 108(3): 1745-1758, 2019.
- [18] R. Rajabioun, "Cuckoo optimization algorithm," *Appl. Soft Comput.*, 11(8): 5508-5518, 2011.
- [19] <https://www.opnet.com>
- [20] C. Perkins et al. "Ad hoc ondemand distance vector (AODV) routing," RFC, 3561, 2003.
- [21] S. Ghasemnezhad, A. Ghaffari, "Fuzzy logic based reliable and real-time routing protocol for mobile ad hoc networks," *Wireless Pers. Comm.*, 98(1): 593-611, 2018.
- [22] G. Wang, "A comparative study of cuckoo algorithm and ant colony algorithm in optimal path problems," in *proc. MATEC Web of Conferences: 03003, 2018.*

BIOGRAPHIES



Shayesteh Tabatabaei received her B.S. and M.S. degrees in Computer Engineering from Islamic Azad University, Shabestar, Iran, in 2006 and 2008 and then continued her study for Ph.D. in Computer Science Branch Software in Islamic Azad University, Science and Research Branch, Tehran, Iran. Her research interests include routing, QoS, security, and clustering for mobile ad hoc

networks and wireless sensor networks.



Hassan Nosrati Nahook received his bachelor's and master's degrees in Sistan and Baluchestan and Islamic Azad universities, Research Sciences Branch, Tehran, Iran in 2010 and 2013, respectively. His research interests include machine learning, Fuzzy systems and methods, evolutionary processing and bioinformatics.

Copyrights

©2020 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

S. Tabatabaei, H. Nosrati Nahook, "A new routing protocol in MANET using cuckoo optimization algorithm," *J. Electr. Comput. Eng. Innovations*, 9(1): 75-82, 2021.

DOI: 10.22061/JECEI.2020.7511.397

URL: http://jecei.sru.ac.ir/article_1487.html





Research paper

Optimal PSS Parameters Design Based on a Novel Objective Function for Small Signal Stability Enhancement

P. Naderi^{1,*}, B. Ehsan Maleki¹, H. Beiranvand²

¹Power Engineering Department, Faculty of Electrical Engineering, Shahid Rajaei Teacher Training University, Tehran, Iran.

²Power Engineering Department, Faculty of Engineering, Lorestan University, Lorestan, Iran.

Article Info

Article History:

Received 22 May 2020
Reviewed 23 July 2020
Revised 19 September 2020
Accepted 21 November 2020

Keywords:

Power system
Power system stabilizer
Variable slope damping scale
Dynamic stability

*Corresponding Author's Email
Address:

p.naderi@sru.ac.ir

Abstract

Background and Objectives: In this paper, a novel objective function is proposed for designing the power system stabilizers (PSSs). Although the object of the previous designs was to enhance the critical modes' stability, the derived stability indices were, to some extent, low and in some cases not acceptable at all. The prospect of attaining higher stability motivated authors to design a new objective function in this study. In all the previous objective functions, the same priority is accorded to all modes, and an objective function is generally defined. A novel function is presented, called Variable Slope Damping Scale (VS DS), based on the assumed variable slope for the straight line in the fan-shaped region, which is an area in the complex plane for determining the eigenvalue placement range, with a reference tip at the negative point. This can be an efficient solution to the low value of critical modes' stability. In general, more damping for critical modes and lower priority for searching non-critical modes are taken as key points. The result of applying VS DS leads to a high value of damping scales for critical modes. The nonlinear simulation results and eigenvalues analysis has demonstrated that the proposed approach in this study is highly effective in damping the most critical modes.

Methods: The proposed method assumes a variable slope for the straight line of the convergence region (specified area for placement of poles) in a fan-shaped type. Indeed, the increase in critical mode's damping scale is taken into account as a key point to introduce a powerful objective function.

Results: The value of the damping scale and also the overall dynamic stability of the test system has increased by using the proposed objective function.

Conclusion: Also, it has been shown that a variable slope convergence region is better than that of a constant slope one to the optimal tuning of WAPSS. In other words, the value of the damping scale with the proposed method over the existing techniques clearly shows that the proposed objective function is more effective than the other ones.

Introduction

In most power plants, PSS is taken into account as efficient equipment to solve oscillations and stability problems in our process [1]-[3]. Effective application of stabilizers is mostly related to the effectiveness of

objective function, especially when several PSSs are considered simultaneously in power system designing. There are lots of objective functions offered to provide a better condition for PSS design. However, all of them consider the same strategy for different modes and

ignore the fact that critical modes play a significant role in the overall stability of the power system. Regarding the significant effects of objective functions, there are some limitations associated with the displacement of the whole eigenvalues in each power system. It goes without saying that this issue has experienced both growing interest and extensive efforts when it came directly to the design of a powerful objective function [4]-[12].

This paper proposes a new objective function known as the VSDS objective function. It is obvious that achieving greater stability is of great importance and all equipment must be designed to meet this goal. The commonly used objective functions proposed in previous works are relatively appropriate, but they are not as effective in providing high stability. Therefore, in this paper, some thoughts and novelties are provided to address this important issue. In [4], a method is proposed for designing a wide-area power system stabilizer based on a new multi-objective function. In this function, critical modes are displaced to improve the system stability, which the stabilizer is designed in the minimum-phase with less control gain. In the objective function presented in [5], there is no limitation in the confrontation of high-frequency oscillation and just all damping factors are placed in a rectangular region. In [6], the performance of various commonly used objective functions including integral square error (ISE), integral time square error (ITSE), integral absolute error (IAE), and integral time absolute error (ITAE) are analyzed. The objective functions are used to tune the proportional-integral-derivative (PID) controller values in various power systems. The results established that their performance change based on the power system size, which is a weak point. The objective function presented in [7], similar to one in [5], the limitation of damping factors is performed, but it still has problems with the high-frequency oscillations. In [8], a multi-objective optimization procedure including both performance and robustness criteria is considered for formulating the dual-dimensional supplementary damping controller (SDC) and accelerating PSS (PSS2B). However, in designing the objective function, there is no difference between system modes, and the same method is used for their displacement. The objective function presented in [9], the problem of high-frequency oscillation has been solved, but there is no investigation on the modes works on damping factor limitation. In [10], a novel performance index is proposed to evaluate a system's ramp response. The study is applied to the tuning of a PSS of a simple power system, at several operating points of the synchronous machine. It explores the process of tuning controllers for conditions that better reflect realistic operating conditions, such as input signals of different shapes. In [11], a novel

objective function based on power system exciter frequency response is proposed to design a robust PSS using heuristic optimization techniques, in order to damp the electromechanical oscillations at very low frequencies (0.1-3 Hz), that often tend to grow with time and cause system instability. The objective function presented in [12], which is called the damping scale objective function, all the problems with the previous well-known objective function such as lightly damped and higher-frequency modes were eliminated somewhat, and the only problem is the slight convergence difficulty. Now, the proposed objective function has all advantages of the previous method, and there are no convergence problems. In this investigation, the novelty is about finding a convergence region that result in a more damping scale for critical modes. Actually, the main idea of this paper is to deal with some particular modes by defining a special convergence region. Moreover, this method scrutinizes some issues like the non-constant slope of the convergence region or less energy consumption for un-critical modes. Also, to enhance the system stability, a combinational input of both local and remote signals was regarded for damping controllers in this study. The GRSA (a newly proposed objective function inspired by general relativity theory) with considering VSDS objective function was applied in PSSs designing, and its positive effects have been investigated. This paper has organized as follows: Second section introduces the power system model and PSSs structure, which is accompanied by a review of both new and customary objective functions. Third section presents simulation results and a comparison in investigating the efficacy of the proposed methods under various system operating conditions. Some conclusions are also explained in next section.

Problem Statement

The problem statement is defined in three parts as follows:

A. Power system model and PSSs structure

The closed-loop nonlinear modeling can be constructed with a set of nonlinear differential-algebraic equations by the following form:

$$\dot{X} = f(X, U) \quad (1)$$

Regardless of the damping controller's structure, the linearized incremental model of the interconnected power system around an equilibrium point is generally shown as follows:

$$\Delta \dot{X} = A \Delta X + B \Delta U \quad (2)$$

This paper implements a commonly utilized lead-lag PSS and the IEEE-type-ST1 excitation system, which are similar to Fig.1, [2]. The considered PSS structure and i^{th} angular speed deviation are denoted as follows:

$$G_i(s) = K_i \frac{T_w s (1 + sT_{1i})(1 + sT_{3i})}{(1 + sT_w)(1 + sT_{2i})(1 + sT_{4i})} \quad (3)$$

$$\Delta\omega_i = (\gamma_L \Delta\omega_{Li} + \gamma_R \sum_k \Delta\omega_{Rk})$$

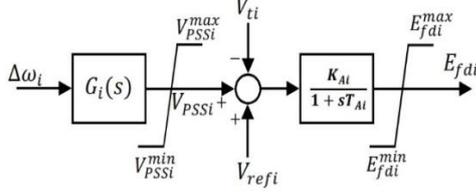


Fig. 1: Overall interconnection of i^{th} PSS with IEEE-type-ST1 excitation system.

In this paper, the input signal of PSSs is not just a local signal, and a combination of local and remote signals is considered as the input signal. In this regard, the second equation, which determines the input signal of the i^{th} generator, is based on the local signal and combination of all remote signals impacting the determined generator. It should be mentioned that remote inputs for each generator will be selected based on a sensitive analysis of its state variables. The weighting coefficients can be calculated by the optimization method, but constant values are also acceptable. The values of parameters are available in the Appendix section.

B. Overview of Previous Objective Function

Up to now, there have been several attempts with the presentation of a different objective function in order to achieve higher small-signal stability. A schematic of these convergence regions in four well-known objective functions is depicted as shaded areas in Fig. 2. The objective function depicted in Fig. 2(d) has been constructed by defining the damping scale function in the form of (4) [12]. As such, this damping scale objective can be written as below (5):

$$\chi \equiv \frac{-(\sigma - \sigma_0)}{\sqrt{(\sigma - \sigma_0)^2 + \omega^2}} \times 100\% \quad (4)$$

$$\text{Min } F_1 = \sum_{y=1}^{n_y} (\chi_0 - \min_{1 \leq q \leq n_q} \chi_q)_y \quad (5)$$

In the mentioned convergence region, damping factors and damping scales of eigenvalues are respectively lower than σ_0 and upper than χ_0 . Furthermore, the slope of the straight line of the fan-shaped region can be calculated as following [12]:

$$\text{Slope}_{\chi_0} = \pm \sqrt{\frac{1}{\chi_0^2} - 1} = \pm \frac{1}{\sqrt{\frac{\zeta_0}{1 - \zeta_0^2} - \frac{\sigma_0}{\omega}}} \quad (6)$$

As it is obvious in Fig. 2(d), there is not any problem associated with the existence of high-frequency or low-frequency lightly damped modes. So, the substantial drawback of prior works has been resolved in this work.

For $\sigma_0 = 0$ in the (6), Slope_{χ_0} will have the same value as Slope_{ζ_0} , and for σ_0 on the negative side of the damping scale axis, the magnitude of Slope_{χ_0} is smaller than that of Slope_{ζ_0} [12].

C. Proposed Objective Function

In the proposed objective function (VSDS) it has been assumed that there is a variable slope for the straight line of the fan-shaped region, which its value is different in each range of damping factor interval $[\sigma_0 \sigma_d]$. This variable slope has been considered for the most important modes and σ_d depicts the endpoint of this range. Indeed, the elevate of the critical mode's damping is taken into account as a key point to introduce a powerful objective function. Considering the incremental slope in the VSDS and its less energy expenditure for uncritical modes, it provides a higher damping scale and frequency limitation. Since the effectiveness of the last introduced well-known objective function has been compared to previous ones in [12], this objective function will be considered as a comparison criterion in designing a new proposed objective function. The objective function must be optimized to limit the eigenvalues considering the different damping scale values in each damping factor interval. The objective function can be expressed as follows:

$$\begin{cases} \text{Min } F_2 = \sum_{y=1}^{n_y} (\sum_{i=1}^{n_q} [\chi'_k - \chi_i])_y \\ \text{for } \sigma_i \in (\sigma'_k \text{ or } [\sigma_d - \infty)) \end{cases} \quad (7)$$

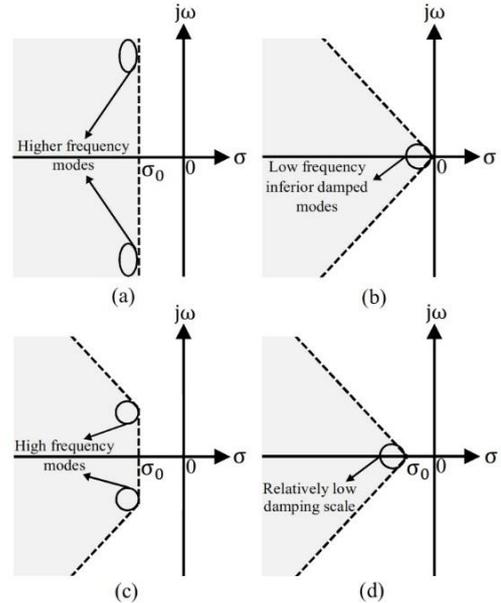


Fig. 2: Convergence regions of four objective functions. (a) Rectangular region of damping factor [5]. (b) Fan-shaped region with tip at the origin of damping ratio [13]. (c) D-shaped region of damping factor and damping ratio [7]. (d) Fan-shaped region with tip at σ_0 and the damping ratio ζ_0 of damping scale [12].

where χ'_k and σ'_k are shown in Fig. 3(a), and they can be computed as follows (indices of i and k are associated with different eigenvalue and interval step numbers, respectively):

$$\chi'_k = \begin{cases} \chi_k & \sigma \in [\sigma_0 \ \sigma_d] \\ \chi_f & \sigma \in [\sigma_d \ -\infty) \end{cases} \quad (8)$$

The parameter of χ_k is defined as follows:

$$\chi_k = \{\chi_0 - (k - 1) * m \mid \sigma \in \sigma'_k\} \quad (9)$$

where,

$$\sigma_{k-1} \leq \sigma'_k < \sigma_k; \quad \sigma_k = \sigma_0 - k * r \quad (10)$$

In (7), all of the operating conditions have taken into consideration by the first sigma operator, and the value of inner parenthesis under different damping factor intervals has determined by the second sigma operator. The values of the required constant parameters such as k ; r ; m and etc. are listed in Appendix (Table A2).

As it is obvious in (11), the slope of the convergence curve is directly in the effect of χ'_k and varies by its change. In a specific interval, χ'_k is equal to χ_k , and it reduces by the rate of $(k-1) \times m$ times. As such, we have various values for χ'_k , and they will be obtained by the same procedure at different intervals (see Fig. 3(a)).

The slope of the convergence region, similar to (6), is computed as:

$$Slope_{\chi'_k} = \pm \sqrt{\frac{1}{(\chi'_k)^2} - 1} \quad (11)$$

According to (11), the slope of the convergence region is variable in different ranges. In the range of $[\sigma_0, \sigma_d]$, the value of $Slope_{\chi'_k}$ is smaller than $Slope_{\chi_0}$ and for $\sigma < \sigma_d$, these values are equal. The evaluation of the proposed objective function in Table 1 demonstrates its superiority, as it has other significant features in addition to the considerable characteristics of the well-known objective functions. It should be mentioned that axis scales of the depicted figures in this section are not exact, and they are considered just for the sake of better presentation of our objective function characteristics (all of the terms/variables are defined in the nomenclature section).

Table 1: Evaluation of well-known objective functions

Convergence region	Component	High Frequency limitation	Damping factor limitation	High damping scale
Fig.2. (a)	σ	No	Yes	No
Fig.2. (b)	ξ	No	No	No
Fig.2. (c)	σ, ξ	No	Yes	No
Fig.2. (d)	χ	Yes	Yes	No
Fig.3. (b)	χ'	Yes	Yes	Yes

Simulation Results and Discussion

The case study here is a 10-machine 39-bus system as a medium multi-machine power system, which is shown in Fig. 4. Generally speaking, the PSS optimum parameters will be obtained by applying objective functions (OFs) of F_1 and F_2 and considering below constrained optimization problems.

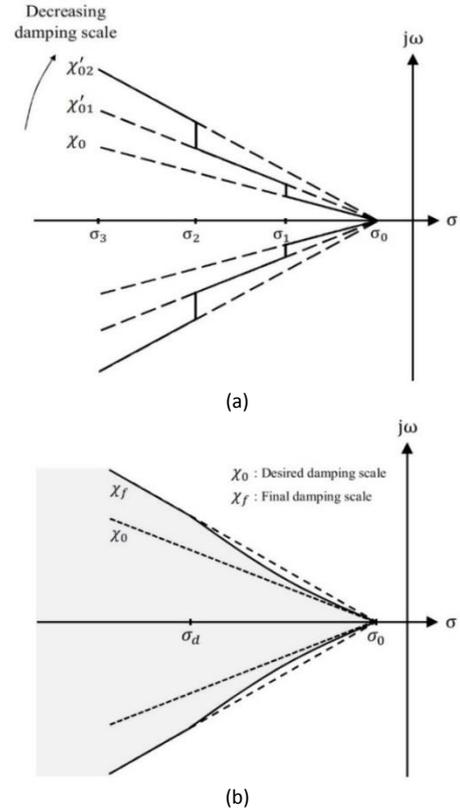


Fig. 3: (a) Slope change of convergence region, (b) Convergence region of the proposed objective function.

Minimize (OFs) subject to:

$$K_i^{min} \leq K_i \leq K_i^{max} \quad (12)$$

$$T_{1i}^{min} \leq T_{1i} \leq T_{1i}^{max} \quad (13)$$

$$T_{2i}^{min} \leq T_{2i} \leq T_{2i}^{max} \quad (14)$$

$$T_{3i}^{min} \leq T_{3i} \leq T_{3i}^{max} \quad (15)$$

$$T_{4i}^{min} \leq T_{4i} \leq T_{4i}^{max} \quad (16)$$

The range of optimizing our parameters is $[0:001 \ 50]$ for K_i and $[0:01 \ 1.0]$ for T_{1i}, T_{2i}, T_{3i} and T_{4i} . Furthermore, the optimization algorithm of GRSA has been chosen to be used in this optimization procedure. In the way of designing the proposed PSSs, OFs (F_1 and F_2) and GRSA are utilized to pursue a comparative study in our objective functions. In this paper, PSSs' parameters are optimized based on two kinds of an objective function, and the design based on each of which is considered as a scenario.

A. Tested system

The single-line diagram of the IEEE 10 machine (New England) power system (Fig. 4) is considered for simulation and the detail of system data is taken from . In this study, in order to achieve the overall dynamic stability of the test system with the presented approaches, all generators, except G2, which is an equivalent power source of the U.S.-Canadian interconnection system, are equipped with PSSs. To design robust PSSs, two different operating conditions in addition to our base case are characterized by the system under both severe loading conditions and critical line outage. Here the "Robust" term is used in this concept that the designed controllers will be responsive to load changes and will keep the system stable. The three cases are defined as:

Base case: normal condition.

Case 1: outage of the line between buses 21-22.

Case 2: outage of the line between buses 21-22 and 10% increase in the loads at even buses.

The above-mentioned cases are chosen based on the authors' knowledge which is obtained by various simulated cases. The obtained results also demonstrate that the above-mentioned cases are two sensitive ones for designing the PSSs. The results of sensitivity analysis, as it is needed for designing PSSs, illustrate the oscillation among given i^{th} and k^{th} generators as follows:

Considering " \sim " as oscillation symbol, $G_1 \sim G_3$, $G_4 \sim (G_2 \& G_6)$, $G_5 \sim G_4$ and $G_7 \sim G_6$. Therefore, all of generators G1, G4, G5 and G7 are select to equip with the WAPSSs installation. The feedback combination signal of the i^{th} generator is $\omega_i - \sum_k \omega_k$.

B. Eigenvalue Analysis and Simulation Results

The worst numerical values of stability characteristic out of eigenvalue's analysis under multiple operating conditions are summarized in Table 2. For the sake of brevity, only some critical modes are presented. Additionally, for each of objective function different values of χ_0 are considered over the maximum obtained value regarded to them. To give more details, 4:08% is considered for F_1 and 8:16% for F_2 both of which are obtained based on GRSA. After describing the accuracy of tuning method, our objective function can be assessed on the basis of root locus analysis at three operating points, which they are drawn into the complex s-plane as shown in Fig. 5. For space constraint reason and the significance of dominant poles, only the eigenvalues in the right side of -4 in horizontal axis are considered. The Root-Locus plot shows the placement of eigenvalues bounded based on the objective functions of F_1 and F_2 with the tip at the σ_0 . According to that, higher values of critical modes' damping scale are clearly

visible for the proposed objective function. In this way, it can be concluded that the proposed method effectively extend the power system stability limitations. Furthermore, convergence characteristic of objective functions under applied optimization technique has been depicted in Fig. 6. The PSS parameters searched by F_2 and utilizing GRSA method, and in addition, the requirement data in the process of evaluating proposed approach are written in Appendix (Tables A1 and A2).

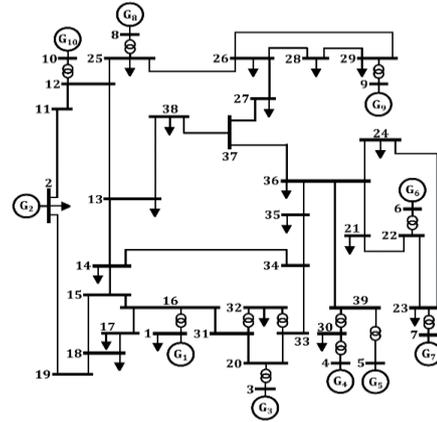


Fig. 4: Single-line diagram of 10-machine 39-bus (New England) Test System.

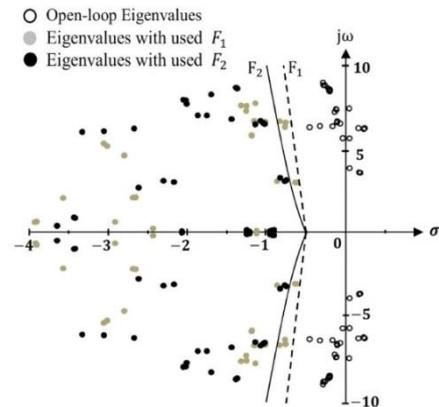


Fig. 5: Dominant eigenvalues of the system result by three operating conditions.

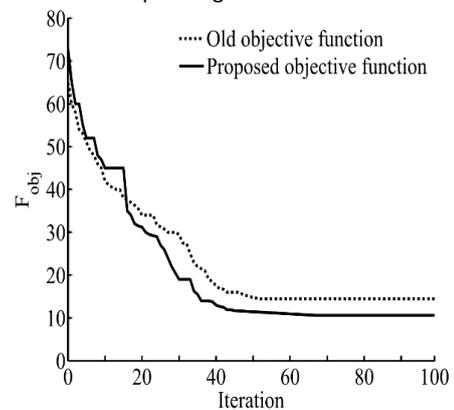


Fig. 6: Convergence characteristic of the proposed objective function.

Table 2: Dynamic stability characteristics of the test system in different conditions

Case type	Base Case				Case 1				Case 2					
	sorted	f (Hz)	σ	$\zeta(\%)$	$\chi(\%)$	f(Hz)	σ	$\zeta(\%)$	$\chi(\%)$	f(Hz)	σ	$\zeta(\%)$	$\chi(\%)$	
Non-PSSs		0.593	0.049	-1.32	-14.5	0.556	0.168	-4.81	-18.7	0.551	0.184	-5.32	-19.38	
		0.959	0.238	-3.95	-12.1	0.959	0.161	-2.68	-10.9	0.866	0.048	-0.88	-10.02	
		1.001	-0.070	1.12	-6.81	0.867	-0.005	0.09	-9.04	1.139	0.013	-0.18	-7.15	
		1.133	-0.123	1.73	-5.29	1.127	-0.115	1.63	-5.41	0.950	0.227	-9.04	-3.80	
F_1	Max σ	1.012	-0.75	11.86	4.08	0.461	-0.62	21.08	4.31	0.463	-0.62	20.90	4.20	
		0.474	-0.86	27.95	12.24	1.008	-0.76	11.95	4.14	0.970	-0.76	12.48	4.38	
		0.001	-1.12	100	99.00	1.018	-1.01	15.7	8.06	1.022	-0.83	12.89	5.21	
	Min ζ	1.012	-0.75	11.86	4.08	1.008	-0.76	11.95	4.14	0.970	-0.76	12.48	4.38	
		1.116	-1.25	17.63	10.72	1.018	-1.01	15.7	8.06	1.022	-0.83	12.89	5.21	
		1.024	-1.15	17.66	10.13	1.162	-1.25	16.9	10.24	1.186	-1.14	15.14	8.57	
	Min χ	1.012	-0.75	11.86	4.08	1.008	-0.76	11.95	4.14	0.463	-0.62	20.9	4.20	
		1.024	-1.15	17.66	10.13	0.461	-0.62	21.08	4.31	0.970	-0.76	12.48	4.38	
		1.116	-1.25	17.63	10.72	1.018	-1.01	15.7	8.06	1.022	-0.83	12.89	5.21	
	F_2	Max σ	0.506	-0.82	25.20	10.27	0.474	-0.76	25.00	9.00	0.484	-0.74	23.90	8.16
			0.029	-0.90	98.01	91.07	0.016	-0.92	99.40	97.25	0.020	-0.95	99.14	96.41
			1.005	-1.04	16.28	8.55	1.026	-1.07	16.46	8.90	0.998	-1.02	16.09	8.30
Min ζ		1.005	-1.04	16.28	8.55	1.321	-1.37	16.29	10.43	0.998	-1.02	16.09	8.30	
		1.329	-1.38	16.41	10.59	1.026	-1.07	16.46	8.90	1.008	-1.03	16.20	8.49	
		1.266	-1.69	20.88	14.90	0.993	-1.14	18.00	10.24	1.328	-1.38	16.33	10.50	
Min χ		1.005	-1.04	16.28	8.55	1.026	-1.07	16.46	8.90	0.484	-0.74	23.90	8.16	
		0.506	-0.82	25.20	10.27	0.474	-0.76	25.00	9.00	0.998	-1.02	16.09	8.30	
		1.329	-1.38	16.41	10.59	0.993	-1.14	18.00	10.24	1.008	-1.03	16.20	8.49	

C. Nonlinear Time-Domain Simulation

The nonlinear time-domain simulation of the power system is conducted, which uses a power flow program to calculate the dynamic initial conditions, and Differential- Algebraic-Equations (DAEs) of the power system is solved with MATLAB/Simulink through using the Ordinary Differential Equation (ODE) solver which is presented in. Additionally, to ascertain the designed PSSs' robustness in the state of variable operating conditions, there is a study which completely investigates the system response by considering three following disturbance (Corresponding to D:1, D:2, and D:3):

- A three-phase short circuit fault at bus 39 at $t = 1s$. The fault is cleared by tripping the line between buses 36-39 after 100 ms.
- 0.05 pu increment in V_{ref} of the G_3 excitation system at $t = 1s$, with a duration of 100 ms.
- 50% load increasing at bus 35 at $t = 1s$ with a duration of 100 ms.

Because of space constraints, some studies are only conducted with considering the first disturbance.

D. Evaluation of Objective Functions

The main goal of this section is to concentrate on the positive effect of the proposed objective function by considering the first scenario under various operating conditions. Since all of the state variables of the system, such as speed response and generators' rotor angle deviation, obviously demonstrate the behavior of the power system, rotor angle deviations are selected to be analyzed here.

The rotor angle of G_3 with respect to G_1 (as reference generator) under the presence of the first disturbance is depicted graphically in Fig. 7(a)-Fig.7 (c). These figures show that the generators' response by applying the old objective functions undergo more oscillations. In contrast, designed PSSs by considering proposed objective function can effectively mitigate the oscillations so that settle faster and reach the new steady-state condition. The latest figure shows that the designed PSSs can be effective even under severe conditions and robustness will be obtained.

E. Evaluation of Performance Indices

In order to do a clear and complete system response for different disturbances, two common performance indices (PI) which are respectively associated with the settling time and overshoot in speed response (ISTSE, and ISE) [9], are described as (17) and (18). It has been shown that these indices are more appropriate ones than others for representing system characteristics.

$$ISTSE (PI_1) = \sum_{i=1}^{n_m} \int_{t=0}^{t=t_{sim}} (t \times \Delta\omega_i(t))^2 dt \quad (17)$$

$$ISE (PI_2) = \sum_{i=1}^{n_m} \int_{t=0}^{t=t_{sim}} (\Delta\omega_i(t))^2 dt \quad (18)$$

Furthermore, Table 3 represents the information about the performance indices of PI_1 and PI_2 by considering all three disturbances in the mathematical equations. Here, it is obvious that the obtained result of performance indices for the second scenario is better in all of the disturbances. As such, from the stability point

of view, it should be noted that there has been supremacy in relation to our proposed technique in different disturbances.

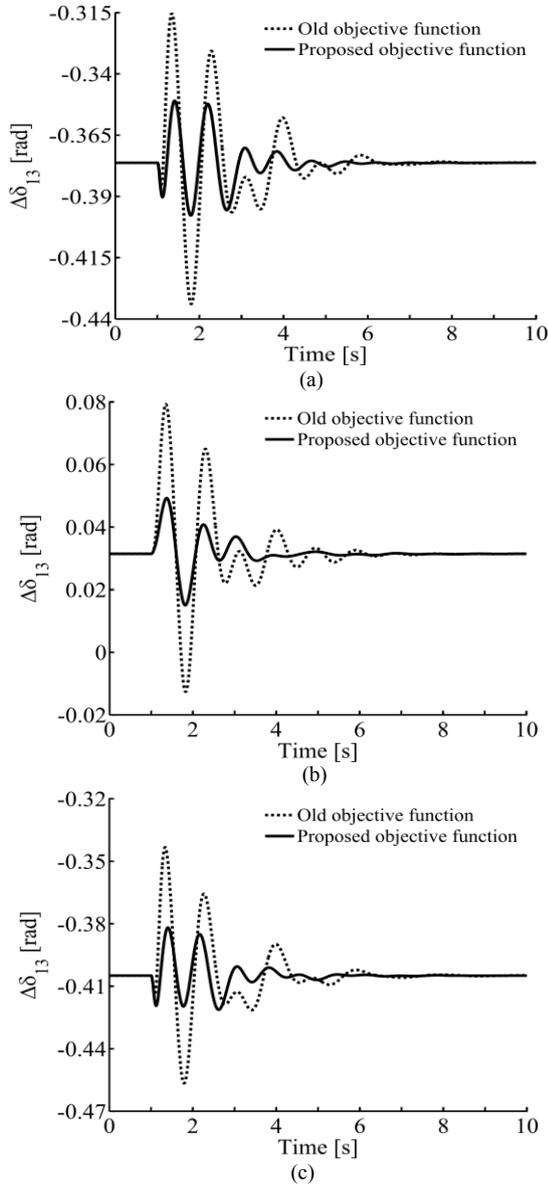


Fig. 7: Rotor angle deviation with the first disturbance (a) Base case for evaluating of objective function (b) Case 1 for evaluating of objective function (c) Case 2 for evaluating of objective function.

Table 3: Performance indices of the system response under different scenarios and disturbances

Scenario	Parameter	Disturbance		
		D.1	D.2	D.3
S.1	PI_1	0.0318	1.0665	0.0124
	PI_2	0.1716	2.5131	0.3427
S.2	PI_1	0.0290	0.9221	0.0117
	PI_2	0.1533	2.1974	0.3226

Conclusion

The purpose of this paper was to introduce a novel objective function and use it in designing powerful PSSs. To overcome the low value of the obtained damping

scale in old objective functions, some additional restrictions have been applied to the convergence region and led to the VSOS objective function. To be more specific, further displacements of critical modes were considered to enhance the overall dynamic stability of the power system. To prove the superiority of the proposed function, a comparison study has been taken place over the values of damping scale and performance indices. It has been shown that a variable slope convergence region is better than that of a constant slope one in obtaining optimal parameters. Small signal analysis and time-domain simulations have been implemented on a multi-machine power system and the effectiveness of the proposed method has been demonstrated.

Appendix

Optimal parameters of the designed PSSs and also requirement data for evaluating the proposed approach are listed in Table A1 and Table A2, respectively.

Table A1: Optimal parameters of the designed PSSs

Parameter	Symbol	Unit	Value
Damping factor interval number	k	-	1:1000
Desired value of damping factor	σ_d	-	-10.5
Factor of gradient reduction	m	-	0.0001
Step length of damping factor intervals	r	-	0.01
Initial value of damping factor	σ_0	-	-0.5
Number of operating conditions	n_y	-	3
Number of steps in slope change	n_k	-	1000
Washout time constant	T_w	Sec	10
Weight of local machine feedback signal	γ_L	-	1
Weight of remote machine feedback signal	γ_R	-	-1

Table A2: Requirement data

Gen	Old Objective Function				
	K	T_1	T_2	T_3	T_4
G_1	16.9523	0.5033	0.1708	0.3986	0.0209
G_3	2.6856	0.3305	0.0206	0.6281	0.2350
G_4	45.0241	0.5252	0.0201	0.4207	0.5239
G_5	4.8984	0.3366	0.0200	0.3934	0.1849
G_6	45.3028	0.6905	0.2952	0.3499	0.0228
G_7	2.895	0.5971	0.0200	0.8119	0.6676
G_8	3.0008	0.4011	0.2050	0.5041	0.0229
G_9	5.5970	0.5146	0.0205	0.1888	0.2132
G_{10}	12.405	0.1861	0.1607	0.4746	0.0244
Gen	Proposed Objective Function				
	K	T_1	T_2	T_3	T_4
G_1	8.9127	0.6506	0.0203	0.5743	0.3967
G_3	6.8628	0.2696	0.1584	0.4414	0.0382
G_4	49.3471	0.3670	0.3838	0.6252	0.0210
G_5	46.7737	0.4394	0.2333	0.3453	0.0200
G_6	5.7158	0.4169	0.3143	0.7340	0.0215
G_7	6.7360	0.3990	0.3317	0.4173	0.0233
G_8	5.2918	0.2736	0.2544	0.6973	0.0209
G_9	30.0610	0.1307	0.0242	0.6043	0.1316
G_{10}	3.4127	0.6069	0.4990	0.4321	0.0221

Author Contributions

B. Ehsanmaleki and H. Beiranvand have found a new method for enhancing existing objective function. H. Beiranvand presented new method for carrying out of nonlinear time-domain simulation. B. Ehsanmaleki and P. Naderi interpreted the results and wrote the manuscript.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

A, B	Constant matrices of state-space equations
C_r	Crossover probabilities
d	Dimension of particle position
Dn	Dimension of cloud
E_{fdi}	Field voltage of i^{th} excitation system (V)
En	Entropy or uncertainty measurement of qualitative concept
Ex	Mathematical expectation of cloud drop
F_1	Damping scale objective function
F_2	Variable slope damping scale objective function
f_{ave}	Average value of function
f_g	Best solution for optimization algorithm till now
f_{rand}	Random selected objective function value in current subspace
FNC	Forward normal cloud generator
G_i	Transfer function of i^{th} PSS
K_{Ai}	Gain of i^{th} excitation system
k	Subscript index of machine numbers
K_i	Optimal value of i^{th} PSS's stabilizer gain
$K_{V,ij}(t)$	Distance of best position and a random particle position in search subspace s
l	Number of particle
M_g	Maximum generation
M_p	Mutation probabilities
m	Factor of gradient reduction
N_p	Population size
n	Number of state variable
n_k	Number of steps in slope change

n_q	Number of system's eigenvalues
n_y	Number of system's operating condition
n_m	Number of test system's machine
q	Subscript index of system's eigenvalues
r	Step length of damping factor intervals
S_s	Search space size
S	Number of search subspaces
$Slope_{\zeta_0}$	Convergence region slope in old objective function ($^\circ$)
$Slope_{\chi_0}$	Convergence region slope in objective function F_1 ($^\circ$)
$Slope_{\chi'_k}$	Convergence region slope in objective function F_2 ($^\circ$)
s	Denotation of complex plane
$sign$	Signum function
S_s	Search space size
T_{Ai}	Time constant of i^{th} excitation system
T_{ji}	j^{th} optimal time constant of i^{th} PSS
T_w	Washout time constant
t_{sim}	Simulation time duration
$T_i(t)$	Best particle position obtained in exploration till now
$T_{ij}(t)$	Best position of j^{th} component of i^{th} particle
$T^{rand,s}$	Random position in search subspace s
$T^{Best,s}$	Best position in search subspace s
$T_j^{Best,s}(t)$	j^{th} component of best position in search subspace s
TG	Global best particle position obtained in tensor till now
TB_i	Best position obtained in the exploration till now
$X_i(t)$	Cloud drop of i^{th} agent
$X_{i,rand}$	Negative random selected cloud drop of i^{th} agent
y	Subscript index of system operating condition
U	Vector of input variable
$V_{ij}(t)$	Velocity vector of i^{th} agent of j^{th} dimension
V_{PSSi}	Output voltage of i^{th} PSS (V)
V_{refi}	Reference voltage of i^{th} generator (V)
V_{ti}	Terminal voltage of i^{th} generator (V)
X	Vector of state variable

γ_L	Weight of local machine feedback signal
γ_R	Weight of remote machine feedback signal
σ	Damping factor of eigenvalue
σ_0	Initial value of damping factor
σ_d	Desired value of damping factor
σ_k	Damping factor of k^{th} step
σ'_k	Damping factor of k^{th} interval
σ_i	Damping factor of i^{th} eigenvalue
ω_{Li}	Angular speed of i^{th} local machine (rad/s)
ω_{Rk}	Angular speed of k^{th} remote machine (rad/s)
ω_i	Angular speed of i^{th} machine (rad/s)
χ_i	Damping scale of i^{th} eigenvalue
χ_q	Damping scale of q^{th} eigenvalue
χ_0	Expected damping scale's value in different conditions
χ_f	Final damping scale value in different conditions
χ'_k	Damping scale's expected value of k^{th} interval
ζ_0	Initial value of damping ratio
ω_t	Weighting factor in step length equation
ε	Very small number
∂	Symbol of deviation
Δ	Symbol of difference
\sum	Symbol of summation

References

[1] A. Sabo, N.I. Abdul Wahab, M. Lutfi Othman, M.Z. Mohd Jaffar, H. Beiranvand, "Optimal design of power system stabilizer for multimachine power system using farmland fertility algorithm," *Int. Trans. Electr. Energy Syst.*, 30(12): e12657, 2020.

[2] W. Peres, F.C. Coelho, J.N. Costa, "A pole placement approach for multi-band power system stabilizer tuning," *Int. Trans. Electr. Energy Syst.*, 30(10):e12548, 2020.

[3] J.A. Lala, C.F. Gallardo, "Adaptive tuning of power system stabilizer using a damping control strategy considering stochastic time delay," *IEEE Access*, 8: 124254-124264, 2020.

[4] M.R. Shakarami, I.F. Davoudkhani, "Wide-area power system stabilizer design based on grey wolf optimization algorithm considering the time delay," *Electr. Power Syst. Res.*, 133: 149-159, 2016.

[5] C. Chang, S. Yang, "Optimal multiobjective planning of dynamic series compensation devices for power quality improvement," *IEE Proc. Gener. Transm. Distrib.*, 148(4): 361-370, 2001.

[6] K. Jagatheesan, B. Anand, K. N. Dey, A.S. Ashour, S.C. Satapathy, "Performance evaluation of objective functions in automatic generation control of thermal power system using ant colony

optimization technique-designed proportional-integral-derivative controller," *Electr. Eng.*, 100(2): 895-911, 2018.

[7] M. Abido, Y. Abdel-Magid, "Robust design of multimachine power system stabilisers using tabu search algorithm," *IEE Proc. Gener. Transm. Distrib.*, 147(6): 387-394, 2000.

[8] Y. Hashemi, H. Shayeghi, M. Moradzadeh, "Multi-objective attuned design of damping controllers using frequency-based functions in smart grids," *Knowledge Based Syst.*, 76:176-88, 2015.

[9] M. Abido, "Optimal design of power-system stabilizers using particle swarm optimization," *IEEE Trans. Energy Convers.*, 17(3): 406-413, 2002.

[10] B. Mohandes, Y.L. Abdelmagid, I. Boiko, "Development of PSS tuning rules using multi-objective optimization," *Int. J. Electr. Power Energy Syst.*, 100: 449-462, 2018.

[11] M. Hassan, M.A. Abido, A. Aliyu, "Design of power system stabilizer using phase based objective function and heuristic algorithm," in *Proc. 2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO)*: 1-6, 2019.

[12] S.K. Wang, "A novel objective function and algorithm for optimal PSS parameter design in a multi-machine power system," *IEEE Trans. Power Syst.*, 28(1): 522-531, 2013.

[13] W.F. Sacco, C.R.E.D. Oliveira, "A new stochastic optimization algorithm based on a particle collision metaheuristic," presented at the 6th World Congresses of Structural and Multidisciplinary Optimization, Rio de Janeiro, Brazil, 2005.

[14] K. Padiyar, *Power System Dynamics Stability and Control*, BS publications, 2008.

[15] H. Beyranvand, E. Rokrok, M. Shakarami, "Matsim: A matpower and simulink based tool for power system dynamics course education," presented at the 31th power system conference (PSC), Tehran, Iran, 2016.

[16] R.D. Zimmerman, C.E. Murillo-Sanchez, R.J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, 26(1): 12-19, 2011.

[17] P.M. Anderson, A.A. Fouad, *Power system control and stability*, John Wiley & Sons, 2008.

Biographies



Peyman Naderi was born in Ahvaz, Iran, in 1975. He received his B.Sc. degree in Electronic Engineering in 1998 and M.Sc. degree in Power Engineering from Chamran University, Iran, Ahvaz in 2001. He has a Ph.D. degree in Power Engineering Science from K.N. Toosi University, Tehran, Iran. His interests are hybrid and electric vehicles, vehicles dynamic, power system transients and power system dynamics. He is currently associate professor in Shahid Rajaei Teacher Training University of Tehran, Iran.



Behzad Ehsan maleki was born on 6st May 1991 in Tehran, Iran. He obtained scholarship from the Ministry of Education, Iran in 2009 to peruse the BSc./B.Tech degree in Electrical Engineering. He received BSc./B.Tech and MSc./M.Tech degree in Electrical Engineering from Lorestan University and Shaheed Rajaei Teacher Training University in 2013 and 2016, respectively. His current research interests are Power System Stability and Flexible Alternating Current Transmission System (FACTS) devices.



Hamzeh Beiranvand was born on 21st March 1989 in Khorramabad, Iran. He obtained scholarship from the Ministry of Education, Iran in 2007 to pursue the BSc./B.Tech degree in Electrical Engineering. He received BSc./B.Tech and MSc./M.Tech degree in Electrical Engineering from Lorestan University in 2011 and 2014, respectively.

He started his Ph.D. in the year 2015 at Lorestan University and currently working on Solid-State Transformers (SSTs) in collaboration with chair of power electronics, Christian-Albrechts-Universität zu Kiel, Germany. His current research interests are Power Electronic Converters, Modular Multi-level Converters (MMCs), and SSTs.

Copyrights

©2020 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

P. Naderi, B. Ehsan Maleki, H. Beiranvand, "Optimal PSS parameters design based on a novel objective function for small signal stability enhancement," J. Electr. Comput. Eng. Innovations, 9(1): 83-92, 2021.

DOI: [10.22061/JECEI.2020.7154.361](https://doi.org/10.22061/JECEI.2020.7154.361)

URL: http://jecei.sru.ac.ir/article_1489.html





Research paper

Fast and Efficient Hardware Implementation of 2D Gabor Filter for a Biologically-Inspired Visual Processing Algorithm

A. Mohammadi Anbaran¹, P. Torkzadeh^{1,*}, R. Ebrahimpour^{2,4}, N. Bagheri^{3,5}

¹Electrical Engineering Department, Faculty of Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran.

²Artificial Intelligence Department, Faculty of Computer Engineering, Shahid Rajaee Teacher Training University; Tehran, Iran.

³Communication Engineering Department, Faculty of Electrical Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran.

⁴School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran.

⁵School of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran.

Article Info

Article History:

Received 24 May 2020
Reviewed 27 July 2020
Revised 22 September 2020
Accepted 18 November 2020

Keywords:

Gabor filter
FPGA
Separable filter
Convolution
HMAX model

*Corresponding Author's Email Address:

p-torkzadeh@srbiau.ac.ir

Abstract

Background and Objectives: Programmable logic devices, such as Field Programmable Gate Arrays, are well-suited for implementing biologically-inspired visual processing algorithms and among those algorithms is HMAX model. This model mimics the feedforward path of object recognition in the visual cortex.

Methods: HMAX includes several layers and its most computation intensive stage could be the S1 layer which applies 64 2D Gabor filters with various scales and orientations on the input image. A Gabor filter is the product of a Gaussian window and a sinusoid function. Using the separability property in the Gabor filter in the 0° and 90° directions and assuming the isotropic filter in the 45° and 135° directions, a 2D Gabor filter converts to two more efficient 1D filters.

Results: The current paper presents a novel hardware architecture for the S1 layer of the HMAX model, in which a 1D Gabor filter is utilized twice to create a 2D filter. Using the even or odd symmetry properties in the Gabor filter coefficients reduce the required number of multipliers by about 50%. The normalization value in every input image location is also calculated simultaneously. The implementation of this architecture on the Xilinx Virtex-6 family shows a 2.83ms delay for a 128×128 pixel input image that is a 1.86X-speedup relative to the last best implementation.

Conclusion: In this study, a hardware architecture is proposed to realize the S1 layer of the HMAX model. Using the property of separability and symmetry in filter coefficients saves significant resources, especially in DSP48 blocks.

Introduction

Feature extraction in object recognition is performed in the primary visual cortex (V1) in the mammalian visual pathway. The V1 simple cells behavior modeling

discovered by Hubel and Wiesel [1] is generally performed by 2 dimensional (2D) Gabor filters. A Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. A pyramid of 2D Gabor filters is a common approach for modeling classical simple cells (S1

layer) of HMAX (Hierarchical model and X) model [2], [3]. HMAX is a computational model of the ventral visual pathway found within the visual cortex, which is responsible for object recognition with an admissible performance. In S1 layer of the HMAX model, which is one of the computationally intensive stages of the model, the real part of the 2D Gabor filter is applied to the input image in four orientations and 16 scales. Applying different filters increases the robustness of the model for changes of object orientation and scale in the input image. This is performed in the S1 layer of the model, which is one of the computationally intensive stages of the model.

In addition to biologically-inspired systems, the 2D Gabor filter is employed in many other applications, such as texture classification, facial expression recognition techniques [4], edge detection [5], and iris recognition [6]. For this reason, speedup in applying a 2D Gabor filter to an image is particularly critical in real-time applications.

Many articles have been written in the last decade with the aim of acceleration of applying Gabor 2D filter to an input image. In [7], a technique has been proposed to integrate the interpolation and the convolution processes of the Gabor filter. The integration of these two processes makes the 2D Gabor filter separable along any direction.

Separating the filter in two dimensions, x and y , converts the 2D filter to two more efficient 1D filters with Gaussian and sinusoidal modulations. However, this method requires a technique for re-sampling an image by an interpolation kernel, a process which involves additional computational complexity.

To convert convolution to multiplication, some articles have used frequency domain and Fourier transforms. To obtain the final output, [8] first transforms input data and filters to Winograd or frequency domain, performs element-wise multiplication, and then applies inverse transformation. This study proposes a novel architecture for implementing fast algorithms on FPGAs.

Utilizing some mathematical techniques in mathematical relationships reduces computation overhead without any loss in accuracy. Some recent studies have computed filters as linear combinations of a smaller number of separable filters, thus greatly reducing computational complexity at no cost in terms of performance [9], [10].

In this study, a filter with rank R converts to an R separable filter or R filter with rank one. Separation in a 2D filter indicates that separation may be achieved by applying two or more 1D filters, which greatly reduces the run-time and computational resource requirements without a loss in accuracy. The idea of separating Gabor

filter kernels is also introduced in [11], but the suggested method only accommodates particularly oriented Gabor filters and so is not generic or flexible enough.

Reference [12] proposes GPU acceleration of the texture feature extraction algorithm by using separable 1D Gabor filters to approximate non-separable Gabor filter kernels.

The complexity and severity of computations in the Gabor filter's real-time applications, such as S1 layer of HMAX, has always posed a challenge in filter implementation. In recent years, researchers have been fascinated by an effective approach to overcoming this challenge: the design of hardware accelerators, which enable massive parallel processing and pipelining [13]-[15].

In addition, FPGA-based accelerators provide fast programming times and cost-effective ways for evaluating algorithms and prototyping, which eliminate fabrication time [16]. Furthermore, FPGA fabricators have developed IP Cores, such as CORDIC Cores, Memory Cores, and math calculation Cores, for instance routing and power, which make designing easier, faster, and more optimal.

A. Our Contributions

As the main contribution of this paper, we proposed a resource efficient version of the S1 layer of the HMAX model for hardware implementation. More precisely, the proposed model has the following properties:

- The 2D Gabor filter is separable in 0° and 90° directions, and in other directions the Gabor filters becomes separable by approximating them using its isotropic ($\gamma = 1$, circular) version. In order to reduce the complexity of computation, separation and symmetry properties have been used in the isotropic version of 2D Gabor filters, which is effectively raises the filtering speed.
- This method is employed in the design of hardware architecture for accelerating the S1 layer of the HMAX model. The effect of this approximation (i.e. $\gamma = 1$) on the accuracy of the HMAX model has been investigated in [17] and it has been shown that it has no effect on its accuracy.
- Pipeline hardware architecture for the proposed S1 layer was designed and implemented on one Virtex-6 FPGA family. By utilizing the separation and symmetry property, there is a desirable reduction in hardware resources, especially in DSP48E1 and memory blocks.
- The normalization value of the results is calculated in parallel with the filter to increase the illumination invariance.

B. Paper Organization

The organization of the present paper is as follows. The second section reviews the S1 layer of a HMAX

model of a classical simple cell emulating V1 simple cell. In the third Section discusses the mathematical relations of the separability and symmetric coefficient of a Gabor filter in special orientations. The next section presents the overall hardware architecture of the S1 layer of the HMAX model according to the contents of the previous section. The next section presents the simulations of the modified model.

The following section provides the results of the proposed architecture and a comparison with a state-of-the-art implementation. Finally, the conclusion of the work is presented.

S1 Layer of HMAX

The S1 layer is the lowest layer of the HMAX model and receives a gray value image as its input. This input image is then applied to a set of Gabor filters as an edge detector filter. The Gabor filters fit very well with the receptive field weight functions found in the simple cells of the primary visual cortex. The following equations describe the real part of the two-dimensional Gabor filter [3]:

$$G(x, y) = e^{-\frac{x^2+y^2\gamma^2}{2\sigma^2}} \cdot \cos\left(\frac{2\pi}{\lambda} X\right) \tag{1}$$

$$X = x \cos \theta + y \sin \theta$$

$$Y = -x \sin \theta + y \cos \theta$$

x and y are determined by the filter size, which spans a range of sizes from 7×7 to 37×37 pixels in the steps of the two pixels. Table 1 of [2] presents all filter parameters, i.e., the aspect ratio, $\gamma=0.3$, orientation θ (0°, 45°, 90°, and 135°), effective width σ , and wavelength λ . Thus, the complete pyramid consists of 4×16=64 filters, leading to 64 different S1 receptive field types (four orientations and 16 scales). Figure 1 shows the 2D Gabor filter bank.

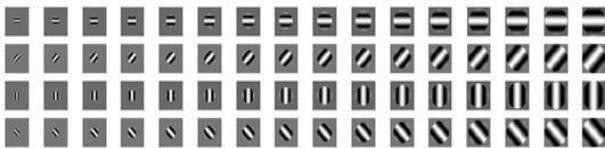


Fig. 1: The 2D Gabor filter bank with four orientations and 16 scales [17].

In the S1 layer of the HMAX model, the 64 Gabor filters should be applied to the input image and then the results should be normalized for illumination invariance. Normalization values are obtained by calculating the root sum square values of the input image pixels at each location where the Gabor filter is applied. Figure 2 provides the pseudo code for computing the S1 layer

response and Fig. 3 shows the result of the S1 layer in various orientations and scales of the Gabor filter.

```

For each scale, s = 7 : 2 : 37
  For each orientation,  $\Theta = 0 : 45 : 135$ 
    Extract sxs Gabor Filter in  $\Theta$  orientation from memory
    Computing the 2D convolution between the input Image and the Gabor Filter
    Computing the Normalization value of the Image in the sxs region
  end
end
result <= 2D convolution / Normalization value
    
```

Fig. 2: Pseudo code of the S1 layer.

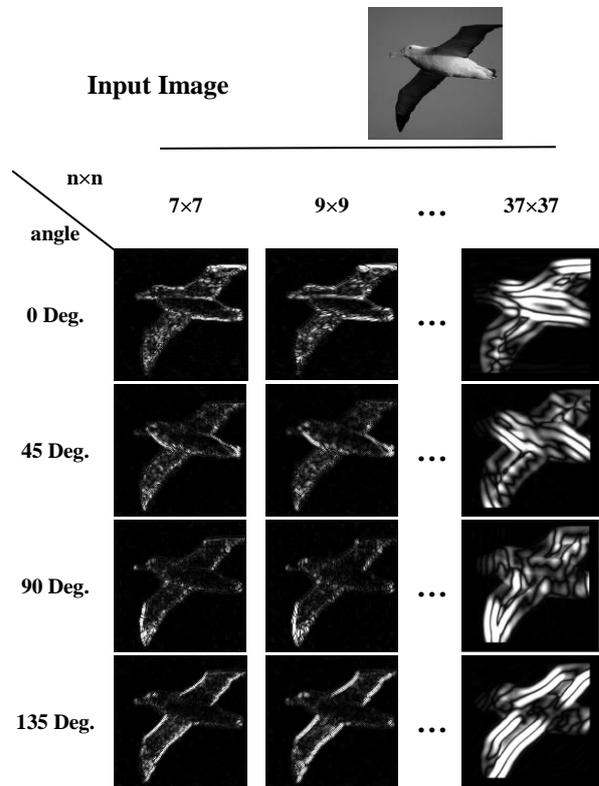


Fig. 3: Applying Gabor filters in 4 orientations and in 16 scales.

Mathematical Relations

The conventional way to apply the Gabor filters in the S1 layer is convolution in the spatial domain. The complexity of convolution depends directly on $M \times N$ and P , where P is the width and height of the filter and M and N are the width and height of the image, respectively.

The complexity of calculating the filter response for one location is P^2 and for the entire image is MNP^2 .

A filter is called separable if it can be expressed as the multiplication of two column and row vectors. Therefore, the convolution in the separable filters can be performed by two one-dimensional filters.

Consequently, the computational complexity decreases to 2PMN or P/2 times.

$$G(x, y) = G_1(x) * G_2(y)$$

$$\begin{aligned} Out(x, y) &= G(x, y) * In(x, y) \\ &= [G_1(x) * G_2(y)] * In(x, y) \\ &= G_1(x) * [G_2(y) * In(x, y)] \end{aligned} \quad (2)$$

where the operator * denote to the convolution, $G(x,y)$ is the 2D Gabor filter and G_1 and G_2 are 1D Gabor filters by column and row representation. $Out(x,y)$ and $In(x,y)$ are the input image and the output of the filter, respectively.

Solving (1) in 0° and 90° shows that the Gabor filter is separable in these orientations.

$$\begin{aligned} G_{\theta=0}(x, y) &= e^{-\frac{x^2+y^2y^2}{2\sigma^2}} \cdot \cos\left(\frac{2\pi}{\lambda}x\right) \\ &= e^{-\frac{y^2y^2}{2\sigma^2}} \cdot \left[e^{-\frac{x^2}{2\sigma^2}} \cos\left(\frac{2\pi}{\lambda}x\right) \right] \\ &= E_2^T(y) * E_1(x) \end{aligned} \quad (3)$$

$$\begin{aligned} G_{\theta=90}(x, y) &= e^{-\frac{y^2+y^2x^2}{2\sigma^2}} \cdot \cos\left(\frac{2\pi}{\lambda}y\right) \\ &= \left[e^{-\frac{y^2}{2\sigma^2}} \cos\left(\frac{2\pi}{\lambda}y\right) \right] \cdot e^{-\frac{y^2x^2}{2\sigma^2}} \\ &= E_1^T(y) * E_2(x) \end{aligned}$$

in which:

$$E_1(x) = e^{-\frac{x^2}{2\sigma^2}} \cos\left(\frac{2\pi}{\lambda}x\right) \quad (4)$$

$$E_2(y) = e^{-\frac{y^2y}{2\sigma^2}}$$

where $E_1^T(y)$, $E_2^T(y)$ and $E_1(x)$, $E_2(x)$ are column and row vectors, respectively, and the * sign designates the convolution of the two column and row vectors.

In any orientation θ other than 0° and 90° :

$$\begin{aligned} G_\theta(x, y) &= \\ &e^{-\frac{(x \cos \theta + y \sin \theta)^2 + y^2(-x \sin \theta + y \cos \theta)^2}{2\sigma^2}} \cos\left(\frac{2\pi}{\lambda}(x \cos \theta + \right. \\ &\left. y \sin \theta)\right) \end{aligned} \quad (5)$$

In order to reduce the complexity of computation and make the Gabor filters separable, we approximate them

using its isotropic ($\gamma=1$, circular) version. The effect of this approximation on the accuracy of the HMAX model is investigated in [17] and substituting $\gamma=1$ does not reduce accuracy. Therefore:

$$\begin{aligned} G_\theta(x, y) &= e^{-\frac{x^2+y^2}{2\sigma^2}} \left[\cos\left(\frac{2\pi}{\lambda}x \cos \theta\right) \cos\left(\frac{2\pi}{\lambda}y \sin \theta\right) \right. \\ &\quad \left. - \sin\left(\frac{2\pi}{\lambda}x \cos \theta\right) \sin\left(\frac{2\pi}{\lambda}y \sin \theta\right) \right] \\ &= e^{-\frac{x^2}{2\sigma^2}} \cos\left(\frac{2\pi}{\lambda}x \cos \theta\right) e^{-\frac{y^2}{2\sigma^2}} \cos\left(\frac{2\pi}{\lambda}y \sin \theta\right) \\ &\quad - e^{-\frac{x^2}{2\sigma^2}} \sin\left(\frac{2\pi}{\lambda}x \cos \theta\right) e^{-\frac{y^2}{2\sigma^2}} \sin\left(\frac{2\pi}{\lambda}y \sin \theta\right) \\ &= E_3^T(x) \cdot E_4^T(y) - O_1^T(x) \cdot O_2^T(y) \end{aligned} \quad (6)$$

in which:

$$\begin{aligned} E_3(x) &= e^{-\frac{x^2}{2\sigma^2}} \cos\left(\frac{2\pi}{\lambda}x \cos \theta\right) \\ E_4(y) &= e^{-\frac{y^2}{2\sigma^2}} \cos\left(\frac{2\pi}{\lambda}y \sin \theta\right) \\ O_1(x) &= e^{-\frac{x^2}{2\sigma^2}} \sin\left(\frac{2\pi}{\lambda}x \cos \theta\right) \\ O_2(y) &= e^{-\frac{y^2}{2\sigma^2}} \sin\left(\frac{2\pi}{\lambda}y \sin \theta\right) \end{aligned} \quad (7)$$

This signifies that the 2D Gabor filter is equal to the minus of two separable filters in each arbitrary orientation.

Especially in 45° and 135° orientations $E_3 = E_4$ and $O_1=O_2$ due to $\sin \theta = \cos \theta$:

$$\begin{aligned} G_{\theta=45}(x, y) &= \\ &= \left[e^{-\frac{x^2}{2\sigma^2}} \cos\left(\frac{\sqrt{2}\pi x}{\lambda}\right) \right] \cdot \left[e^{-\frac{y^2}{2\sigma^2}} \cos\left(\frac{\sqrt{2}\pi y}{\lambda}\right) \right] \\ &\quad - \left[e^{-\frac{x^2}{2\sigma^2}} \sin\left(\frac{\sqrt{2}\pi x}{\lambda}\right) \right] \cdot \left[e^{-\frac{y^2}{2\sigma^2}} \sin\left(\frac{\sqrt{2}\pi y}{\lambda}\right) \right] \end{aligned} \quad (8)$$

$$\begin{aligned} G_{\theta=135}(x, y) &= \\ &= \left[e^{-\frac{x^2}{2\sigma^2}} \cos\left(\frac{\sqrt{2}\pi x}{\lambda}\right) \right] \cdot \left[e^{-\frac{y^2}{2\sigma^2}} \cos\left(\frac{\sqrt{2}\pi y}{\lambda}\right) \right] \\ &\quad + \left[e^{-\frac{x^2}{2\sigma^2}} \sin\left(\frac{\sqrt{2}\pi x}{\lambda}\right) \right] \cdot \left[e^{-\frac{y^2}{2\sigma^2}} \sin\left(\frac{\sqrt{2}\pi y}{\lambda}\right) \right] \end{aligned} \quad (9)$$

Then:

$$G_{\theta=45}(x, y) = E_3^T(x) * E_3(y) - O_1^T(x) * O_1(y)$$

$$G_{\theta=135}(x, y) = E_3^T(x) * E_3(y) + O_1^T(x) * O_1(y)$$
(10)

in witch:

$$E_3(x) = e^{-\frac{x^2}{2\sigma^2}} \cos\left(\frac{\sqrt{2}\pi x}{\lambda}\right)$$

$$O_1(x) = e^{-\frac{x^2}{2\sigma^2}} \sin\left(\frac{\sqrt{2}\pi x}{\lambda}\right)$$
(11)

Therefore, in these cases, the 2D Gabor filter transforms to plus or minus of two separate 1D filter. In addition, each 1D filter can be halved by using the even or odd symmetry properties in the functions of (10). This will reduce the required number of multipliers by about 50%.

The use of separability and symmetry properties improves storage capacity of Gabor filter coefficients. For example, storing a 37×37 filter requires 1,699 memory locations. However, in the proposed design, $19+19=38$ memory locations are utilized. For a total of 64 filters in S1 layer of HMAX, instead of $\sum_{k=7,9,\dots,37} 4k^2 = 145,664$ memory locations, $\sum_{k=7,9,\dots,37} 4 * 2 * (k + 1)/2 = 5,888$ are utilized which k is the filters dimension. This represents a 95.8% reduction in the memory requirement for storing Gabor filter coefficients.

Hardware Architecture of the Accelerator

This section describes the hardware accelerator architecture of the HMAX model's S1 layer. This design is the result of an accurate study of the S1 layer structure and the arrangement of the steps involved in this layer. The design is part of the HMAX model and is ultimately utilized with this model.

A gray scale image as the accelerator input is stored in an internal memory. In addition, the Gabor filter coefficients are pre-calculated with the functions presented in the previous section and stored in the FPGA internal memory.

Figure 3 presents the block diagram of the entire S1 layer architecture. The design has two paths. In the first path, the input image data is applied to the 1D filter and the outputs are stored in an intermediate RAM. In the second path, instead of the input image, the stored values of the first path are applied to the 1D filter with a vertical sweep. Since the input image is required to apply other filters, the memory allocated for the input image cannot be used for intermediate values. Except for input

image RAM and normalization values calculator (the normalization value is the same for all filters), this architecture has been replicated four times (for orientations 0° , 45° , 90° and 135°). The results of applying 0° and 90° Gabor filter are directly divided by the normalization value, but according to (10), the results of applying E_3 and O_1 filters must be + and - for Gabor filters 135° and 45° , respectively. The normalization value is calculated with the processing of the filter, simultaneously.

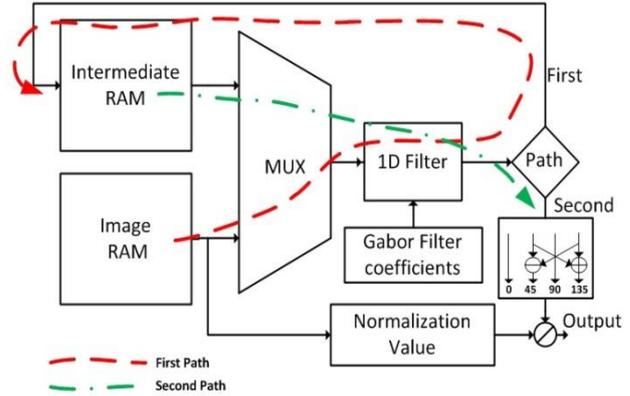


Fig. 4: The block diagram of the S1 layer architecture.

Figure 5 shows the architecture of a 1D filter. At each clock pulse, one pixel from a row of the input image is entered into a filter. Therefore, in the 37×37 filter, after 37 clock pulses, the data is available and the accelerator starts to work.

By exploiting the even or odd symmetry in the coefficients of the 1D filters, the pixel values of the input image are added or subtracted and then applied to the filter coefficients. Finally, the outputs of 19 multipliers, resulting of 37×37 filter, are added together using a pipelined version of the adder tree. Using multi-stage pipeline structure, the 1D filter output is prepared in one position of the input image in one clock pulse.

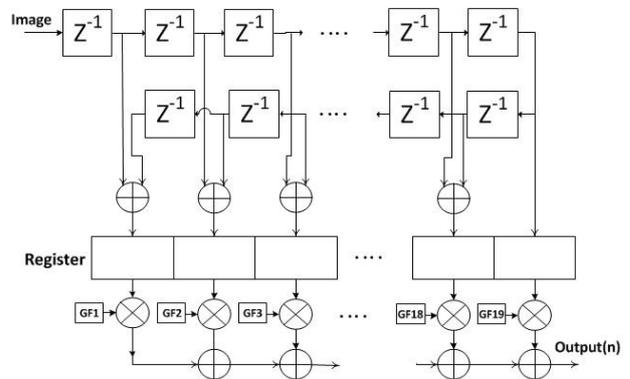


Fig. 5: Architecture of a 1D filter (with even symmetry). Z^{-1} represents a one-sample delay and GF is a Gabor Filter coefficient.

Figure 6 shows the architecture of the normalization values calculator. Initially, the image values are squared and entered into a shift register. A pipeline adder collects the register values and the results of the first path are stored in an intermediate RAM.

In the second path, the stored values are vertically swept and entered into the shift register. The square root of the results are used as normalization value as shown in Fig. 5.

The “Image RAM” shown in Fig. 6 and Fig. 4 is the same.

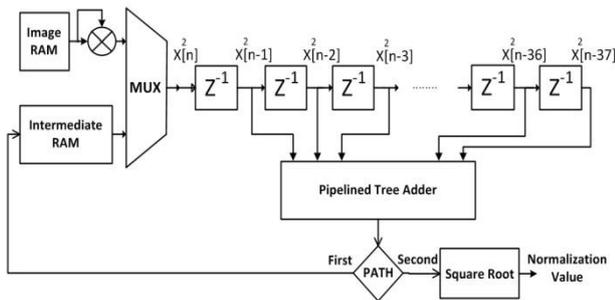


Fig. 6: Architecture of the normalization values calculator.

With the employment of the Xilinx IP cores available in ISE Design Suite software, the design of the circuit is facilitated and its square root, division, multiplication, and RAM components become more efficient. In order to more parallel processing in the hardware, four 1D filters with orientations of 0°, 45°, 90° and 135° have been instantiated to generate the results at four angles simultaneously.

Simulation

The present study investigates three different simulations on various images from well-known datasets, such as Caltech 101 [18]. The first one is the S1 layer from the standard HMAX model written by Serre et al. [2], whose source code is available online. This is a simulation in MATLAB and serves as a basis for comparing our extension code results.

The second simulation is the modified code in MATLAB, which converts a 2D filter to two 1D filter by use of the separability and symmetry properties of isotropic Gabor filters ($\gamma=1$). As expected, the result of this simulation is exactly equivalent to the first simulation or the standard model, due to the mathematical proof given in the third section.

With the usage of ModelSim software, the VHDL code of the designed architecture is simulated as the third simulation.

This simulation tests the functional operation and timing characteristics of the circuit (Fig. 7). The functional simulation is useful for checking the

fundamental correctness of the designed circuit and the accuracy of the results due to the bit width limitations.

Hardware implementation of circuit with Xilinx ISE determines the timing and resource consumption. The timing simulation evaluates the speed performance by considering the latency of wires and logic components and tests the mathematical operation of the circuit. The target platform for simulation and synthesis is a mid-range commercial Xilinx FPGA, i.e. XC6VLX240T of the Virtex-6 family.

Implementation Results and Comparisons

By considering accuracy, efficiency, as well as resource and power consumption, this section evaluates the accelerator proposed. Implementation is performed on a Xilinx XC6VLX240T device, which hosts approximately 150,000 LUTs, a 14.5 Mb built-in RAM, and 768 DSP48. The design is written in VHDL language and synthesized and implemented with ISE 14.7 software.

The input image is assumed to be 128×128 pixels and converts to a gray scale with an integer number between 0 and 255 (8 bits) assigned to every pixel. The input image is stored in an internal memory. The output is expressed with four 16-bit integer numbers in conjunction with the 14-bit address of the image pixel and a signal for output validation (Fig. 7). The output is not stored in the internal memory due to its consumption of excessive storage space.

For each of the Gabor filter coefficients, a 16-bit fixed point number was assigned; 8 bits for integer part and 8 bits for fractional part.

Fixed point expression was selected since this expression uses fewer physical resources and fewer clock cycles than do floating point numbers. At the end of the first and second path, 8 bits from the low significant of the result numbers are removed. So the number stored in the intermediate RAM and the result of the second path will be integer. The bit width of intermediate values is adjusted such that it never overflows in the worst-case scenario.

Multiplication, division, RAM, and square root components are generated by Core Generator software. A state machine controls the entire circuit via multiple counters for counting the rows and columns of the input image.

Implementation of the total circuit indicates that the design can run at a minimum period of 5.38nsec or a maximum frequency of 185 MHz.

Since the processing of a single line of input images with the pipeline structure takes about N clock cycles, processing the first path of filter in four orientations requires about MN clocks, where M and N are the width and height of the input image, respectively.

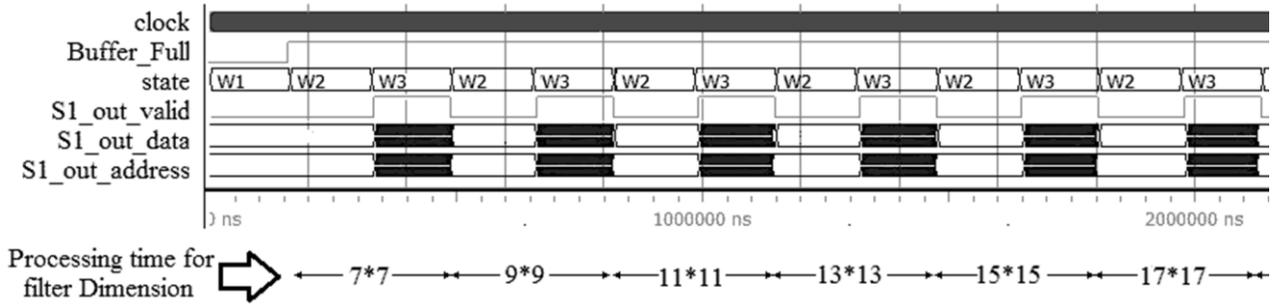


Fig. 7: Simulation of the S1 layer with ModelSim (states: W1: load image, W2: first path of filter, W3: second path of filter).

Therefore, process of first and second path of the filter and produce the final result take about 2MN clocks. Thus applying 16 filters take about:

$$\text{Del} = 16 \times 2MN \times 5.38 \text{ nsec} \quad (12)$$

Del is the total delay of our hardware for each input image with $M \times N$ pixels. This time is 2.82msec for 128×128 pixel image. Simulation of the circuit shows 2.84msec delays, resulting in a maximum throughput of 354 frames per second (fps). Consequently, the average processing time for one filter is $2.84\text{ms}/64 = 44\mu\text{s}$. Obviously, this circuit is one layer of the HMAX model and its throughput should be comparable by the throughput of the rest of the circuit. In this calculation, the loading time of the input image is eliminated due to the pipeline structure of the circuit.

Table 1 presents the total hardware resources consumed by the implemented circuit on a XC6VLX240T device. The limiting factor in further parallel processing is the amount of memory available. Table 1 is for a 128×128 pixel image. The major impact of increasing the input image's size is on the RAM used as it is directly related to the image's dimensions. Therefore, if the image is enlarged to 256×256 , then the memory consumption is quadrupled. Regardless of the amount of memory required, one of the advantages of the proposed architecture is that increasing the size of the input image has no effect on the number of multipliers required. Number of multipliers is affected by the largest filter dimension or 37. It should be noted that not utilizing the separability and symmetry property, we will require $37 \times 37 \times 4 = 5,476$ multiplications.

The estimated dynamic power consumption is 1.02W and the quiescent power consumption is 2.97W due to leakage by use of Xilinx XPower Analyzer software.

For comparison of circuit speeds, Table 2 provides three different implementations of the HMAX model's S1 layer released in recent years. In addition to the standard HMAX model which is discussed in the present paper, there is an extension model developed by Mutch et al. [3]. In this model, Gabor filters with same sizes

(11×11) are used for all scales, applying them to scaled versions of the image. The advantage of this model is in reducing computational complexity, so it has been used in many implementations in recent years.

Table 1: Hardware resource usage

Resource	Used	Available	Percent
DSP48E1	77	768	10%
RAMB36E1	57	416	13%
RAMB18E1	5	832	1%
Slice Reg.	15,598	301,440	5%
LUTs	9,619	150,720	6%

Reference [19] implements the whole Mutch HMAX model in the embedded Power PC, applies four 11×11 Gabor filter to 10 images pyramid in the S1 layer, and reports a 511ms delay. The hardware implementation of this study achieved a maximum delay of 260msec. In the hardware proposed in [14], implements the application of an 11×11 Gabor filter, in four directions, to a 12 images pyramid, from 256×256 to 38×38 image dimensions, and reports a hardware delay of 56.3 msec. In an article similar to the current work, the S1 layer of the standard HMAX model with a 128×128 image is implemented on programmable hardware, achieving a favorable Throughput of 190 fps. with a separable Gabor filter [13].

In a comparable work, two 7×7 and 9×9 Gabor filters are applied to an input image in four directions [20]. As reported, 75 and 143 fps could be processed in their FPGA and std_cell implementations for full-HD (1920×1080) images, respectively. Even though the image dimensions are about 126 times larger than those of the present study's implementation, only the two smallest filters, out of the 16, are used in this implementation. In the present work, there are 16 scales filter from 7×7 to 37×37 which include 36,416 multiply accumulate while in 7×7 and 9×9 filters 130 multiply accumulate are required.

Table 2: Speed comparisons among various implementations

	Image Size (pixel)	Filter Size	Throughput (fps)	Description
Debole [19]	140×140 ~ 30×30 (10*)	11×11	3.85	Mutch HMAX model, 11 orientation
Masshri [14]	256×256 ~ 38×38 (12*)	11×11	17.8	Mutch HMAX model
Orchard [13]	128×128	7×7 ~ 37×37 (16*)	190	standard HMAX
Licciardo [20]	1920×1080	7×7 and 9×9	75/143	FPGA/Std_cell implementation
This paper	128×128	7×7 ~ 37×37 (16*)	353	standard HMAX

* Number of images in the pyramid

Therefore, the complexity ratio in these two works in filters will be $36,416/130=280$, which is much larger than image dimensions complexity ratio (126). By other calculation, according to [12], delay of our circuit for 1920×1080 pixel images is about 357msec. Since the difference in complexity of the filters used in the two implementations is 280 times, therefore, it can be concluded that if our circuit hardware resources are used for [20] implementation, the circuit delay will be $357/280=1.34$ msec which is 9.9X speedup.

Comparisons of [14], [19], [20] are unfair due to their differences in, for example, the model, filter size, input image size, and implementation platform. Despite this, Table 2 presents these implementations since they are the most recently designed for increasing the response speed.

The design that is most resembling to that of present article is [13], which is an architecture about 1.86X slower than the current work's on the identical hardware platform.

Conclusion

In this study, we proposed a new method for applying an isotropic 2D Gabor filter to an input image. With the conversion of one 2D filter to two 1D filters (the separability property), the computational complexity is reduced. Besides, to determine the efficiency and resource requirements of the proposed method, pipeline structure of the S1 layer of the HMAX model has been implemented, which consists of 64 2D Gabor filter in different scales and orientation as well as normalization value calculator circuit.

The architecture implemented on a fast and mid-range commercial FPGA platform, i.e. XC6VLX240T, for which the achieved throughput was 353 fps. Using this method saves FPGA resources, specially the number of DSP48E1 and RAM blocks, as well as reducing hardware delays.

The hardware design of the S1 layer of the standard HMAX model is performed with the assistance of Xilinx IP cores.

The proposed S1 layer architecture accelerates by more than 1.86X compared to the most similar work.

Using two one-dimensional filters and using a pipeline structure between the two filters can be introduced as a future work.

Author Contributions

A. Mohammadi Anbaran and N. Bagheri carried out the hardware design, simulation and implementation of the circuit and wrote the manuscript. P. Torkzadeh and R. Ebrahimpour carried out the data analysis and edit the manuscript. A. Mohammadi Anbaran collected the data and interpreted the results.

Acknowledgment

We thank the editor and all anonymous reviewers.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

<i>HMAX</i>	Hierarchical model and X
<i>FPGA</i>	Field Programmable Gate Arrays
<i>2D</i>	2 Dimension
<i>GF</i>	Gabor Filter
<i>CORDIC</i>	COordinate Rotation Digital Computer

References

- [1] D.H. Hubel, T.N. Wiesel, "Receptive fields of single neurones in the cat's striate cortex," *J. Physiol.*, 148(3): 574-591, 1959.
- [2] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, T. Poggio, "Robust object recognition with cortex-like mechanisms," *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(3): 411-426, 2007.
- [3] J. Mutch, D.G. Lowe, "Multiclass object recognition with sparse, localized features," in *Proc. Computer Vision and Pattern Recognition conf. (CVPR'06)*: 11-18, 2006.

[4] S. Rajan, P. Chenniappan, S. Devaraj, N. Madian, "Facial expression recognition techniques: a comprehensive survey," *IET Image Proc.*, 13(7): 1031-1040, 2019.

[5] R. Mehrotra, K.R. Namuduri, N. Ranganathan, "Gabor filter-based edge detection," *Pattern Recognit.*, 25(12): 1479-1494, 1992.

[6] S. Liu, Y. Liu, X. Zhu, Z. Liu, G. Huo, T. Ding, et al., "Gabor filtering and adaptive optimization neural network for iris double recognition," in *proc. Chinese Conf. on Biometric Recognition: 441-449*, 2018.

[7] G. Amayeh, A. Tavakkoli, G. Bebis, "Accurate and efficient computation of gabor features in real-time applications," in *proc. Int. Symposium on Visual Computing: 243-252*, 2009.

[8] Y. Liang, L. Lu, Q. Xiao, S. Yan, "Evaluating fast algorithms for convolutional neural networks on FPGAs," *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.*, 39(4): 857-870, 2020.

[9] A. Sironi, B. Tekin, R. Rigamonti, V. Lepetit, P. Fua, "Learning separable filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, 37(1): 94-106, 2014.

[10] J. Khosravi, M. Shams Esfandabadi, R. Ebrahimpour, "Image registration based on sum of square difference cost function," *J. Electr. Comput. Eng. Innovations*, 6(2): 263-271, 2018.

[11] V. Areekul, U. Watchareeruetai, S. Tantaratana, "Fast separable gabor filter for fingerprint enhancement," in *Proc. Int. Conf. on Biometric Authentication: 403-409*, 2004.

[12] W.-M. Pang, K.-S. Choi, J. Qin, "Fast Gabor texture feature extraction with separable filters using GPU," *J. Real-Time Image Proc.*, 1(12): 5-13, 2016.

[13] G. Orchard, J.G. Martin, R.J. Vogelstein, R. Etienne-Cummings, "Fast neuromimetic object recognition using FPGA outperforms GPU implementations," *IEEE Trans. Neural Networks Learn. Syst.*, 24(8): 1239-1252, 2013.

[14] A. Al Maashri, M. Cotter, N. Chandramoorthy, M. DeBole, C.-L. Yu, V. Narayanan, et al., "Hardware acceleration for neuromorphic vision algorithms," *J. Signal Process. Syst.*, 70(2): 163-175, 2013.

[15] S. Moini, B. Alizadeh, M. Emad, R. Ebrahimpour, "A resource-limited hardware accelerator for convolutional neural networks in embedded vision applications," *IEEE Trans. Circuits Syst. II Express Briefs*, 64(10): 1217-1221, 2017.

[16] F. Abdi, P. Amiri, M.H. Refan, "Low computational complexity and high computational speed in leading DCD ERLS algorithm," *J. Electr. Comput. Eng. Innovations*, 7(1): 19-26, 2019.

[17] S. Chikkerur, T. Poggio, "Approximations in the hmax model," *Computer Science and Artificial Intelligence Laboratory, Tech. Rep. MIT-CSAIL-TR-2011-021*, 2011.

[18] L. Fei-Fei, R. Fergus, P. Perona, "Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories," *Comput. Vision Image Understanding*, 106(1): 59-70, 2007.

[19] M. DeBole, Y. Xiao, C.-L. Yu, A. Al Maashri, M. Cotter, C. Chakrabarti, et al., "FPGA-accelerator system for computing biologically inspired feature extraction models," in *proc. 45th Asilomar Conference on Signals, Systems and Computers (ASIOMAR): 751-755*, 2011.

[20] G.D. Licciardo, C. Cappetta, L. Di Benedetto, "Design of a Gabor filter HW accelerator for applications in medical imaging," *IEEE Trans. Compon. Packag. Manuf. Technol.*, 8(7): 1187-1194, 2018.

Biographies



Alireza Mohammadi Anbaran received a B.Sc. and a M.Sc. degree in electronic engineering from Shahrood University of Technology and Iran University of Science and Technology, in 2000 and 2003, respectively. He is currently pursuing the Ph.D. degree in the Faculty of Electrical and Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran. His research interest includes digital circuit implementation, object recognition, digital signal processing and ASIP design.



Pooya Torkzadeh was born in Isfahan, in 1980. He received the B.Sc. degree from Isfahan University of Technology (IUT), Iran, in 2002 and the M.Sc. and Ph.D. degrees from the Sharif University of Technology (SUT), Iran, in 2004 and 2011 respectively both in electrical engineering. He is currently faculty member of Islamic Azad University, Science and Research Branch. Since 2003 he has joint to Sharif Integrated Circuit

and System Group (SICAS) working on continuous/discrete time sigma-delta modulators with low power consumption for low power appliances. He has received many patents in the field of sigma-delta modulator designing and optimizing. He is the author and coauthor of more than 25 published international journal and conference papers on analog and digital integrated circuits. His research interests include ADC signal converters, low power phase locked loops (PLLs) with ultra-low phase noise amount for broad-band applications.



Reza Ebrahimpour is a professor at the department of computer engineering, Shahid Rajaei Teacher Training University, Tehran, Iran. He obtained Ph.D. degree in the field of Cognitive Neuroscience from the SCS, IPM, Tehran, Iran in July 2007. Dr. Ebrahimpour is the author or co-author of more than 100 international journal and conference publications in his research areas, which include Cognitive and Systems

Neuroscience, Human and Machine Vision, Decision Making and Object Recognition.



Nasour Bagheri is an associate professor at the Electrical Engineering Department, Shahid Rajaei Teacher Training University, Tehran, Iran. His research interests include cryptology, digital designing and hardware implementation. A record of his publication is available at Google Scholar: <https://scholar.google.com/citations?userD32llx44AAAAJ&hlDen>.

Copyrights

©2020 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

A. Mohammadi Anbaran, P. Torkzadeh, R. Ebrahimpour, N. Bagher, "Fast and efficient hardware implementation of 2D Gabor filter for a biologically-inspired visual processing algorithm," J. Electr. Comput. Eng. Innovations, 9(1): 93-102, 2021.

DOI: [10.22061/JECEI.2020.7548.404](https://doi.org/10.22061/JECEI.2020.7548.404)

URL: http://jecei.sru.ac.ir/article_1490.html





Research paper

Modelling and Optimization of Channel Allocation for Power Line Communications Access Networks in the Presence of In-Line and In-Space Interference

M. Sheikh-Hosseini^{1,*}, S.M. Nosratabadi²

¹Department of Computer and Information Technology, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran.

²Department of Electrical Engineering, Sirjan University of Technology, Sirjan, Iran.

Article Info

Article History:

Received 26 April 2020

Reviewed 27 June 2020

Revised 22 July 2020

Accepted 23 November 2020

Keywords:

Broadband power line communications (B-PLC) access network

PLC channel allocation (PLC-CA) problem

In-line and in-space neighboring schemes

Co-site interference

Genetic and Shuffled frog-leaping algorithms

*Corresponding Author's Email Address:

m.sheikhhosseini@kgut.ac.ir

Abstract

Background and Objectives: Broadband power line communications (PLC) is a promising candidate for implementing access network of different telecommunication Technologies. Planning process of the PLC access network is subdivided into two main optimization problems of the generalized base station placement and PLC channel allocation.

Methods: This paper studies the latter one for an actual PLC network by taking both in-line and in-space neighboring schemes into account for the first time and modeling the PLC channel allocation according to them. In this regards, different aspects of this problem are first introduced in details and then our suggested models for them are presented and numerically evaluated.

Results: Specifically, for each pair of the broadband-PLC cells, in-line neighboring is modeled either by one or zero indicating the cells are neighbor or not; in-space neighboring is suggested to be a number from the interval [0 1] according to physical vicinity of cell's wirings; and consequently aggregate neighboring intensity will be a number from [0 2]. Subsequently, the network interference is defined as a function of neighboring intensity and assigned frequency sets to the neighbor cells; so that the more neighboring intensity is increased and the more distance between the sets is decreased, the more interference is imposed on the PLC network. Eventually, the meta-heuristic methods of Genetic and shuffled frog-leaping algorithms are exploited to solve resulting PLC channel allocation problem via minimizing the interference.

Conclusion: In general, the results confirmed the success of the suggested method in modeling PLC channel allocation problem in actual scenarios, tracking the network interference in these situations, providing an optimal solution for them, and including all previous research as a comprehensive method.

Introduction

Power Line Communications (PLC) technology utilizes existing electrical infrastructure to provide and deliver various telecommunications services. History of this

technology dates back to the 1900's, and its fast and low-cost deployment together with the wide spread availability of power grids has been the most promoting factors encouraging power companies to develop PLC

technology over the years [1]-[3]. Compared with other wireline technologies like fiber optic and coaxial cables, PLC has lower deployment cost, fast installation time and much more pervasive structure [4], [5]. Furthermore, since PLC utilizing already implemented infrastructure, its deployment cost is comparable with that one of wireless technologies [4], [5]. Moreover, to increase coverage and enable signal propagation through concrete walls in different floors, ultra-wideband PLC has been suggested as an alternative for the existing in-building wireless solutions [6], [7].

Different PLC technologies can be classified into three main categories of ultra-narrowband PLC (UN-PLC), narrowband-PLC (N-PLC) and broadband-PLC (B-PLC), where the ultra-narrowband and narrowband ones use frequency bands below 500 kHz and support data rates between a few hundreds of bits per second (b/s) up to a few Mb/s. However, the operation frequencies of B-PLC systems are more than 1 MHz and may reach to 250 MHz. B-PLC systems provide data rate ranging from several Mb/s to several hundred Mb/s [4], [5], [8]. PLC technology has wide application area including from management and monitoring tasks for power utilities, smart grid applications, implementing access network of different telecommunication systems as a last mile solution, to smart homing, in-home and in-building networking, and in-vehicle communications [9]-[11].

In the last two decades, PLC have also received increasing attention as a promising solutions for telecommunication infrastructure needed by smart grid and a new subcategory, called high data rate N-PLC, has been emerged and standardized by the PRIME, G3-PLC, IEEE and ITU-T alliances for applications such as automatic meter reading, vehicle-to-grid communications and networking of consumer appliances [12], [13]. Furthermore, it is worth to be noted that the integration of the PLC with wireless technologies such as radio frequency and visible light communications systems have also received great attention and explored to enhance performance metrics like coverage and reliability [14], [15].

However, power lines are very harsh medium and suffer from some constraints such as deep attenuation, multipath propagation and impulsive noise [16], [17]. To overcome them, orthogonal frequency division multiplexing (OFDM) has been received increasing interest as an anti-multipath technique [18] and adapted to the B-PLC and high data rate N-PLC standards [17].

As outlined earlier, B-PLC can be exploited to implement the last mile (or access network) of various telecommunication technologies. In a B-PLC access network, the low voltage grid is subdivided into several cells and these cells are equipped with PLC base stations (BS), modems and repeaters to establish connection of

telecommunications backbone network with end-user subscribers via an OFDM-based system. Typical structure of a B-PLC access network is conceptually illustrated in Fig. 1, in which the cell's areas are distinguished by dotted lines and the BS terminal of each cell is located in the low voltage transformer station or a street cabinet. A more detailed structure of a typical B-PLC cell is shown in Fig. 2, where the BS terminal will connect the telecommunications backbone network to low voltage grid through an access point. At low voltage side, PLC modems over power outlets are utilized to connect the end-user devices to the BS terminal and the backbone network. In addition, several repeaters may be used in each cell to provide BS coverage for all users. As a result, this design enables injection of telecommunication signals into the low voltage grid and extraction of them from the grid.

Planning process of B-PLC access network can be subdivided into two different optimization problems: the generalized base station placement (GBSP) problem and the PLC channel allocation (PLC-CA) problem [19]. The main task of former one is to optimize network cost and quality of service by determining the optimum number of cells for serving all users, the most appropriate locations of the BS terminals, the set of required repeaters and subset of users must be served by each cell or BS terminal. When GBSP problem is solved, the PLC-CA problem will rise to optimally share the PLC resources such as available OFDM frequencies between different B-PLC cells. Expressing with more details, in order to increase network capacity, the frequency reuse technique is used to assign bandwidth to different cells and this results in assignment of same or near frequencies to neighbor cells and imposing co-site interference on the B-PLC network. PLC-CA solution must manage this challenge by providing a channel allocation which keeps this interference at authorized level.

A. Related Works and Our Motivations

As outlined, design of the B-PLC access network has been first addressed in [19] by converting this process into problems of the GBPS and PLC-CA. This work was paved the way for subsequent research on these problems. In the case of the GBPS problem, this challenge has been modeled as an optimization problem in [20]-[22] and solved by minimizing the network costs and delay. The challenge of PLC-CA is first addressed in [23], [24], and modeled as an optimization problem. Subsequently, the authors have generalized their works in [23], [24] and investigating PLC-CA problem more precisely in [25]. Furthermore, it is worth to be noted that dynamic resource allocation, by assuming dynamic structure for the B-PLC cells, has been developed to the B-PLC access networks in [26].

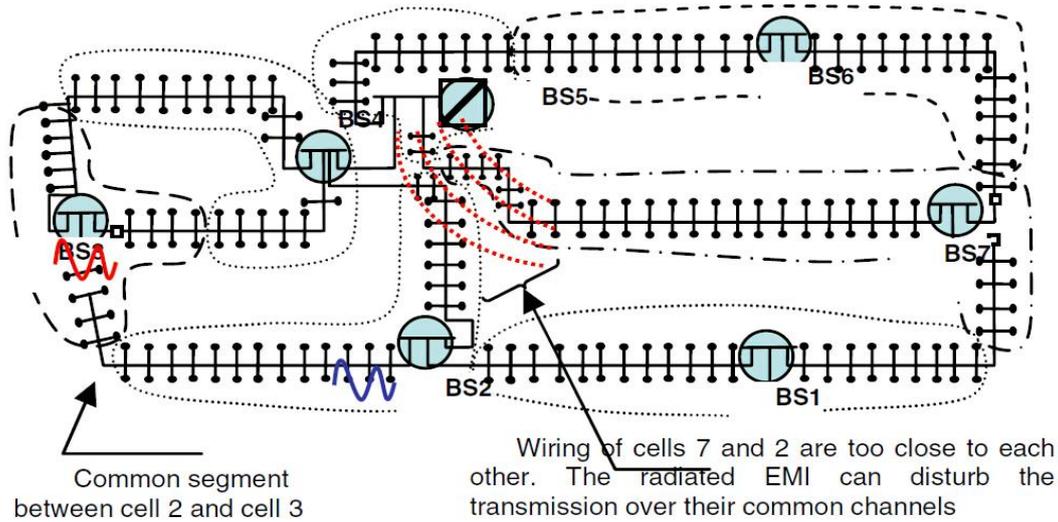


Fig. 1: Typical structure of a B-PLC access network [23].

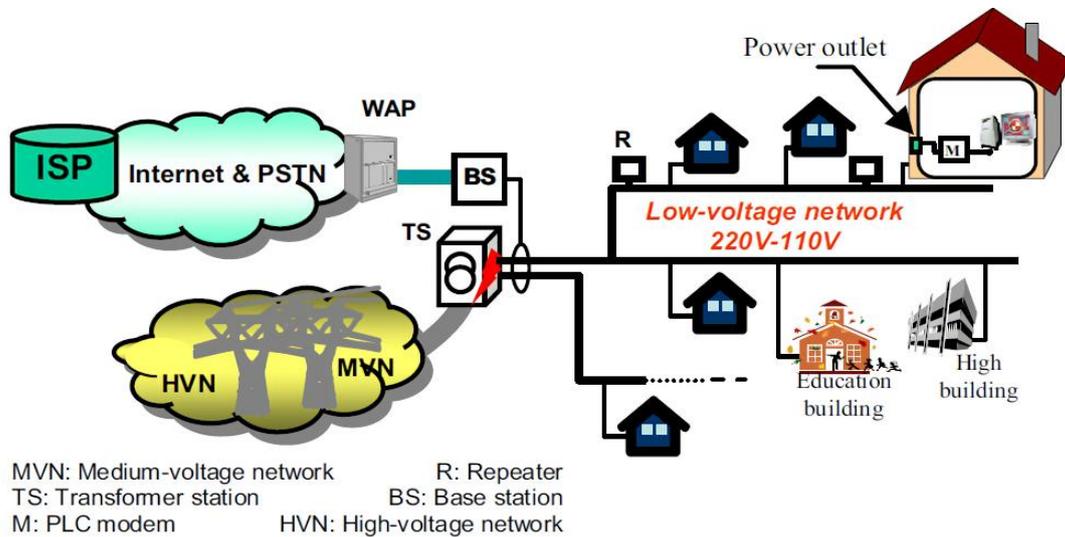


Fig. 2: Typical structure of a B-PLC cell [20].

However, similar to that one of [23]-[25], we investigate the channel allocation for the fixed B-PLC network and dynamic resource allocation falls beyond its scope.

Unfortunately, the provided analysis in [23]-[25] suffers from some limitations, since these works directly applied existing wireless solutions to the B-PLC networks without considering the differences between them. Expressing with more details, whereas there exist some similarities between B-PLC and wireless access networks, but the physical medium, size, geometrical shape and neighboring schemes of the cells are different in wireless and PLC networks.

In fact, PLC uses a physical medium with a specific wiring topology, and the PLC cells (as demonstrated in Fig. 2) are not similar to that one's of the wireless networks, i.e. hexagonal cells in free space. More specifically, in contrast to wireless, there exist two

different neighboring schemes for the B-PLC access networks. These are in-line and in-space neighboring. In-line neighboring is due to existence of common segments or wirings between two cells and generates conducted disturbances in the form of electrical signal. However, in-space neighboring is due to physical vicinity of wirings and generates disturbances in the term of electromagnetic signals. For instance, as demonstrated in Fig. 1, the cells 3 and 7 are in-line and in-space neighbors of the cell 2, respectively. Thereby, the channel and resource allocation in wireless scenarios (like [27] and [28]) is different from that one of the B-PLC networks and these are two distinct problems needing different models, analyses and solutions.

However, these differences have not taken into account in [23]-[25] and they have considered the B-PLC network similar to a wireless environment, which is only in expose of the in-line neighboring scheme.

Consequently, the PLC-CA have been modeled and analyzed in this regard and the impact of co-site interference imposed by in-space neighboring scheme has been ignored. This is while in-space neighboring is the main distinguishing feature of the PLC access and wireless networks. Therefore, it is necessary and vital to investigate and analyze the PLC-CA problem for the actual B-PLC networks by taking these differences into account. The lacks of this analysis has motivated us to devote this research on investigation of the PLC-CA problem over actual networks by taking both in-line and in-space neighboring schemes into account.

B. Our Contributions and Paper Organization

In order to overcome the above-mentioned limitations, this paper is devoted to investigating PLC-CA problem for the actual B-PLC access network by taking both in-line and in-space neighboring schemes into account and allocating channels accordingly. In this regard, in-line neighboring is modeled either by one or zero similar to [23]-[25]; however, in-space neighboring is considered as a real number from the interval [0 1] according to physical vicinity of cell's wirings; and the aggregate neighboring intensity is assumed as the summation of these two numbers. Then, the co-site interference of the B-PLC access networks is defined as a function of the aggregate neighboring intensity and the assigned frequency sets to the cells, so that the interference is increased when the value of neighboring intensity is increased and the distance of assigned frequency sets is decreased. Finally, the resulting co-site interference is minimized using genetic algorithm (GA) and shuffled frog-leaping algorithm (SFLA), and the assigned channels to cells are extracted. In general, the suggested method has advantage over the existing studies in terms of considering actual PLC network, providing more suitable and accurate model for the co-site interference, and evaluating the performance using two evolutionary methods and various realizations. However, since providing full details requires the problem model and its formulation and parameters, our contributions and novelties and their modifications over existing works have been postponed until the second part of the next section and highlighted in details at the beginning of this part.

The rest of paper is organized as follows: In Section II, we present the main results of this study by modeling PLC-CA problem as an optimization problem and suggesting our models for neighboring schemes and co-site interference. Optimum performance of these models is numerically evaluated and verified in Section III. Finally, Section VI includes conclusion.

Notations: Throughout this paper, matrices are defined by boldface letters like \mathbf{M} , where m_{ij} indicates its (i, j) element. The notations \mathcal{A} are used to define

sets, where $|\mathcal{A}|$ indicates the set cardinality, i.e. the number of elements in the sets.

Main Results

In this section, the PLC-CA problem is first formalized as an optimization problem and then the suggested models for different aspects of this problem are presented.

C. Modeling PLC-CA Problem as an Optimization Problem

Generally speaking, planning process of B-PLC access network includes dividing the low voltage power grid into several cells; equipping each cell with a BS terminal, PLC modems and repeaters; determining the subset of users have to be served by each cell; and assigning OFDM channels (frequencies) needed by each cell to serve its users.

In this paper, available B-PLC cells and OFDM frequencies are denoted by the sets $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_{|\mathcal{C}|}\}$ and $\mathcal{F}_s = \{\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_s\}$ respectively, where $|\mathcal{C}|$ and $|\mathcal{F}_s|$ are the cardinality of these sets. The i^{th} cell, i.e. \mathcal{C}_i , is modeled by $\mathcal{C}_i = \{B_i, \mathcal{U}_s^i, \mathcal{R}_s^i, \mathcal{F}_{CA}^i\}$, where B_i indicates to BS terminal and its access point, \mathcal{R}_s^i determines its repeaters set and \mathcal{U}_s^i represents the subset of users have to be served by this cell. These three quantities, i.e. B_i , \mathcal{R}_s^i and \mathcal{U}_s^i , are determined by solving the GBSP problem and the task of PLC-CA is to discover \mathcal{F}_{CA}^i , where \mathcal{F}_{CA}^i represents the set of assigned frequencies to cell \mathcal{C}_i .

In general, any solution of the PLC-CA problem can be represented using a binary matrix of size $|\mathcal{C}| \times |\mathcal{F}_s|$. Throughout this paper, we have named this matrix by \mathbf{M}^{CA} , where $m_{ij}^{CA} = 1$ means that the j^{th} frequency of the \mathcal{F}_s has been assigned to the cell \mathcal{C}_i and $m_{ij}^{CA} = 0$ means this frequency has not allocated to this cell. Note that the index values of non-zero elements in i^{th} row of \mathbf{M}^{CA} are the members of assigned frequency set to cell \mathcal{C}_i , i.e. \mathcal{F}_{CA}^i .

The requirements and constraints must be taken into account for solving PLC-CA problem are as below:

- First, due to some technical and regulatory restrictions, cell \mathcal{C}_i may has a forbidden frequency set as \mathcal{F}_{Forbid}^i . Thus, its allowable frequency set is $\mathcal{F}^i = \mathcal{F}_s - \mathcal{F}_{Forbid}^i$, and the constraint of $\mathcal{F}_{CA}^i \subseteq \mathcal{F}^i$ must be met in solving PLC-CA problem.
- To guarantee serving of all users, the number of assigned frequencies to cell \mathcal{C}_i must be as $|\mathcal{F}_{Req}^i| = A_{User} |\mathcal{U}_s^i|$, where $|\mathcal{U}_s^i|$ is the number of allocated users to this cell and A_{User} is a constant number representing the average number of frequencies needed by each user to provide the traffic requested in its downlink and uplink channels. This imposes another constraint to the PLC-CA problem as $|\mathcal{F}_{CA}^i| \geq$

$$|F_{Req}^i|.$$

- Moreover, to optimally solve PLC-CA problem, it is necessary to minimize the co-site interference. Since the in-line and in-space neighboring are the sources of the co-site interference over B-PLC access networks, PLC-CA solution must keep the summation of these two interferences at authorized level by increasing distance between assigned frequencies to neighbor cells as much as possible. Hence, a $|\mathcal{C}| \times |\mathcal{C}|$ matrix, called interference distance matrix \mathbf{M}^{ID} , is needed to determine the threshold limits of interference for each pair of cells. Indeed, elements of this matrix are the thresholds determining assigned frequencies to cells leading to co-site interference or not. Expressing more precisely, for the two neighbor cells of \mathcal{C}_i and $\mathcal{C}_{i'}$, if distance between the frequency f_i of \mathcal{C}_i and $f_{i'}$ of $\mathcal{C}_{i'}$ is less than $m_{ii'}^{ID}$, i.e. $|f_i - f_{i'}| < m_{ii'}^{ID}$, then this assignment of frequency leads to network interference. But, when $|f_i - f_{i'}| \geq m_{ii'}^{ID}$, this assignment has no interference effect on the network.

According to the above considerations, the PLC-CA problem can be generally formalized as a combinatorial optimization problem with the following specifications:

- Inputs:
 1. The output of GBSP problem, which is the set of cells $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_{|\mathcal{C}|}\}$ with their details of $\{B_i, U_s^i, \mathcal{R}_s^i\}$.
 2. The sets \mathcal{F}_s and \mathcal{F}_{Forbid}^i , to compute $\mathcal{F}^i = \mathcal{F}_s - \mathcal{F}_{Forbid}^i$ for the cells.
 3. The interference distance matrix \mathbf{M}^{ID} .
 4. The aggregate neighborhood matrix \mathbf{M}^{NB} which is summation of in-line and in-space neighboring matrices.
 5. The average number of frequencies needed by each cell, i.e. A_{User} , to compute $|F_{Req}^i| = A_{User} |U_s^i|$ for each cell.
- Output:
 1. The solution matrix \mathbf{M}^{CA} , to extract the assigned frequency sets of different B-PLC cells, i.e. \mathcal{F}_{CA}^i for $i = 1, 2, \dots, |\mathcal{C}|$.
- Objectives:
 1. Minimization of the co-site interference of the B-PLC access network.
- Constraints:
 1. $\mathcal{F}_{CA}^i \subseteq \mathcal{F}^i$
 2. $|F_{Req}^i| \leq |\mathcal{F}_{CA}^i| \leq |\mathcal{F}^i|$

D. Suggested Models for the Different Aspects of the PLC-CA Problem

In the sequel, our proposed methods for the modeling of neighboring schemes, co-site interference and interference distance matrix are provided in a respective order. Compared with the existing study of

[23]-[25], the contributions and novelties of our research can be highlighted as:

- In [23]-[25], in-line neighboring scheme has only considered for the B-PLC access networks and for each pair of B-PLC cells, the neighboring is modeled either by zero or one. However, we investigate PLC-CA problem for an actual PLC network in the presence of both in-line and in-space neighboring schemes. Since the binary model of [23]-[25] has no longer holds, we suggest a fuzzy model where the aggregate neighboring intensity for each cell pair will be real number from the interval [0 2].
- In [23]-[25], elements of interference distance matrix \mathbf{M}^{ID} are considered equal to a constant number. In our study, these elements are intelligently defined as a function of aggregate neighboring intensity of the cells. As it will be described later, this definition makes solver of PLC-CA problem more sensitive to cells having greater neighboring intensity and increases the distance between assigned frequency sets to them.
- In [23]-[25], the amount of co-site interference imposed on network by different neighbor cells are considered equal to each other's. Here, as a more appropriate and efficient model, the co-site interference is defined as a function of neighboring intensity of the B-PLC cells and the assigned frequency sets to them. According to this, for each pair of neighbor cells, the more neighboring intensity is increased and the more distance between assigned frequency sets is decreased, the more interference is imposed on network.
- It is also worth to be noted that a general formulation is provided for the co-site interference which subsumes many realizations such as existing ones of [23]-[25].
- In addition to the GA, SFLA are also utilized to verify the optimum performance of the suggested models for different realizations of interference function, neighboring models and forbidden sets.
 - 1) *Formulation of the Neighboring Schemes*

The proposed models for in-line, in-space and aggregate neighboring schemes can be summarized as follows:

 - Similar to [23]-[25], in-line neighboring is modeled by a $|\mathcal{C}| \times |\mathcal{C}|$ binary matrix, whose entries are zero or one. Here, this matrix is denoted by \mathbf{M}^{NBL} , where $m_{ii'}^{NBL} = 1$ implies to in-line neighboring of cells \mathcal{C}_i and $\mathcal{C}_{i'}$, and $m_{ii'}^{NBL} = 0$ indicates these cells are not in-line neighbor.
 - Since in-space neighboring is due to physical vicinity of cell's wirings and its strength changes according to the degree of this vicinity, it is modeled by a fuzzy

approach and the elements of $|\mathcal{C}| \times |\mathcal{C}|$ in-space neighboring matrix of \mathbf{M}^{NB_S} are chosen from the interval [0 1]. In this manner, $m_{ii'}^{NB_S} = 0$ indicates cells \mathcal{C}_i and $\mathcal{C}_{i'}$ are not in-space neighbor; however, if these cells are in-space neighbor, then the value of $m_{ii'}^{NB_S}$ will be a non-zero number from (0 1]. Note that the amount of $m_{ii'}^{NB_S}$ directly depends on the vicinity degree of cell's wirings and the more physical vicinity of wirings is increased, $m_{ii'}^{NB_S}$ is more increased until reaches its upper limit equal to one, which is equivalent to in-line neighboring.

- Eventually, the aggregate neighboring is defined by $\mathbf{M}^{NB} = \mathbf{M}^{NB_L} + \mathbf{M}^{NB_S}$, which is a $|\mathcal{C}| \times |\mathcal{C}|$ matrix whose elements belonging to interval [0 2] indicating the neighboring intensity of the different B-PLC cells.

Note that since the cells are not the neighbor of themselves, the main diagonal elements of matrices \mathbf{M}^{NB_L} , \mathbf{M}^{NB_S} and \mathbf{M}^{NB} are always equal to zero.

II) Formulation of the Co-Site Interference

In this part, we present our proposed models for the co-site interference and interference distance matrix. In order to provide more insight, the function of co-site interference is first formalized for the two B-PLC cells of \mathcal{C}_i and $\mathcal{C}_{i'}$, and then generalized to the whole network.

The co-site interference generated between the cells \mathcal{C}_i and $\mathcal{C}_{i'}$ or imposed on network by the cells \mathcal{C}_i and $\mathcal{C}_{i'}$ can be formalized as follows:

$$I_{ii'}(F, m_{ii'}^{NB}) = \begin{cases} \sum_{f \in \mathcal{F}_{CA}^i} \sum_{f' \in \mathcal{F}_{CA}^{i'}} G_1(F, m_{ii'}^{NB}) & F \leq m_{ii'}^{ID} \\ 0 & F > m_{ii'}^{ID} \end{cases} \quad (1)$$

in which $I_{ii'}$ represents the co-site interference imposed on network by the cells \mathcal{C}_i and $\mathcal{C}_{i'}$. \mathcal{F}_{CA}^i and $\mathcal{F}_{CA}^{i'}$ are the assigned frequency sets to the cells \mathcal{C}_i and $\mathcal{C}_{i'}$ respectively. $F = |f_i - f_{i'}|$ is the distance between members of these sets, $m_{ii'}^{ID}$ represents the (i, i') element of interference distance matrix \mathbf{M}^{ID} , and the function of $G_1(F, m_{ii'}^{NB})$ indicates to the co-site interference for the specific values of F and $m_{ii'}^{NB}$.

As outlined earlier, $m_{ii'}^{ID}$ denotes a threshold value determining assigned frequencies to the cells \mathcal{C}_i and $\mathcal{C}_{i'}$ leading to co-site interference or not. As described below, to properly model interference distance matrix, we define this threshold as a function of neighboring intensity using $m_{ii'}^{ID} = G_2(m_{ii'}^{NB})$.

The continuous functions of $G_1(F, x)$ and $G_2(x)$ must meet following necessary conditions:

- First of all, since co-site interference does not generate by non-neighbor cells, $I_{ii'}$ must be equal to

zero for any $m_{ii'}^{Neigh} = 0$. Consequently, $G_1(F, x)$ must be such that $G_1(F, x)|_{x=0} = 0$.

- Since the amplitude of the co-site interference directly depends on the $G_1(F, x)$, this function must be such that (i): the more neighboring intensity is increased, the more co-site interference is imposed on network, (ii): the more distance between assigned frequency sets of neighbor cells is decreased, the more co-site interference is produced. Hence, $G_1(F, x)$ must be a strictly increasing function of the x and a strictly decreasing function of the F .
- In [23]-[25], the elements of the interference distance matrix \mathbf{M}^{ID} are considered equal to a constant number. Here, to provide a more efficient model, these elements are defined as a strictly increasing function of neighboring intensity using $G_2(m_{ii'}^{NB}) = m_{ii'}^{ID}$. According to this model, if the neighboring intensity of the cells \mathcal{C}_{i_1} and $\mathcal{C}_{i_1'}$ being more than that of \mathcal{C}_{i_2} and $\mathcal{C}_{i_2'}$, i.e. $m_{i_1 i_1'}^{NB} > m_{i_2 i_2'}^{NB}$, then $m_{i_1 i_1'}^{ID} > m_{i_2 i_2'}^{ID}$. Thereby, more co-site interference will be counted using (1) for the cells \mathcal{C}_{i_1} and $\mathcal{C}_{i_1'}$ in comparison with that of \mathcal{C}_{i_2} and $\mathcal{C}_{i_2'}$. Consequently, the interference function is more sensitive to the cells having greater neighboring intensity and this finally increases the distance between assigned frequencies to these cells during the process of interference minimization.

In the following, these conditions are more clarified by considering some special realizations for the $G_1(F, x)$ and $G_2(x)$.

- $G_1(F, x) = a$ and $G_2(x) = b$, for any positive number of a and b : This realization does not satisfy the above-mentioned conditions and is actually similar to that one considered in [23]-[25]. Since dependency of interference function to neighboring intensity has been ignored in this realization, taking in-space neighboring into account has no impact on solution of the PLC-CA problem and the performance of this realization in the presence and the absence of in-space neighboring is identical and similar to that one reported in [23]-[25] for in-line neighboring.
- Realizations of the $G_1(F, x)$: Any continuous function satisfying the first two of above conditions will be a realization for $g_1(F, x)$, where $g_1(F, x) = \frac{x^2}{1 + |\mathcal{F}_S|}$ and $g_1(F, x) = (1 + x^2)e^{-\frac{F}{|\mathcal{F}_S|}}$ are the two examples that are included for the numerical simulations of next section.
- Realizations of the $G_2(x)$: Examples like $G_2(x) = x + |\mathcal{F}_S|$ and $G_2(x) = |\mathcal{F}_S|x$ are two realizations of $G_2(x)$ satisfying the last condition of above and also utilized for the numerical simulations.

Eventually, the whole co-site interference of the B-PLC access network is obtained from (1) by summation over all cells as:

$$I = \frac{1}{|\mathcal{C}||\mathcal{F}_s|} \sum_{i \in \mathcal{C}} \sum_{i' \in \mathcal{C}} I_{ii'} \quad (2)$$

in which $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_{|\mathcal{C}|}\}$ and $\mathcal{F}_s = \{\mathcal{F}_1, \mathcal{F}_2, \dots, |\mathcal{F}_s|\}$ are the sets of available B-PLC cells and the available OFDM frequencies with the cardinality of $|\mathcal{C}|$ and $|\mathcal{F}_s|$, respectively.

In the next section, the evolutionary algorithms are utilized to solve PLC-CA problem via minimizing this interference function.

Numerical Results and Discussion

This section is dedicated to the numerical results and analyzing the performance of the proposed model for the optimization problem of PLC-CA. In this regard, GA and SFLA, as two population-based meta-heuristic methods, are exploited and implemented using MATLAB software. It is worth to be noted that GA has been inspired from natural selection and corrects an initial random population in the consecutive iterations using operations like elitism, selection, crossover, mutation, until converging to the solution [29], [30]. In order to prevent convergence to local optima in some problems, new advanced algorithms like SFLA have been proposed, which inspired by the collective evolution of frogs in obtaining food supply and benefits from local search and information mixing in its evolution process to find the global optimum and prevent convergence to local optima [31]. It must be noted that since the PLC-CA problem is a combinatorial optimization problem, the binary versions of GA and SFLA are exploited to solve the PLC-CA problem and optimally discover the assigned frequency sets to the cells.

Performance of the suggested model is evaluated for two different B-PLC test networks, where the first one is consisted of three B-PLC cells and utilizes an OFDM modulation having ten frequencies; however, the second one relies on six B-PLC cells and employs OFDM modulation of 64 frequencies. In order to diversify the results and allow comparison with the previous studies, the experiments of each test network have been performed for the two different scenarios as detailed in Table 1. As can be seen from the Table 1, the first scenario investigates the performance in the absence of forbidden sets for the B-PLC cells; However, the second one takes the forbidden sets into account. Moreover, numerical simulations of each scenario have also been performed for the two different cases; case 1: which only considers in-line neighboring scheme similar to that one of previous studies in [23]-[25], and case 2: where takes both in-line and in-space neighboring schemes into account. Further, to provide a fair comparison in these

two cases, the numerical simulations of them have also been performed for the identical interference functions of $G_1(F, x)$ and $G_2(x)$.

Table 1: Specifications of the considered scenarios for the performance evaluation of the B-PLC test networks

Scenario	Case	$\mathcal{F}_{\text{Forbid}}^i$	Interference Functions		Neighboring Matrices	
			G_1	G_2	\mathbf{M}^{NB_L}	\mathbf{M}^{NB_S}
First	1	No	$\frac{ \mathcal{F}_s x^2}{F + \mathcal{F}_s }$	$x + \mathcal{F}_s $	Yes	No
	2	No	$\frac{ \mathcal{F}_s x^2}{F + \mathcal{F}_s }$	$x + \mathcal{F}_s $	Yes	Yes
Second	1	Yes	$\frac{1 + x^2}{e^{ \mathcal{F}_s }}$	$ \mathcal{F}_s x$	Yes	No
	2	Yes	$\frac{1 + x^2}{e^{ \mathcal{F}_s }}$	$ \mathcal{F}_s x$	Yes	Yes

The in-line and in-space neighboring matrices of the first B-PLC test network are as below and its other details are provided in Table 2.

$$\mathbf{M}^{NB_L} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{M}^{NB_S} = \begin{bmatrix} 0 & 0.2 & 0 \\ 0.5 & 0 & 0.1 \\ 0 & 0.3 & 0 \end{bmatrix} \quad (3)$$

Table 2: Specifications of the first B-PLC test network

$ \mathcal{C} = 3, \mathcal{F}_s = 10$		
\mathcal{C}_i	$ \mathcal{F}_{\text{Req}}^i $	$\mathcal{F}_{\text{Forbid}}^i$
\mathcal{C}_1	5	{9, 10}
\mathcal{C}_2	5	{2, 3}
\mathcal{C}_3	6	{5, 10}

As can be seen from the Table 2, this test network totally needs 16 frequencies and the PLC-CA solution must assign them from the available ten frequencies by taking the forbidden sets and neighboring model into account. In the other words, the solution matrix \mathbf{M}^{CA} must such that while the assigned frequency sets to the cells, i.e. \mathcal{F}_{CA}^i , belonging to their allowable frequency sets $\mathcal{F}^i = \mathcal{F}_s - \mathcal{F}_{\text{Forbid}}^i$, the frequencies of the neighbor cells be far away from each other's as much as possible.

For the numerical simulations, a random population of size $N = 40$ is first generated for both GA and SFLA, and fitness value for the elements of this population are calculated according to the network's interference given in (2). Then, for the GA, the members (chromosomes) of this population exchanged their information through GA operations including elitism, selection, crossover of rate 0.9 and mutation of rate 0.1, to produce next generation of chromosomes and this process is continued until an

optimal solution is obtained or stopping criterion is met. For the SFLA, the generated random population is subdivided into $m = 10$ memeplexes, each contains $n = N/m = 4$ members (frogs). Then, a local search is performed for each memeplex to improve the frog with the worst fitness. When tenth iterations of local search are performed, all frogs of population are mixed together and partitioned into $m = 10$ memeplexes again to repeat the local search. This process is continued iteration by iteration until an optimal solution is obtained or stopping criterion is met.

In the cases of first B-PLC test network, the curves of the best fitness (co-site interference) versus iteration (for both GA and SFLA) have been depicted in Fig. 3 for the first scenario and Fig. 4 for the second one.

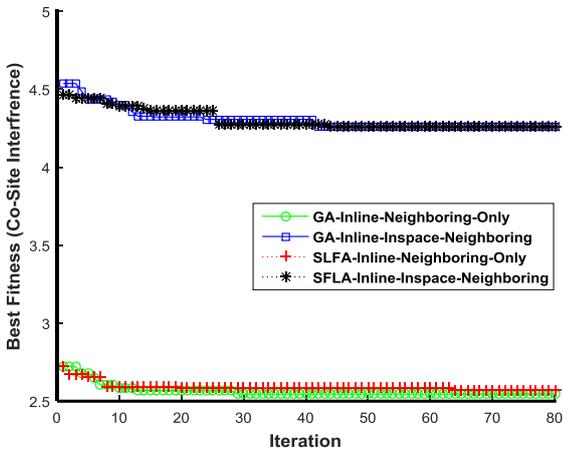


Fig. 3: First B-PLC test network: curves of the co-site interference versus iterations for the first scenario.

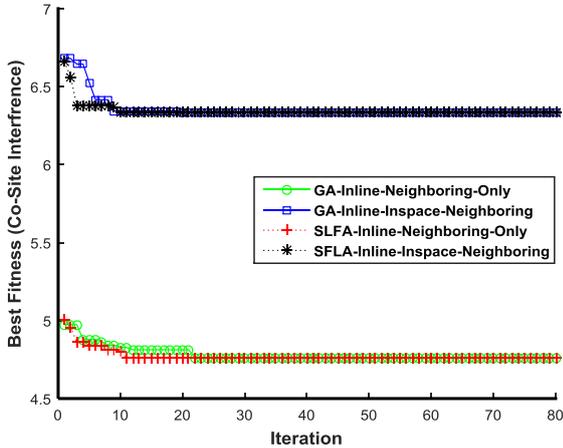


Fig. 4: First B-PLC test network: curves of the co-site interference versus iterations for the second scenario.

Further, the solution matrix of the first scenario takes one of the following forms for each realization of GA and SFLA:

$$M^{CA} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (4)$$

However, in the case of the second scenario, the solution matrix has only below format:

$$M^{CA} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \quad (5)$$

As demonstrated in Fig. 3 and Fig. 4, the suggested interference function can successfully model and track the network interference according to the considered neighboring schemes in the two considered cases for each scenario. Moreover, as evident from (4) and (5), the resulting solution matrices indicating to an optimum channel allocation for the both scenarios and also for both GA and SFLA. This is because the assigned frequency sets to the neighboring cells have the maximum distance from each other's. Expressing in details, these results and their modifications with respect to existing studies can be highlighted as below:

- First, since the case 1 of each scenario only considers in-line neighboring scheme and in case 2 both in-line and in-space neighboring are taken into account, these two cases compare the results of our suggested model for PLC-CA problem in which co-site interference varies depending on neighboring intensity and assigned frequency sets to the cells with that one in which interference variations is only due to changes in assigned frequency sets. Due to taking both in-line and in-space neighboring schemes into account for the actual B-PLC networks of the case 2, it is evident that the amount of co-site interference imposed on this real network must be more than of case 1. As shown in Fig. 3 and Fig. 4, the suggested interference function of (2) can properly model and successfully track the variations of co-site interference in these two cases. Noted that the case 1 of each scenario can be considered as an improved version of the assumed model in [23]-[25]. This is due to the fact both ones, i.e. case 1 and the works of [23]-[25], relying on in-line neighboring scheme; However, the co-site interference of case 1 varies according to the distance between assigned frequency sets, but this dependency has not considered in [23]-[25]. Thus, the results of the cases 1 and 2 provide a comparison between the results of previous studies and our method.
- Second, both GA and SFLA leading to optimum channel allocation for all considered cases and scenarios. For instance, as evident from (4) and (5), there is no common frequencies between assigned sets to the cells 1 and 2 of the first test network. Note that existence of one common frequency for cells 2 and 3 of (4) is due to the fact that these two cells totally needing 11 frequencies, this is while the available OFDM channels are equal to $|\mathcal{F}_S| = 10$. Similarly, existence of two common frequencies for

these cells in (5) is due to limited number of OFDM channels and also due to considering forbidden sets which restricts it more cells. It is also important to be noted that when the solution matrices are compared for two different pairs of neighbor cells at intermediate iterations, it is observed that distance between assigned frequency sets to cells of pair with greater neighboring intensity is more than that of other pair. This is due to the fact that we defined the interference distance matrix so that the considered solver be more sensitive to the cells having greater neighboring intensity and as a result, the frequencies allocated to these cells have more distance from each other in intermediate steps where the optimal solution for the network has not been obtained yet.

- Third, as demonstrated in Fig. 3 and Fig. 4, both GA and SFLA are finally converged to a similar co-site interference. However, they reach to this value in different iterations due to difference in computational complexity. Expressing more precisely, if the solution matrix of the GA and SFLA, i.e. M^{CA} , is compared before convergence at an intermediate iteration, it is observed that SFLA providing a more optimal solution than the GA. However, it should be noted that this comparing is not entirely accurate because SFLA also benefits from an iteration-based local search in each iteration. Therefore, it can be concluded that the SFLA provides better solutions in intermediate iterations due to performing local search and wasting the time; however, the both ones eventually leading to an optimal solution.
- Forth, curves of co-site interference versus iterations have an expected behavior for both GA and SFLA and follows a non-increasing function. Further, when forbidden sets of the frequencies are considered in the second scenario of Table 1, both GA and SFLA successfully avoids assignment of these sets to the B-PLC cells and the channels are allocated according to the allowable frequency sets $\mathcal{F}^i = \mathcal{F}_s - \mathcal{F}_{Forbid}^i$.
- Eventually, the comprehensiveness of the suggested interference function is also confirmed by performing numerical simulations for two different realizations of both $G_1(F, x)$ and $G_2(x)$. Note that these functions are utilized in (1) and (2) to define co-site interference.

Here, the second test network is considered to evaluate performances of the suggested model for a more complex structure. Note this network is consisted of the six B-PLC cells and exploits OFDM modulation of $|\mathcal{F}_s| = 64$ frequencies.

The other details of this network are provided in Table 3 and its in-line and in-space neighboring matrices are given in (6) and (7).

$$M^{NB_L} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad (6)$$

$$M^{NB_S} = \begin{bmatrix} 0 & 0 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0.7 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 & 0 & 0 \\ 0.9 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (7)$$

Table 3: Specifications of the second B-PLC test network

$ \mathcal{C} = 6, \mathcal{F}_s = 64$		
\mathcal{C}_i	$ \mathcal{F}_{Req}^i $	\mathcal{F}_{Forbid}^i
\mathcal{C}_1	20	{61,62,63,64}
\mathcal{C}_2	15	{1,2,3,4,5}
\mathcal{C}_3	32	\emptyset
\mathcal{C}_4	40	\emptyset
\mathcal{C}_5	27	{1,2,10,44,45}
\mathcal{C}_6	27	{60,61,62,63,64}

For the numerical simulations of the second B-PLC test network, population size has been considered equal to $N = 120$ and crossover and mutation rates of the GA are set to 0.9 and 0.2, respectively. For the SFLA, this population is subdivided into $m = 12$ memplexes of $n = N/m = 10$ members. Curves of the best fitness versus iterations have been depicted in Fig. 5 for the first scenario and Fig. 6 for the second one.

Furthermore, according to the solution matrix of the first scenario, the assigned sets to the neighboring cells 1 and 4 have one of the following forms:

$$\begin{cases} \mathcal{F}_{CA}^1 = \{1,2, \dots, 20\}, \mathcal{F}_{CA}^4 = \{64,63, \dots, 25\} \\ \mathcal{F}_{CA}^1 = \{64,63, \dots, 45\}, \mathcal{F}_{CA}^4 = \{1,2, \dots, 40\} \end{cases} \quad (8)$$

However, when the forbidden sets are taken into the account for the second scenario, the solution is:

$$\mathcal{F}_{CA}^1 = \{1,2, \dots, 20\}, \mathcal{F}_{CA}^4 = \{64,63, \dots, 25\} \quad (9)$$

In similar manner, for the first scenario, the assigned sets to the neighboring cells of 5 and 6 belonging to one of the following groups:

$$\begin{cases} \mathcal{F}_{CA}^5 = \{1, \dots, 27\}, \mathcal{F}_{CA}^6 = \{64, \dots, 38\} \\ \mathcal{F}_{CA}^5 = \{64, \dots, 38\}, \mathcal{F}_{CA}^6 = \{1, \dots, 27\} \end{cases} \quad (10)$$

However, due to considering forbidden sets in the second scenario, the solution will be as:

$$\begin{cases} \mathcal{F}_{CA}^5 = \{64,63, \dots, 46,43,42, \dots, 36\} \\ \mathcal{F}_{CA}^6 = \{1, \dots, 27\} \end{cases} \quad (11)$$

As can be seen from Fig. 5 and Fig. 6, the curves of

the co-site interference have their expected behavior as non-increasing functions and track the network interference according to the considered cases in these scenarios. Furthermore, as evident from (8)-(11), there is no common frequencies between assigned sets to the neighboring cells, i.e. cells 1 and 4, and cells 5 and 6. Moreover, the assigned sets to these cells have maximum distance from each other. In general, it is observed that the numerical results of the second test network are consistent with that one of the first test network and all mentioned points for the results of first test network are valid for the second one as well. Thereby, it is concluded that the suggested method has the ability to model PLC-CA problem in actual scenarios, tracking the network interference in these situations and providing an optimal solution for them. Further, it is confirmed the suggested method, as a comprehensive model, includes all previous research.

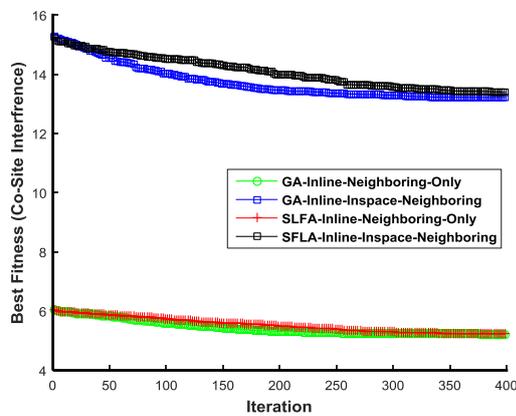


Fig. 5: Second B-PLC test network: curves of the co-site interference versus iterations for the first scenario.

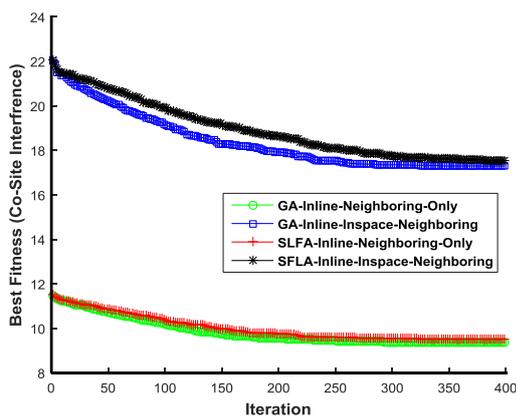


Fig 6: Second B-PLC test network: curves of the co-site interference versus iterations for the second scenario.

Conclusion

In this paper, the problem of channel allocation was investigated for the B-PLC access network. In order to overcome the limitations of the existing study and provide a comprehensive analysis, an actual B-PLC access that is in expose of both in-line and in-space neighboring

schemes was assumed and the problem modelling, formulations, and suggested solutions are planned in this regard. To achieve these goals, PLC-CA problem is first formalized as an optimization problem; in-line neighboring is considered as zero number or one; in-space neighboring is modeled as a real number from the interval [0 1]; and the aggregate neighboring intensity is assumed as the summation of these two numbers. Then, the co-site interference of the B-PLC access networks is defined as a function of the aggregate neighboring intensity and the assigned frequency sets to the cells, so that when the value of neighboring intensity is increased and the distance of assigned frequency sets is decreased, the interference is increased. Finally, the meta-heuristic methods of GA and SFLA were employed to solve the resulting PLC-CA problem and evaluate its performance for different scenarios and compare it with the existing research. The results were confirmed that suggested method can successfully model channel allocation problem in actual B-PLC networks and provides an optimum solution for both GA and SFLA by avoiding the assignment of forbidden frequencies and increasing the distance between neighboring cells as much as possible. It was also revealed that this method includes and improves all previous research and worked as a comprehensive alternative for tracking co-site interference in different situations.

Author Contributions

M. Sheikh-Hosseini designed and modeled the suggested methods for the PLC-CA problem and wrote the manuscript.

S. M. Nosratabadi designed the experiments and edited the manuscript.

Acknowledgment

The authors would like to acknowledge the financial support of Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran, under Contract number 96/3166.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

<i>PLC</i>	Power Line Communications
<i>B-PLC</i>	Broadband-PLC
<i>N-PLC</i>	narrowband-PLC
<i>PLC-CA</i>	PLC Channel Allocation
<i>BS</i>	Base Station

GBSP	Generalized Base Station Placement
GA	Genetic Algorithm
SFLA	Shuffled Frog-Leaping Algorithm
OFDM	Orthogonal Frequency Division Multiplexing

References

- [1] M. Schwartz, "Carrier-wave telephony over power lines: Early history," *IEEE Commun. Mag.*, 47(1): 14-18, 2009.
- [2] S. Mudrievskiy, "Power line communications: State of the art in research, development and application," *AEU Int. J. Electron. Commun.*, 68(7): 575-577, 2014.
- [3] C. Cano et al., "State of the art in power line communications: From the applications to the medium," *IEEE J. Sel. Areas Commun.*, 34(7): 1935-1952, 2016.
- [4] S. Galli, A. Scaglione, Z. Wang, "For the grid and through the grid: the role of power line communications in the smart grid," *Proc. IEEE*, 99(6): 998-1027, 2011.
- [5] M. Sheikh-Hosseini, "Evolution of power line communications: From a fixed telephone system to telecommunication technology of smart energy grid," *Ir Iran. J. Eng. Educ.*, 20(80): 71-96, 2019.
- [6] S. Chen et al., "Ultra-Wideband Powerline Communication (PLC) above 30MHz," *IET Commun.*, 3(10): 1587-1596, 2009.
- [7] K. Ramanathan, N.J.R. Muniraj, "DWT-IDWT-based MB-OFDM UWB with digital down converter and digital up converter for power line communication in the frequency band of 50 to 578 MHz," *Ann. Telecommun.*, 70: 181-196, 2015.
- [8] M. Sheikh-Hosseini, G.A. Hodtani, M. Molavi-kakhki, "Capacity analysis of PLC point-to-point and relay channels," *Trans. Emerging Telecommun. Technol.*, 27(2): 200-215, 2016.
- [9] L. Lampe, A.M. Tonello, T.G. Swart, *Power Line Communications: Principles, Standards and Applications from multimedia to smart grid*. John Wiley & Sons, 2016.
- [10] E. Ancillotti, R. Bruno, M. Conti, "The role of communication systems in smart grids: Architectures, technical solutions and research challenges," *Comput. Networks*, 36(17): 1665-1697, 2013.
- [11] A. Pittolo et al., "In-vehicle power line communication: Differences and similarities among the in-car and the in-ship scenarios," *IEEE Veh. Technol. Mag.*, 11(2): 43-51, 2016.
- [12] M. Yigit et al., "Power line communication technologies for smart grid applications: A review of advances and challenges," *Comput. Networks*, 70: 366-383, 2014.
- [13] G. Lopez et al., "The role of power line communications in the smart grid revisited: Applications, challenges, and research initiatives," *IEEE Access*, 7: 117346-117368, 2019.
- [14] M. Kashef, M. Abdallah, M.N. Al-Dhahir, "Transmit power optimization for a hybrid PLC/VLC/RF communication system," *IEEE Trans. Green Commun. Networking*, 2(1): 234-245, 2017.
- [15] Y.S. Reddy et al., "Optimisation of indoor hybrid PLC/VLC/RF communication systems," *IET Commun.*, 14(1): 117-126, 2020.
- [16] M. Sheikh-Hosseini, G.A. Hodtani, "On the capacity of additive white mixture gaussian noise channels," *Trans. Emerging Telecommun. Technol.*, 30(7): e3585, 2019.
- [17] E. Biglieri, "Coding and modulation for a horrible channel," *IEEE Commun. Mag.*, 41(5): 92-98, 2003.
- [18] M. B. Noori Shirazi, A. Golestani, H. Ahmadian Yazdi, A. Habibi Daronkola, "Analysis and comparison of PAPR reduction techniques in OFDM systems," *J. Electr. Comput. Eng. Innovations*, 3(1): 37-45, 2015.
- [19] A. Haidine, I. Mellado, R. Lehnert, "PANDeMOO: Powerline communications access network designer based on multi-objective optimisation," presented at the IEEE Symp. ISPLC, Vancouver, Canada, 2005.
- [20] A. Haidine, R. Lehnert, "Solving the generalized base station placement problem in the planning of broadband power line communications access networks," presented at the 2th IEEE Conf. AccessNets., Ottawa, Ont., Canada, 2007.
- [21] A. Haidine, R. Lehnert, "Placement of base stations in broadband power line communications access networks by means of multi-criteria optimization," presented at the 3th IEEE Conf. AccessNets., Las Vegas, USA, 2008.
- [22] A. Haidine, R. Lehnert, "Optimisation of base station location in broadband power line communications access networks," *Int. J. Comm. Network Distr. Syst.*, 5(4): 347-374, 2010.
- [23] A. Haidine, R. Lehnert, "Analysis of the channel allocation problem in broadband power line communications access network," presented at the IEEE Symp. ISPLC, Pisa, Italy, 2007.
- [24] A. Haidine, R. Lehnert, "Modeling of channel allocation in broadband powerline communications access networks as a multi-criteria optimization problem," presented at the 3th IEEE Conf. AccessNets., Las Vegas, USA, 2008.
- [25] A. Haidine, R. Lehnert, "The channel allocation problem in broadband power line communications access networks: analysis, modeling and solutions," *Int. J. Autonom. Adapt. Comm. Syst.*, 3(4): 396-418, 2010.
- [26] L.P. Do, R. Lehnert, "Dynamic resource allocation protocol for large PLC networks," presented at the IEEE Symp. ISPLC, Beijing, China, 2012.
- [27] R. Kulshrestha, "Channel allocation and ultra-reliable communication in CRNs with heterogeneous traffic and retrials: A dependability theory-based analysis," *Comput. Commun.*, 158: 51-63, 2020.
- [28] H. Zhang, H. Li, J.H. Lee, "Efficient subchannel allocation based on clustered interference alignment in ultra-dense femtocell networks," *China Commun.*, 14(4): 1-10, 2017.
- [29] M.H. Refan, A. Dameshghi, "RTDGPS implementation by online prediction of GPS position components error using GA-ANN model," *J. Electr. Comput. Eng. Innovations*, 1(1): 43-50, 2013.
- [30] Y. Rohani, Z. Torabi, S. Kianian, "A novel hybrid genetic algorithm to predict students' academic performance," *J. Electr. Comput. Eng. Innovations*, 8(2): 219-232, 2020.
- [31] T.H. Huynhm "A modified shuffled frog leaping algorithm for optimal tuning of multivariable PID controllers," in *Proc. 2008 IEEE International Conference on Industrial Technology: 1-6*, 2008.

Biographies



Mohsen Sheikh-Hosseini was born in Iran in 1983. He received the B.S. degree in electrical engineering from Shahid Bahonar University of Kerman, Kerman, Iran in 2007, and then he received the M.S. and Ph.D. degrees both in telecommunications from Ferdowsi University of Mashhad, Mashhad, Iran in 2009 and 2014, respectively. He is currently with the Department of Computer and Information Technology, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran. His research interests include wireless communications, power line communications, wireless sensor networks, smart grids, and machine learning applications in telecommunications.



Seyyed Mostafa Nosratabadi received the B.Sc. degree from Shahid Bahonar University of Kerman, Kerman in 2008, M.Sc. degree from University of Kashan in 2011, both in electrical engineering. He received Ph.D. degree in electronics engineering from university of Isfahan, Iran, in 2016 and then joined to Department of Electrical Engineering, Sirjan University of Technology, Sirjan. His research

interests are analysis of power systems, power electronic, smart grid and micro grid.

Copyrights

©2020 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

M. Sheikh-Hosseini, S.M. Nosratabadi, "Modelling and optimization of channel allocation for power line communications access networks in the presence of in-line and in-space interference," *J. Electr. Comput. Eng. Innovations*, 9(1): 103-114, 2021.

DOI: [10.22061/JECEI.2020.7459.395](https://doi.org/10.22061/JECEI.2020.7459.395)

URL: http://jecei.sru.ac.ir/article_1491.html





Research paper

Stochastic Block NIHT Algorithm for Adaptive Block-Sparse System Identification

Z. Habibi^{1,*}, H. Zayyani², M.S.E Abadi³

¹Research Institute for Information and Communications Technologies, Academic Center for Education, Culture and Research, Tehran, Iran.

²Faculty of Electrical and Computer Engineering, Qom University of Technology, Qom, Iran.

³Electronics Engineering Department, Faculty of Electrical Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran.

Article Info

Article History:

Received 20 April 2020
Reviewed 12 June 2020
Revised 24 August 2020
Accepted 24 November 2020

Keywords:

Compressive Sensing (CS)
System identification
Block-sparse system
Adaptive filter

*Corresponding Author's Email
Address:
zhabibi9819@gmail.com

Abstract

Background and Objectives: Compressive sensing (CS) theory has been widely used in various fields, such as wireless communications. One of the main issues in the wireless communication field in recent years is how to identify block-sparse systems. We can follow this issue, by using CS theory and block-sparse signal recovery algorithms.

Methods: This paper presents a new block-sparse signal recovery algorithm for the adaptive block-sparse system identification scenario, named stochastic block normalized iterative hard thresholding (SBNIHT) algorithm. The proposed algorithm is a new block version of the SSR normalized iterative hard thresholding (NIHT) algorithm with an adaptive filter framework. It uses a search method to identify the blocks of the impulse response of the unknown block-sparse system that we wish to estimate. In addition, the necessary condition to guarantee the convergence for this algorithm is derived in this paper.

Results: Simulation results show that the proposed SBNIHT algorithm has a better performance than other algorithms in the literature with respect to the convergence and tracking capability.

Conclusion: In this study, one new greedy algorithm is suggested for the block-sparse system identification scenario. Although the proposed SBNIHT algorithm is more complex than other competing algorithms but has better convergence and tracking capability performance.

Introduction

Compressive sensing (CS) theory has been widely used in various fields, such as mathematics, signal processing, and wireless communications [1]-[3]. One of the main issues in the wireless communication field in recent years is how to identify block-sparse systems. According to [4]-[8], by using the CS theory and proposing block-sparse signal recovery algorithms, we can follow this issue. In the sparse systems, the impulse response of the system has a small number of non-zero coefficients. In

some practical cases, the impulse response of the system may include a small number of subspaces (block) (see Fig. 3), which is known as a block-sparse system. With this description, we can model a block-sparse system with impulse response $\mathbf{h} \in R^N$ as

$$\mathbf{h} = [h_1, \dots, h_d, h_{d+1}, \dots, h_{2d}, \dots, h_{N-d+1}, \dots, h_N]^T \quad (1)$$

$$\mathbf{h}^T [1] \quad \mathbf{h}^T [2] \quad \mathbf{h}^T [L]$$

where $\mathbf{h}^T [i]$ ($i \in \{1, 2, \dots, L\}$), indicates the i -th block of the \mathbf{h} , d is the block length, and $N = dL$. The \mathbf{h} vector is termed K block-sparsity, where $K \in \{1, 2, \dots, L\}$, is the maximum number of blocks involving nonzero components.

Three major approaches sparse signal recovery (SSR) algorithms including convex optimization algorithms [4], [9], greedy pursuit algorithms [5], [8], and stochastic gradient descent (SGD) algorithms [10] have been recently developed to implement the block-sparse signal reconstruction.

In order to reconstruct block-sparse signals, in the convex optimization class, in [4], [9], the basis pursuit (BP) algorithm has been extended the l_1 -minimization to a mixed l_2/l_1 -norm minimization in the recovery algorithm. Also an improved algorithm, named dynamic recovery of block-sparse signal (D-BSS) has been proposed in [11]. In [12], the proposed algorithm benefits from l_0/l_2 penalty to reconstruct the block-sparse signal. However the mentioned algorithms show high computational cost, and so they are not suitable for the large-scale scenario [13].

Recently, a block version of the greedy pursuit algorithms such as matching pursuit (MP), orthogonal MP (OMP), stagewise OMP (StOMP), and iterative hard thresholding (IHT), named BMP, BOMP [4], BStOMP [8], and BIHT [5] respectively, have been introduced that show a better performance than their original versions for the block-sparse signal recovery scenario. Moreover, the extended version of the compressive sampling matching pursuit (CoSaMP) algorithm has been presented in [5]. Also, the block normalized iterative hard thresholding (BNIHT) algorithm, has been proposed in [14]. Although the greedy pursuit class algorithms have low computational complexity, its performance degrades under strong background noise [13].

The SGD-based sparse adaptive filtering algorithms such as zero-attracting least-mean-square (ZA-LMS) algorithm and l_0 -norm least mean square (l_0 -LMS) algorithms in [15], which have shown moderate computational complexity and robustness against noise [13], prove a better performance than other the types of the SSR algorithms in the sparse signal reconstruction process. Using this approach, several SGD-based algorithms have been proposed for the block-sparse signal reconstruction in recent years. In the block-sparse LMS (BS-LMS) algorithm [16], and block zero-point attracting projection (BZAP) [17], a penalty of block sparsity as a mixed norm of the adaptive tap-weights as $l_{2,0}$ -norm in [16], and a mixed $l_{1,0}$ -norm in [17] have been added to the cost function of the LMS and ZAP algorithms respectively. Moreover, block zero attracting LMS (BZA-LMS) and the block l_0 -norm LMS (Bl $_0$ -LMS) algorithms which can sense the block-sparse structure

information of the block-structured sparse signal, have been presented in [18].

In order to improve the convergence performance of the solution of the block-sparse system identification problem, this paper presents a new block version of the NIHT algorithm with an adaptive filter framework, named stochastic block normalized iterative hard thresholding (SBNIHT) algorithm. In this work, we use a search method to identify the blocks of the impulse response of the block-sparse system that we wish to estimate. Simulation results show that the proposed algorithm has a faster convergence rate and a better tracking capability than other the competing algorithms in the literature.

Adaptive Filter Framework for Sparse Signal Recovery

Based on the CS theory, by using the proper dictionary matrix $\mathbf{A} \in R^{M \times N}$ ($M \ll N$) that satisfies the restricted isometry property (RIP), we can compress the sparse signal $\mathbf{s} \in R^N$ to a down-sampling signal $\mathbf{q} \in R^M$ [19]. To achieve more accurate results, the unavoidable background noise \mathbf{v} e.g., additive white Gaussian noise (AWGN) can be considered as an additional term in the SSR equation. Therefore, an SSR problem is formulated by the following underdetermined equation

$$\mathbf{q} = \mathbf{A}\mathbf{s} + \mathbf{v} \quad (2)$$

In order to achieve less complexity cost and robustness against background noise, we can use an adaptive signal processing framework to solve the SSR problem [20].

In the sparse systems identification scenario which can be considered as an SSR problem, the desired signal with adaptive filtering framework is achieved as

$$d(n) = \mathbf{u}^T(n)\mathbf{h}_o + \eta(n) \quad (3)$$

where $\mathbf{u}(n) = [u(n), u(n-1), \dots, u(n-M+1)]^T$, is the input signal vector, $(\cdot)^T$ denotes the transpose, \mathbf{h}_o is the impulse response of the unknown system with the length M that we wish to estimate, and $\eta(n)$ indicates the additive white Gaussian noise. Also, the recursion error is obtained as

$$e(n) = d(n) - \mathbf{u}^T(n)\mathbf{h}(n) \quad (4)$$

where $\mathbf{h}(n) = [h_1(n), h_2(n), \dots, h_M(n)]^T$, denotes the iterative reconstruction impulse response vector. In Table 1 and Fig. 1, we can see the SSR problem that is effectively solved by the adaptive framework, while each row vector \mathbf{a}_j in the dictionary matrix \mathbf{A} in (2) plays the role of $\mathbf{u}^T(n)$ in the adaptive framework in (3), and the

components of the compressed measurement vector \mathbf{q} in (2) that is introduced as q_j , is regarded as $d(n)$ in (3) [13].

Table 1: Corresponding variables between CS problem and adaptive framework [13].

CS Problem	Adaptive Framework
$\mathbf{a}_j, j \in \{1, 2, \dots, M\}$	$\mathbf{u}^T(n)$
$\mathbf{s}(n)$	$\mathbf{h}(n)$
$q_j = \mathbf{a}_j \mathbf{s} + v_j$	$d(n) = \mathbf{u}^T(n) \mathbf{h}_o + \eta(n)$

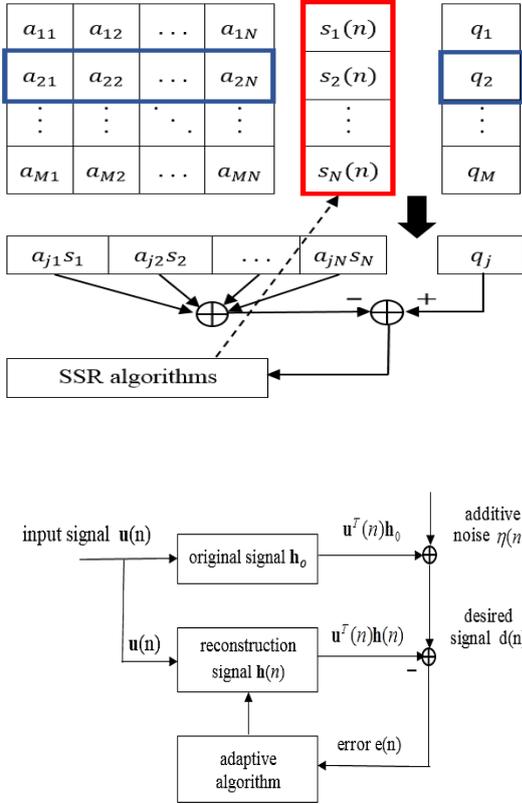


Fig. 1: Adaptive framework for CS problem [13].

Proposed Algorithm

To improve the convergence performance of the solution of the block-sparse system identification problem, this paper by considering this problem as an SSR problem, presents a new block version of the greedy NIHT algorithm with an adaptive filter framework. The proposed stochastic block normalized iterative hard thresholding (SBNIHT) algorithm uses a search method to identify the blocks of the impulse response of the unknown the block-sparse system.

To proceed, we make the following assumptions:

A1: The impulse response of the unknown block-sparse system has at most K blocks of nonzero coefficients;

A2: The maximum length of the blocks including nonzero coefficients is L ;

A3: The numbers of zero coefficients between two adjacent blocks are at least L .

The proposed algorithm benefits from a cost function as:

$$J = \frac{1}{2} E\{|e(n)|^2\} \quad (5)$$

where $E\{\cdot\}$ denotes the expectation, and $e(n)$ is the recursion error which is obtained as (4).

Based on (4), and dropping the time index for simplicity, i.e., $\mathbf{R} \equiv \mathbf{R}(n)$, we can expand (5) for a time-varying system as

$$\begin{aligned} J(\mathbf{h}, \mathbf{R}, \mathbf{r}) &= \frac{1}{2} E\{|d - \mathbf{u}^T \mathbf{h}|^2\} \\ &= \frac{1}{2} (E\{d^2\} - 2\mathbf{h}^T \mathbf{r} + \mathbf{h}^T \mathbf{R} \mathbf{h}) \end{aligned} \quad (6)$$

where $\mathbf{R} = E\{\mathbf{u}(n)\mathbf{u}(n)^T\}$ is the autocorrelation matrix of the input vector $\mathbf{u}(n)$, and $\mathbf{r} = E\{\mathbf{u}(n)d(n)\}$ is the cross-correlation of the input vector $\mathbf{u}(n)$ and the desired signal $d(n)$.

Assuming that the unknown system has a time-varying nature, we use an exponentially time-averaged window to obtain the \mathbf{R} matrix and the \mathbf{r} vector

$$\mathbf{R}(n) = \lambda \mathbf{R}(n-1) + \mathbf{u}(n)\mathbf{u}(n)^T \quad (7)$$

$$\mathbf{r}(n) = \lambda \mathbf{r}(n-1) + \mathbf{u}(n)d(n)^* \quad (8)$$

where $\lambda \in (0, 1]$ is the forgetting factor, and $(\cdot)^*$ denotes the complex conjugate. It is noteworthy that we consider the term *stochastic* for the proposed algorithm because the \mathbf{R} matrix and \mathbf{r} vector in (7) and (8), are stochastic quantities.

In the following, we obtain the negative gradient of (6) at the n -th iteration as

$$-\nabla_{\mathbf{h}} J(\mathbf{h}, \mathbf{R}, \mathbf{r}) = \mathbf{r} - \mathbf{R} \mathbf{h} \quad (9)$$

By using the steepest descent principle, we have

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \mu(n) \nabla_{\mathbf{h}} J(\mathbf{h}, \mathbf{R}, \mathbf{r}), \quad (10)$$

where $\mu(n)$ denotes the step-size of the n -th time iteration. By substituting (9) into (10), we can rewrite (10) as

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu(n)(\mathbf{r}(n) - \mathbf{R}(n)\mathbf{h}(n)) \quad (11)$$

In the following to make a better estimate of the unknown block-sparse system, we apply a block-sparsity constraint as a hard thresholding operator $H_k(\cdot)$ (similar to the NIHT algorithm [21]) to the right-hand side of (11)

$$\mathbf{h}(n+1) = H_k(\mathbf{h}(n) + \mu(n)(\mathbf{r}(n) - \mathbf{R}(n)\mathbf{h}(n))) \quad (12)$$

where, by defining

$$\bar{\mathbf{h}}(n) = H_k(\mathbf{h}(n)) \quad (13)$$

and

$$\begin{aligned} \mathbf{p}(n) &= H_k(\mathbf{r}(n) - \mathbf{R}(n)\mathbf{h}(n)) \\ &= \mathbf{r}(n) - \mathbf{R}(n)\bar{\mathbf{h}}(n) \end{aligned} \quad (14)$$

and substituting (13) and (14) into (12), we can rewrite the update formula of tap-weights for the proposed SBNIHT algorithm as

$$\mathbf{h}(n+1) = \bar{\mathbf{h}}(n) + \mu(n)\mathbf{p}(n) \quad (15)$$

where the vector $\bar{\mathbf{h}}(n)$ in (13-15) is an estimated impulse response vector of the adaptation step n , $(\mathbf{h}(n))$, which its components that do not belong to the support set $\Lambda^{(n)}$, have been zeroed. In here, we consider the support set $\Lambda^{(n)}$ as the indexes sets of the K blocks with the length L involving nonzero components in the estimated vector $\mathbf{h}(n)$. By using the line search optimization method, we can consider the step-size of the SBNIHT algorithm in (15) as

$$\mu(n) = \frac{\mathbf{p}_{|\Lambda^{(n)}}(n)^T \mathbf{p}_{|\Lambda^{(n)}}(n)}{\mathbf{p}_{|\Lambda^{(n)}}(n)^T \mathbf{R}_{|\Lambda^{(n)}}(n) \mathbf{p}_{|\Lambda^{(n)}}(n)} \quad (16)$$

where $\mathbf{p}_{|\Lambda^{(n)}}(n)$ is a sub-vector of the $\mathbf{p}(n)$ vector that contains only the elements with the indexes sets of the support set $\Lambda^{(n)}$ and $\mathbf{R}_{|\Lambda^{(n)}}(n)$ is the sub-matrix of the \mathbf{R} matrix that includes corresponding columns and rows, indexed by support set $\Lambda^{(n)}$.

A. Identify the Support Set Λ

In the following we present a new method with a search approach to identify the support set Λ at each iteration. To achieve this purpose, we define at first, a new vector named $\tilde{\mathbf{h}}$, equal to the iterative reconstruction \mathbf{h} vector

$$\tilde{\mathbf{h}} = \mathbf{h} \quad (17)$$

Then, according to Fig. 2, we find the index of the component with the largest absolute value of the magnitude in the vector $\tilde{\mathbf{h}}$

$$\Omega = \arg \max(|\tilde{\mathbf{h}}|, 1) \quad (18)$$

Then, in the vector $\tilde{\mathbf{h}}$, between the L blocks with the length L that contain the obtained index set Ω (see Fig. 2), we find the index set of the block with the largest l_2 -norm value

$$\Delta = \arg \max(\{\|\tilde{\mathbf{h}}_{[1]}\|_2, \|\tilde{\mathbf{h}}_{[2]}\|_2, \dots, \|\tilde{\mathbf{h}}_{[L]}\|_2\}, 1) \quad (19)$$

Then, we change the magnitude of the components of the obtained block $\tilde{\mathbf{h}}_{[\Delta]}$ as

$$\tilde{\mathbf{h}}_{[\Delta]} = \{\tilde{\mathbf{h}}_{(\Omega+\Delta-L)}, \tilde{\mathbf{h}}_{(\Omega+\Delta-L+1)}, \dots, \tilde{\mathbf{h}}_{(\Omega+\Delta-1)}\} \quad (20)$$

into zero value

$$\tilde{\mathbf{h}}_{|\Psi_m} = 0 \quad (21)$$

where $m \in \{1, 2, \dots, K\}$, and

$$\Psi_m = \{\Omega + \Delta - L, \Omega + \Delta - L + 1, \dots, \Omega + \Delta - 1\} \quad (22)$$

includes the indexes sets of the vector $\tilde{\mathbf{h}}_{[\Delta]}$. In order to find the support set Λ at each time iteration, we repeat the stages (18-22), until the finding K blocks. Therefore, the support set Λ at each time iteration is obtained by merging indexes sets which are resulted in each stage as the subspace Ψ_m where $m \in \{1, 2, \dots, K\}$

$$\Lambda^{(n)} = \Psi_1 \cup \Psi_2 \cup \dots \cup \Psi_K \quad (23)$$

Then, the block sparse vector $\bar{\mathbf{h}}(n+1)$ can be obtained by pruning the elements which do not belong to the support set $\Lambda^{(n)}$

$$\bar{\mathbf{h}}_{|\Lambda^{(n)}}(n+1) = \mathbf{h}_{|\Lambda^{(n)}}(n+1), \quad \bar{\mathbf{h}}_{|\Lambda^{(n)c}}(n+1) = 0 \quad (24)$$

According to the above stages, we can summarize the steps of the proposed the algorithm at each time instant as:

- Step 1)** Update the proxy vector via (14);
- Step 2)** Define the step-size value by using the line search optimization method as (16), and estimate the unknown vector \mathbf{h} via the gradient update equation (15);
- Step 3)** Identify the support set Λ that shows the set of the positions of the nonzero components corresponding to the K blocks with the length L of the estimated vector \mathbf{h} ;
- Step 4)** Prune the elements which do not belong to the support set Λ , for the estimated vector \mathbf{h} .

The steps of the SBNIHT algorithm for the n -th time iteration are summarized in Table 2.

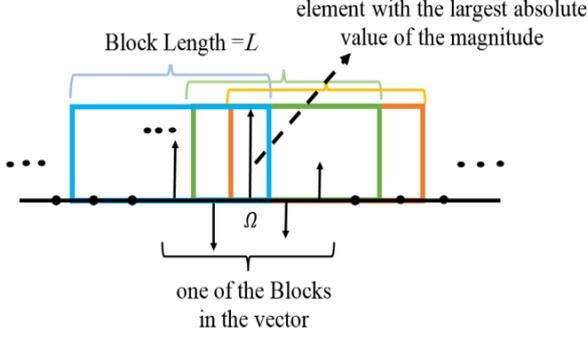


Fig. 2: Method of the finding a block in the block-sparse vector via the proposed algorithm.

Convergence Analysis

To see how the recursive update $\mathbf{h}(n)$ converges toward \mathbf{h}_o , we rewrite (20) as

$$\mathbf{h}(n+1) = (\mathbf{I} - \mu(n)\mathbf{R}(n))\bar{\mathbf{h}}(n) + \mu(n)\mathbf{r}(n) \quad (25)$$

where \mathbf{I} is the M -by- M identity matrix. Next, from (15) and by subtracting \mathbf{h}_o from both sides of (25), we can rewrite (29) as

$$\mathbf{h}(n+1) - \mathbf{h}_o = (\mathbf{I} - \mu(n)\mathbf{R}(n))(\bar{\mathbf{h}}(n) - \mathbf{h}_o) \quad (26)$$

the $\bar{\mathbf{h}}(n+1)$ is the best approximation to $\mathbf{h}(n+1)$, if we have

$$\begin{aligned} \|\bar{\mathbf{h}}(n+1) - \mathbf{h}_o\|_2 &\leq \|\mathbf{h}(n+1) - \mathbf{h}_o\|_2 \\ &= \|(\mathbf{I} - \mu(n)\mathbf{R}(n))(\bar{\mathbf{h}}(n) - \mathbf{h}_o)\|_2 \end{aligned} \quad (27)$$

Defining the vector $\mathbf{v}(n)$ as

$$\mathbf{v}(n) = \bar{\mathbf{h}}(n) - \mathbf{h}_o \quad (28)$$

and substituting this in (27), we obtain

$$\begin{aligned} \|\mathbf{v}(n+1)\|_2 &\leq \|(\mathbf{I} - \mu(n)\mathbf{R}(n))\mathbf{v}(n)\|_2 \\ &\leq \|\mathbf{I} - \mu(n)\mathbf{R}(n)\|_2 \|\mathbf{v}(n)\|_2 \end{aligned} \quad (29)$$

where $\|\mathbf{I} - \mu(n)\mathbf{R}(n)\|_2$ denotes the spectral norm of $\mathbf{I} - \mu(n)\mathbf{R}(n)$ matrix, and defined as

$$\begin{aligned} \|\mathbf{I} - \mu(n)\mathbf{R}(n)\|_2 &= \sqrt{\lambda_{\max}[(\mathbf{I} - \mu(n)\mathbf{R}(n))(\mathbf{I} - \mu(n)\mathbf{R}(n))^*]} \\ &= \sqrt{\lambda_{\max}[\mathbf{I} - \mu^2(n)\mathbf{R}(n)\mathbf{R}^*(n) - 2\mu(n)\mathbf{R}(n)]} \\ &= \sqrt{1 + \mu^2(n)\lambda_{\max}[\mathbf{R}(n)] - 2\mu(n)\lambda_{\max}[\mathbf{R}(n)]} \\ &= \sqrt{(1 - \mu(n)\lambda_{\max}[\mathbf{R}(n)])^2} \\ &= |1 - \mu(n)\lambda_{\max}[\mathbf{R}(n)]| \end{aligned} \quad (30)$$

The proposed algorithm converges in mean square sense, if we have in (29).

Table 2: SBNIHT algorithm for the n -th time iteration

Input: Maximum block length L ; length of unknown system M ; maximum number of blocks involving nonzero coefficients K ; Input vector $\mathbf{u} \in R^{M \times 1}$; desired signal d .

Initialize: $\bar{\mathbf{h}} = \mathbf{0} \in R^{M \times 1}$, $\mathbf{R}(0) = \mathbf{0} \in R^{M \times M}$,
 $\mathbf{r}(0) = \mathbf{0} \in R^{M \times 1}$, $\Lambda^{(1)} = [\mathbf{1} : KL]$

Step 1: gradient update

$$\begin{aligned} \mathbf{R}(n) &= \lambda \mathbf{R}(n-1) + \mathbf{u}(n)\mathbf{u}(n)^T \\ \mathbf{r}(n) &= \lambda \mathbf{r}(n-1) + \mathbf{u}(n)d(n)^* \\ \mathbf{p}(n) &= \mathbf{r}(n) - \mathbf{R}(n)\bar{\mathbf{h}}(n-1) \end{aligned}$$

Step 2: line search optimization

$$\begin{aligned} \mu(n) &= \frac{\mathbf{p}_{|\Lambda}^{(n)}(n)^T \mathbf{p}_{|\Lambda}^{(n)}(n)}{\mathbf{p}_{|\Lambda}^{(n)}(n)^T \mathbf{R}_{|\Lambda}^{(n)}(n) \mathbf{p}_{|\Lambda}^{(n)}(n)} \\ \mathbf{h}(n) &= \bar{\mathbf{h}}(n-1) + \mu(n)\mathbf{p}(n) \end{aligned}$$

Step 3: support set update

$$\tilde{\mathbf{h}} = \mathbf{h}(n+1);$$

for $m=1:K$

$$\Omega = \arg \max(|\tilde{\mathbf{h}}|, 1)$$

for $i=1, 2, \dots, L$

$$\begin{aligned} \tilde{\mathbf{h}}_{[i]} &= \{\tilde{\mathbf{h}}_{(\Omega+i-L)}, \tilde{\mathbf{h}}_{(\Omega+i-L+1)}, \dots, \tilde{\mathbf{h}}_{(\Omega+i-1)}\} \\ \|\tilde{\mathbf{h}}_{[i]}\|_2 &= \sqrt{\sum_{j=i-L}^{i-1} \tilde{\mathbf{h}}_{(\Omega+j)}^2} \end{aligned}$$

end

$$\Delta = \arg \max(\{\|\tilde{\mathbf{h}}_{[1]}\|_2, \|\tilde{\mathbf{h}}_{[2]}\|_2, \dots, \|\tilde{\mathbf{h}}_{[L]}\|_2\}, 1)$$

$$\Psi_m = \{\Omega + \Delta - L, \Omega + \Delta - L + 1, \dots, \Omega + \Delta - 1\}$$

$$\tilde{\mathbf{h}}_{|\Psi_m} = 0,$$

end

$$\Lambda^{(n)} = \Psi_1 \cup \Psi_2 \cup \dots \cup \Psi_K$$

Step 4: pruning

$$\bar{\mathbf{h}}_{|\Lambda}^{(n)}(n) = \mathbf{h}_{|\Lambda}^{(n)}(n), \quad \bar{\mathbf{h}}_{|\Lambda}^{(n)c}(n) = 0$$

$$\frac{\|\mathbf{v}(n+1)\|_2}{\|\mathbf{v}(n)\|_2} < 1 \quad (31)$$

Therefore, based on (29) and (30), a necessary mean-square convergence condition obtained as

$$\|\mathbf{I} - \mu(n)\mathbf{R}(n)\|_2 = |1 - \mu(n)\lambda_{\max}[\mathbf{R}(n)]| < 1 \quad (32)$$

The inequalities (36) can be expanded as

$$-1 < 1 - \mu(n)\lambda_{\max}[\mathbf{R}(n)] < 1 \quad (33)$$

Therefore, based on (33), a necessary mean-square convergence condition obtained as

$$0 < \mu(n) < \frac{2}{\lambda_{\max}[\mathbf{R}(n)]} \quad (34)$$

Complexity

In Table 3, the computational complexity of the proposed algorithm is compared with the other state-of-the-art algorithms in the literature. The comparison is carried out in terms of multiplications per adaptation step of the algorithm.

For the proposed SBNIHT algorithm, we have an increase in complexity compared with the other competing algorithms including BZA-LMS, [18] BS-PNLMS [22], BS-MPNLMS [23], and $l_{2,0}$ -SMPNLMS [24], due to the existence of the gradient $\mathbf{p}(n)$ and vector product $\mathbf{R}_{i\Lambda}(n)\mathbf{p}_{i\Lambda}(n)$.

But instead of this increasing complexity, the proposed algorithm can provide a much better solution for the block-sparse system identification problem compared to other competing algorithms.

Table 3: Comparison in terms of multiplications

Adaptive Algorithm	Complexity Order per Adaptation Step
BZA-LMS	$O(M)$
BS-PNLMS	$O(M)$
BS-MPNLMS	$O(M)$
$l_{2,0}$ -SMPNLMS	$O(M)$
BNIHT	$O(MKL)$
SBNIHT	$O(MKL)$

Results and Discussion

In this section, the proposed algorithm is compared with the algorithms BZA-LMS, BS-PNLMS, BS-MPNLMS, $l_{2,0}$ -SMPNLMS, and BNIHT in the application of block-sparse system identification.

The unknown system is a network echo path with the length $M=512$, and the adaptive filter has the same length.

In order to evaluate the tracking capability, in all simulations, we switch the echo path from the one-cluster in Fig. 3(a) to the two-clusters in Fig. 3(b) in iteration 5000.

The input vector is a white Gaussian sequence or an AR(1) signal with a pole at 0.9 or a speech signal. The background noise $\eta(n)$ is a white Gaussian process with a signal-to-noise ratio (SNR) of 30 dB.

All the results are averaged over 30 independent trials.

The normalized mean square deviation (NMSD), which is used to compare the convergence and the tracking capability performance, is defined as

$$NMSD_{(dB)} = 10 \log_{10}(\|\mathbf{h}(k) - \mathbf{h}_0\|^2 / \|\mathbf{h}_0\|^2) \quad (39)$$

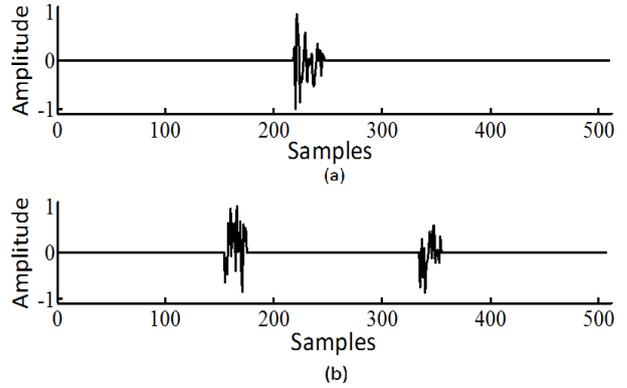
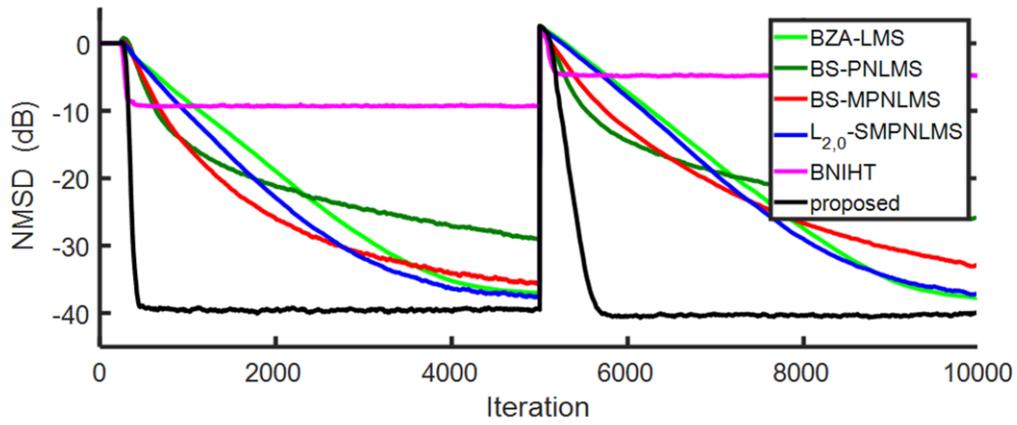


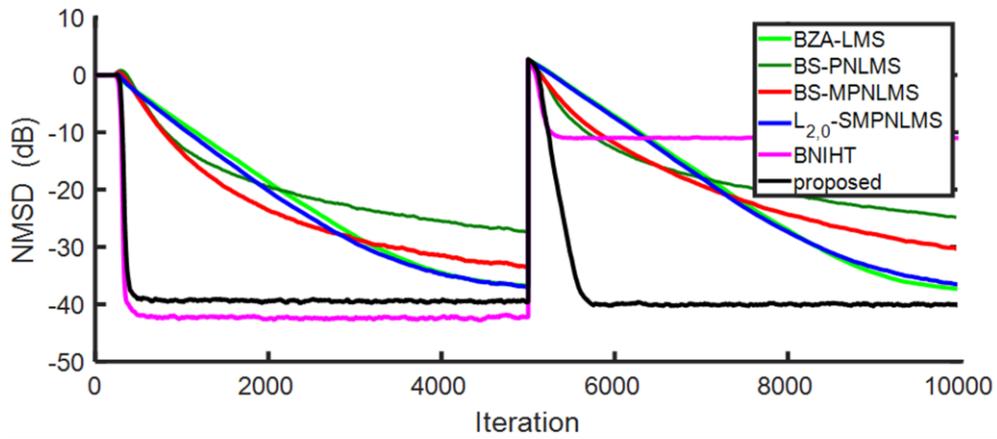
Fig. 3: Two types of measured acoustic-echo-channels as the unknown block-sparse systems, (a) one cluster, (b) two clusters.

In the following, we assume that the impulse response of the unknown system has a maximum of 2 blocks with nonzero coefficients where the maximum block length of these 2 blocks is 32, and the number of zero coefficients between two adjacent blocks are at least 32. According to the mentioned system, we consider in all simulations, our system as Fig. 3, which is one-cluster or two-clusters with a maximum block length 28. By using the above assumptions, we consider for the proposed SBNIHT algorithm, the maximum number of the blocks as $K=2$, and the maximum block length as $L=32$. Also for a better comparison, the number of the blocks for the BNIHT algorithm is considered as 2, as same as the SBNIHT algorithm. In addition, we chose the value of the parameter of the competing algorithms in all the simulations, according to the best values in their references, in such a way that all the algorithms have the same steady-state NMSD with the maximum convergence speed in achieving such a steady-state level.

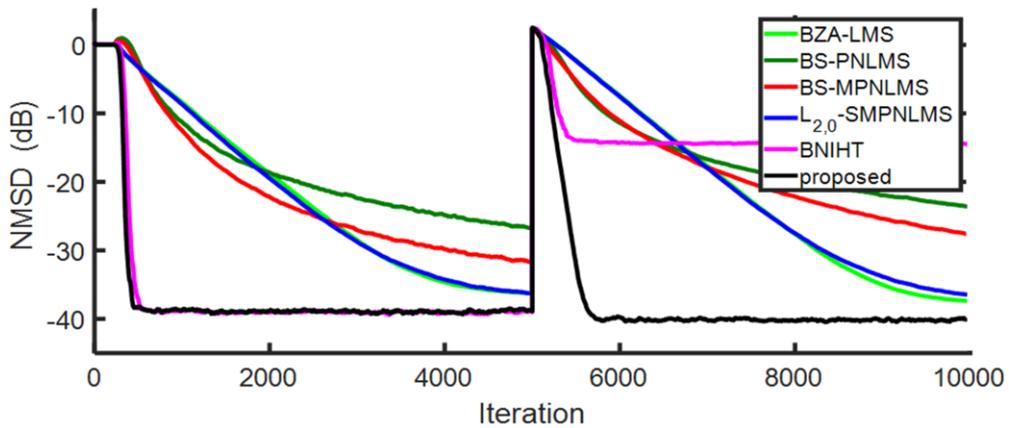
Figure 4 shows the NMSD curves of the BZA-LMS, BS-PNLMS, BS-MPNLMS, $l_{2,0}$ -SMPNLMS, BNIHT, and the proposed SBNIHT algorithms for a white Gaussian sequence input signal, by considering the different block lengths $L \in \{8,16,32\}$ for the BZA-LMS, BS-PNLMS, BS-MPNLMS, BNIHT and $l_{2,0}$ -SMPNLMS algorithms. For the fair comparisons, the value of the algorithm parameters in Fig. 4 and Fig. 5 are selected as: BZA-LMS ($\mu = 0.002$, $\delta = 0.8, \lambda = 1 \times 10^{-7}$), BS-PNLMS ($\mu = 0.8, q = 0.01$), BS-MPNLMS ($\mu = 0.02, \beta = 0.5, \delta = 0.1$), and $l_{2,0}$ -SMPNLMS ($\beta = 20, \delta = 0.01, \kappa = 0.01$). It is observed that the performance of the proposed algorithm is evidently better than other competing algorithms, in terms of convergence and tracking capability.



(a)

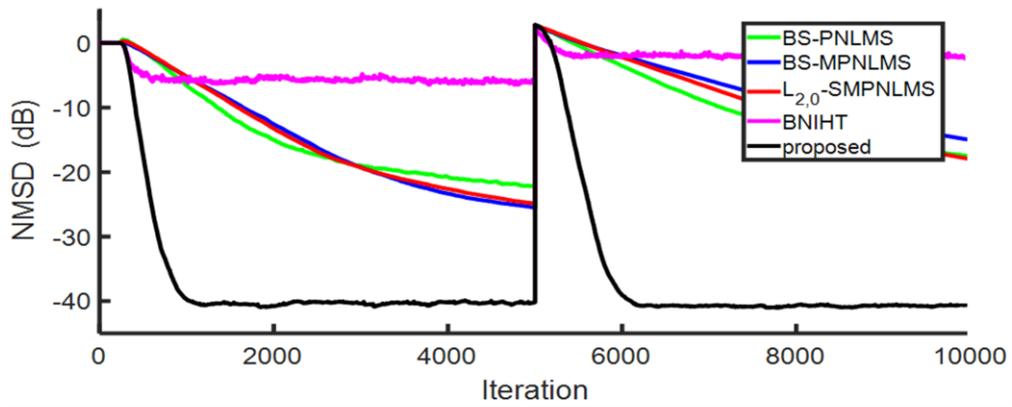


(b)

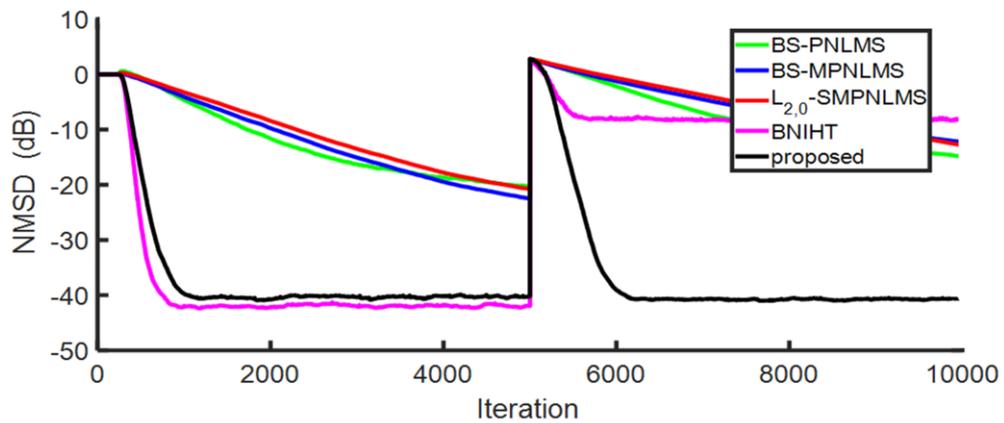


(c)

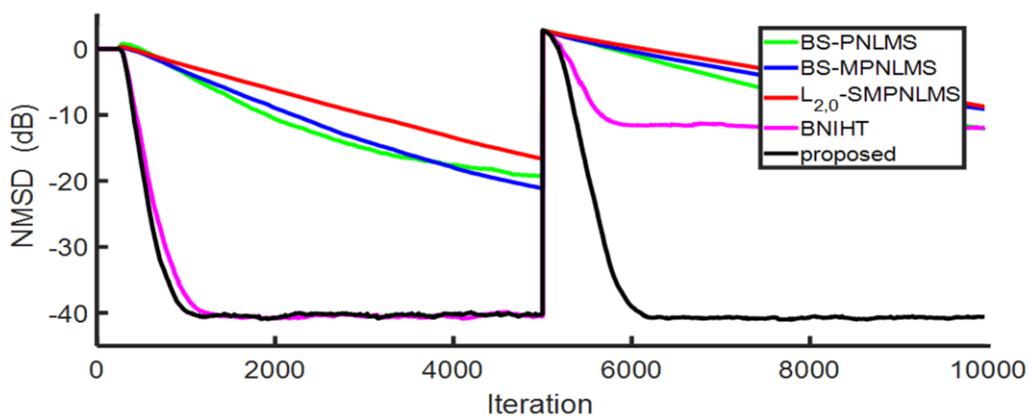
Fig. 4: NMSD learning curves of the several algorithms and the proposed algorithm with a WGN input signal for the different block length of competing algorithms BZA-LMS, BS-PNLMS, BS-MPNLMS, $L_{2,0}$ -SMPNLMS and BNIHT, (a) $L=8$, (b) $L=16$, (c) $L=32$.



(a)

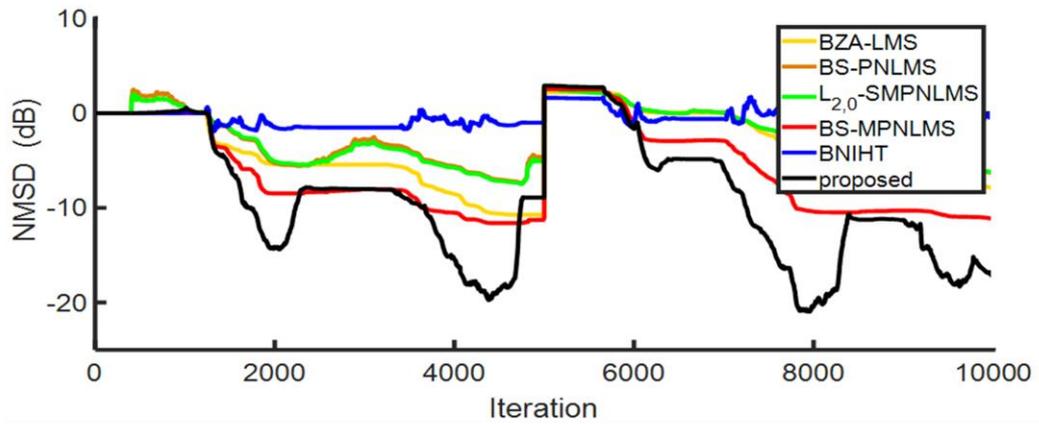


(b)

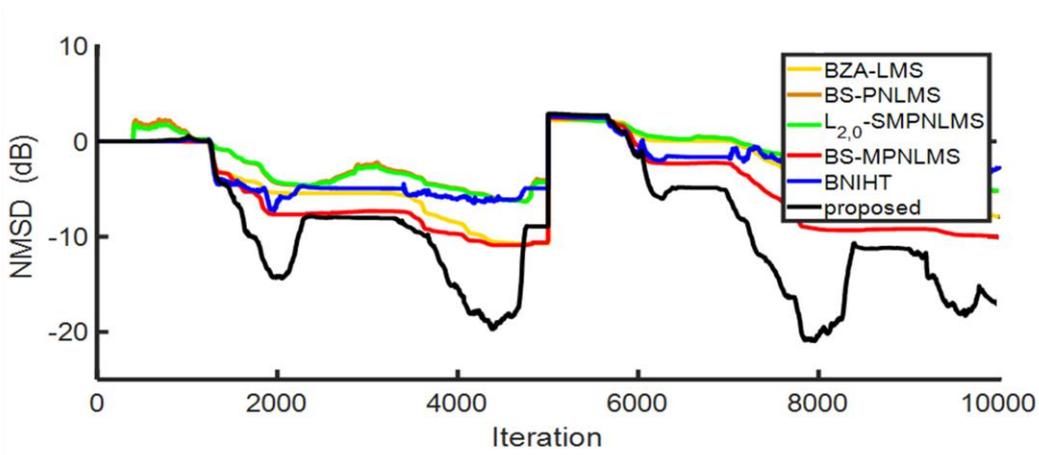


(c)

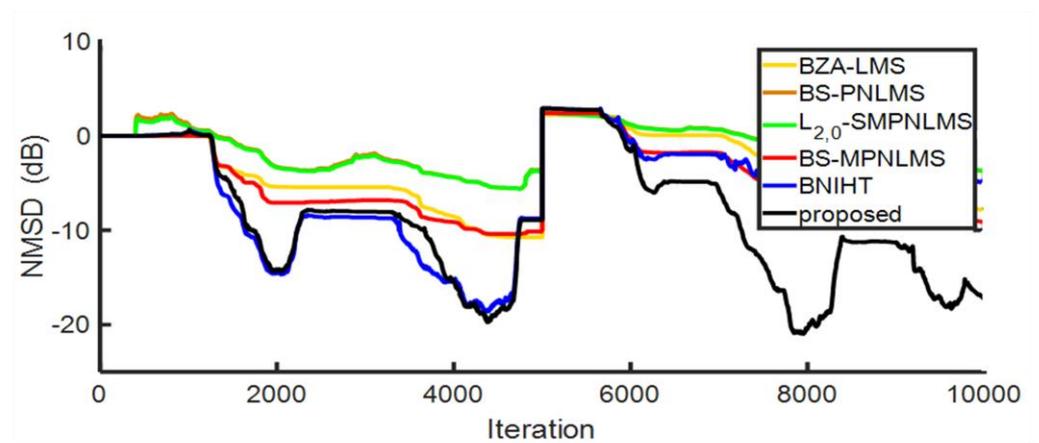
Fig. 5: NMSD learning curves of the several algorithms and the proposed algorithm with an AR(1) input signal for the different block length of competing algorithms BS-PNLMS, BS-MPNLMS, $L_{2,0}$ -SMPNLMS and BNIHT, (a) $L=8$, (b) $L=16$, (c) $L=32$.



(a)



(b)



(c)

Fig. 6: NMSD learning curves of the several algorithms and the proposed algorithm with a speech input signal for the different block length of competing algorithms BZA-LMS, BS-PNLMS, BS-MPNLMS, $l_{2,0}$ -SMPNLMS and BNIHT, (a) $L=8$, (b) $L=16$, (c) $L=32$.

Figure 5 shows the NMSD curves of the BS-PNLMS, BS-MPNLMS, $l_{2,0}$ -SMPNLMS, BNIHT and the proposed SBNIHT algorithms for an AR(1) input signal generated by filtering white Gaussian noise using a first-order $H(z) = 1/(1-0.9z^{-1})$ system.

As same as Fig. 4, we consider the different block lengths $L \in \{8,16,32\}$ for the BS-PNLMS, BS-MPNLMS, BNIHT, and $l_{2,0}$ -SMPNLMS algorithms. As same as the white input signal in Fig. 4, we can see that for the colored AR(1) input signal, the performance of the proposed algorithm is evidently better than other competing algorithms, in terms of convergence and tracking capability.

We do not consider the BZA-LMS algorithm in Fig. 5, because it shows a bad convergence performance for the colored AR(1) input signal.

Figure 6 shows the NMSD curves of the BZA-LMS, BS-PNLMS, BS-MPNLMS, BNIHT and SBNIHT algorithms for a speech input signal, by considering the different block lengths $L \in \{8,16,32\}$ for the BZA-LMS, BS-PNLMS, BNIHT, and BS-MPNLMS algorithms. For fair comparisons, the value of the algorithm parameters in Fig. 6 are selected as: BZA-LMS ($\mu = 0.1, \delta = 0.8, \lambda = 1 \times 10^{-7}$), BS-PNLMS ($\mu = 0.1, q = 0.01$), BS-MPNLMS ($\mu = 0.9, \beta = 0.1, \delta = 0.1$). It is also observed that the performance of the proposed algorithm is evidently better than other competing algorithms, in terms of convergence and tracking capability.

Also, we can see that for all the input signals, the BNIHT algorithm shows a bad tracking capability.

Conclusion

Adaptive filter algorithms have been widely used in various fields, such as system identification, channel equalization, and noise cancellation.

In many system identification scenarios for example acoustic echo path, the impulse response of the system is block-sparse.

In the block-sparse systems, the coefficients of the system are in the form of a single cluster or multi-cluster, wherein a cluster is a gathering of nonzero coefficients. To solve the block-sparse system identification problem, we use the normalized iterative hard thresholding (NIHT) algorithm that is one of the effective algorithms in the compressive sensing (CS) field, with an adaptive filter framework as a basis of the our work in this paper. The proposed algorithm named stochastic block normalized iterative hard thresholding (SBNIHT) algorithm is a new block version of the greedy NIHT algorithm with an adaptive filter framework.

The SBNIHT algorithm uses a new search method to identify the blocks of the impulse response of the unknown block-sparse system. In addition, in this paper, the necessary condition to guarantee the convergence of SBNIHT is derived. Although the proposed SBNIHT algorithm is more complex than other state-of-the-art algorithms in the literature, but Simulation results demonstrate that the proposed algorithm has better convergence and tracking capability.

Author Contributions

Z. Habibi proposed algorithm, did the simulations, interpreted the results, and wrote the manuscript. H. Zayyani corrected the proofing of the article. M.S.E. Abadi supported the article.

Acknowledgment

This work is completely self-supporting, thereby no any financial agency's role is available.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

CS	Compress Sensing
IHT	Iterative Hard Thresholding
NIHT	Normalized IHT
BIHT	Block-IHT
BNIHT	Block-NIHT
BS	Block-Sparse
SSR	Sparse Signal Recovery
SGD	Stochastic Gradient Descent
BP	Basis Pursuit
MP	Matching Pursuit
OMP	Orthogonal MP
StOMP	Stagewise OMP
CoSaMP	Compressive Sampling Matching Pursuit
ZA-LMS	Zero-Attracting Least-Mean-Square
l_0 -LMS	l_0 -norm Least Mean Square
BS-LMS	Block-Sparse LMS
ZAP	Zero-point Attracting Projection
BZAP	Block-ZAP
BZA-LMS	Block Zero Attracting LMS
Bl_0 -LMS	Block l_0 -norm LMS
RIP	Restricted Isometry Property

References

- [1] R.G. Baraniuk, "More is less: signal processing and the data deluge," *Science*, 331(6018): 717-719, 2011.
- [2] Z. Gao, L. Dai, Z. Lu, C. Yuen, Z. Wang, "Super-resolution sparse MIMO-OFDM channel estimation based on spatial and temporal correlations," *IEEE Commun. Lett.*, 18(7): 1266-1269, 2014.
- [3] Z. Zhang, Z. Pi, B. Liu, "TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise," *IEEE Trans. Biomed. Eng.*, 62(2): 522-531, 2014.
- [4] Y.C. Eldar, P. Kuppinger, H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, 58(6): 3042-3054, 2010.
- [5] R.G. Baraniuk, V. Cevher, M.F. Duarte, C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, 56(4): 1982-2001, 2010.
- [6] D. Zhang, Y. Zhang, G. Zheng, C. Feng, J. Tang, "Fast super-resolution estimation of DOA and DOD in bistatic MIMO radar with off-grid targets," *Int. J. Electron.*, 105(5): 866-881, 2018.
- [7] D.L. Donoho, Y. Tsaig, I. Drori, J.L. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, 58(2): 1094-1121, 2012.
- [8] B.X. Huang, T. Zhou, "Recovery of block sparse signals by a block version of StOMP," *Signal Process.*, 106: 231-244, 2015.
- [9] M. Stojnic, F. Parvaresh, B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," *IEEE Trans. Signal Process.*, 57(8): 3075-3085, 2009.
- [10] J. Jin, Y. Gu, S. Mei, "A stochastic gradient approach on compressive sensing signal reconstruction based on adaptive filtering framework," *IEEE J. Sel. Top. Signal Process.*, 4(2): 409-420, 2010.
- [11] J. Ren, C. Wei, L. Yu, H. Zhang, H. Sun, "Dynamic recovery for block sparse signals," *Signal Process.*, 130: 197-203, 2017.
- [12] Y. Jiao, B. Jin, X. Lu, "Group Sparse Recovery via the l0 (l2) Penalty: Theory and Algorithm," *IEEE Trans. Signal Process.*, 65(4): 998-1012, 2016.
- [13] X. Zhang, W. Xu, J. Lin, and Y. Dang, "Block normalised iterative hard thresholding algorithm for compressed sensing," *Electron. Lett.*, 55(17): 957-959, 2019.
- [14] S. He, Y. Lin, Y. "Cauchy distribution function-penalized LMS for sparse system identification," *Circ. Syst. Signal Process.*, 38(1): 470-480, 2019.
- [15] S. Jiang, Y. Gu, "Block-sparsity-induced adaptive filter for multi-clustering system identification," *IEEE Trans. Signal Process.*, 63(20): 5318-5330, 2015.
- [16] J. Liu, J. Jin, Y. Gu, "Efficient recovery of block sparse signals via zero-point attracting projection," in *Proc. 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*: 3333-3336, 2012.
- [17] C. Ye, G. Gui, S.Y. Matsushita, L. Xu, "Block sparse signal reconstruction using block-sparse adaptive filtering algorithms," *J. Adv. Comput. Intell. Intell. Informat.*, 20(7): 1119-1126, 2016.
- [18] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, 52: 1289-1306, 2006.
- [19] T.S. Alexander, "Adaptive signal processing: theory and applications," Springer Science & Business Media, 2012.
- [20] S. Foucart, "Hard thresholding pursuit: an algorithm for compressive sensing," *SIAM J. Numer. Anal.*, 49(6): 2543-2563, 2011.
- [21] T. Blumensath, M.E. Davies, "Normalized iterative hard thresholding: Guaranteed stability and performance," *IEEE J. Sel. Top. Signal Process.*, 4(2): 298-309, 2010.
- [22] J. Liu, S.L. Grant, "Proportionate adaptive filtering for block-sparse system identification," *IEEE/ACM Trans. Audio Speech Lang. Process.*, 24(4): 623-630, 2015.
- [23] Z. Jin, X. Ding, Z. Jiang, Y. Li, "An Improved μ -law Proportionate NLMS Algorithm for Estimating Block-Sparse Systems," in *Proc. 2019 IEEE 2nd International Conference on Electronic Information and Communication Technology (ICEICT)*: 205-209, 2019.
- [24] Z. Jin, Y. Li, Z. Jiang, "A Block-sparse Set-Membership PNLMs Algorithm Based on Hybrid Zero-Attraction Penalty," in *Proc. 2019 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*: 216-219, 2019.

Biographies



ZAHRA Habibi was born in 1984. She received her B.S. degree from Amirkabir University of Technology, Tehran, Iran, in 2007 and her M.S. degree from Malek Ashtar University of Technology, Tehran, Iran, in 2011. She is currently pursuing her Ph.D. degree with the Research Institute for Information and Communications Technologies, Academic Center for Education, Culture and Research, Tehran, Iran. Her research interests include adaptive filters, sparse systems identification.



applications.

HADI Zayyani was born in 1978. He received his B.Sc., M.Sc., and Ph.D. degrees from Sharif University of Technology, Tehran, Iran. He is currently with the Faculty of Electrical and Computer Engineering, Qom University of Technology, Qom, Iran. His research interests include statistical signal processing, sparse signal processing, compressed sensing, adaptive filters and their



Mohammad Shams Esfand Abadi was born in 1978. He received his B.S. degree from Mazandaran University, Mazandaran, Iran and his M.S. degree from Tarbiat Modares University, Tehran, Iran in 2000 and 2002, respectively, and his Ph.D. degree from Tarbiat Modares University, Tehran, Iran in 2007. He is currently with the faculty of Electrical and Computer Engineering, Shahid Rajaei Teacher Training University, Tehran, Iran. His research interests include digital filter theory and adaptive signal processing algorithms.

Copyrights

©2020 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.



How to cite this paper:

Z. Habibi, H. Zayyani, M.S.E Abadi, "Stochastic block NIHT algorithm for adaptive block-sparse system identification," J. Electr. Comput. Eng. Innovations, 9(1): 115-126, 2021.

DOI: [10.22061/JECEI.2020.7525.401](https://doi.org/10.22061/JECEI.2020.7525.401)

URL: http://jecei.sru.ac.ir/article_1492.html





PAPER TYPE? (Research paper, short paper, Review paper *et al.*)

Instructions and Formatting Rules for Authors of Journal of Electrical and Computer Engineering Innovations, JECEI

F. Author, S. Author, T. Author

Affiliations of the Authors: (Department, Faculty, University(Institution), City, Country)

Article Info	Abstract
<p>Article History: Received Reviewed Revised Accepted</p> <hr/> <p>Keywords: The author(s) shall provide up to 6 keywords to help identify the major topics of the paper</p> <hr/> <p>*Corresponding Author's Email Address:</p>	<p>Background and Objectives: This section should be the shortest part of the abstract and should very briefly outline the following information: 1-What is already known about the subject, related to the paper in question. 2- What is not known about the subject and hence what the study intended to examine (or what the paper seeks to present). In most cases, the background can be framed in just 2–3 sentences, with each sentence describing a different aspect of the information referred to above; sometimes, even a single sentence may suffice. The purpose of the background, as the word itself indicates, is to provide the reader with a background to the study, and hence to smoothly lead into a description of the methods employed in the investigation.</p> <p>Methods: The methods section is usually the second-longest section in the abstract. It should contain enough information to enable the reader to understand what was done, and how.</p> <p>Results: The results section is the most important part of the abstract and nothing should compromise its range and quality. This is because readers who peruse an abstract do so to learn about the findings of the study. The results section should therefore be the longest part of the abstract and should contain as much detail about the findings as the journal word count permits.</p> <p>Conclusion: This section should contain the most important take-home message of the study, expressed in a few precisely worded sentences. Usually, the finding highlighted here relates to the primary outcome measure; however, other important or unexpected findings should also be mentioned. It is also customary, but not essential, for the authors to express an opinion about the theoretical or practical implications of the findings, or the importance of their findings for the field. Thus, the conclusions may contain three elements: 1- The primary take-home message 2-The additional findings of importance 3-The perspective.</p>

Introduction

This document provides an example of the desired layout for JECEI paper and can be used as a template for Microsoft Word versions 2003 and later. It contains information regarding desktop publishing format, type sizes, and typefaces. Style rules are provided to explain

how to handle equations, units, figures, tables, abbreviations, and acronyms. Sections are also devoted to the preparation of appendixes, acknowledgments, references, and authors' biographies. For additional information including electronic file requirements for text and graphics, please refer to www.autjournal.com.

Doi:

Technical Work Preparation

Please use automatic hyphenation and check your spelling. Additionally, be sure your sentences are complete and that there is continuity within your paragraphs. Check the numbering of your graphics and make sure that all appropriate references are included.

A. Template

This document may be used as a template for preparing your technical work.

B. Format

If you choose not to use this document as a template, prepare your technical work in single-spaced, double-column format, on paper 21.6×27.9 centimeters (8.5×11 inches or 51×66 picas). Set top and bottom margins to 25 millimeters (0.98 inch) and left and right margins to about 20 millimeters (0.79 inch). Do not violate margins (i.e., text, tables, figures, and equations may not extend into the margins). The column width is 82 millimeters (3.2 inches). The space between the two columns is 6 millimeters (0.24 inch). Paragraph indentation is 4.2 millimeters (0.17 inch). Use full justification. Use either one or two spaces between sections, and between text and tables or figures, to adjust the column length.

C. Typefaces and Sizes

Please use a proportional serif typeface such as Calibri and embed all fonts. Table 1 provides samples of the appropriate type sizes and styles to use.

D. Section Headings

A primary section heading is enumerated by a Roman numeral followed by a period and is centered above the text. A primary heading should be in capital letters.

A secondary section heading is enumerated by a capital letter followed by a period and is flush left above the section. The first letter of each important word is capitalized and the heading is italicized.

A tertiary section heading is enumerated by an Arabic numeral followed by a parenthesis. It is indented and is followed by a colon. The first letter of each important word is capitalized and the heading is italicized.

A quaternary section heading is rarely necessary, but is perfectly acceptable if required. It is enumerated by a lowercase letter followed by a parenthesis. It is indented and is followed by a colon. Only the first letter of the heading is capitalized and the heading is italicized.

E. Figures and Tables

Figure axis labels are often a source of confusion. Try to use words rather than symbols. As an example, write the quantity "Magnetization," or "Magnetization, M ," not just " M ." Put units in parentheses. Do not label axes only with units. As in Fig. 1, write "Magnetization (kA/m)" or "Magnetization ($\text{kA}\cdot\text{m}^{-1}$)," not just "kA/m." Do not label axes with a ratio of quantities and units. For

example, write "Temperature (K)," not "Temperature/K." Figure labels should be legible, approximately 8- to 10-point type.

Large figures and tables may span both columns, but may not extend into the page margins. Arrange these one column figures and tables at either top or end of a page, or at the end of the paper right before the references. Figure captions should be below the figures; table captions should be above the tables. Do not put captions in "text boxes" linked to the figures. Do not put borders around your figures. Use Insert | Reference | Caption to number your tables and figures, and use Insert | Reference | Cross-reference to refer to their numbers.

Table 1: Samples of Calibri sizes and styles used for formatting a pes technical work

Point Size	Purpose in Paper	Special Appearance
9	Table text, figure text footnotes, subscripts, superscripts, references, bio, Figure caption, keywords	Table Title
10	Body text, equations, author affiliation, abstract	<i>Subheadings</i>
11		Section Titles
12	<i>Author Name</i>	

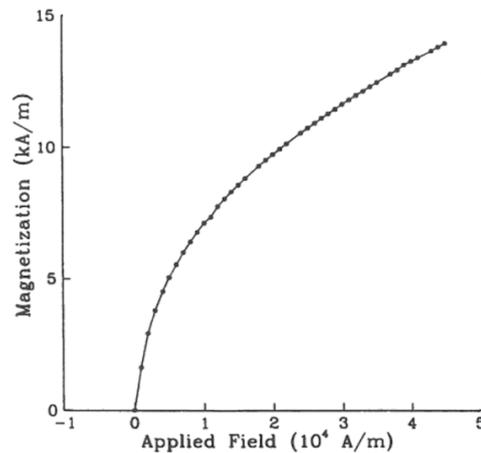


Fig. 1: Magnetization as a function of applied field. (Note that there is a colon after the figure number followed by two spaces.)

All figures and tables must appear near, but not before, their first mention in the text. Use the abbreviation "Fig. 1," even at the beginning of a sentence.

To insert images in Word, use Insert | Picture | From File.

F. Numbering

Number reference citations consecutively in square

brackets [1]. The sentence punctuation follows the brackets [2]. Multiple references [2], [3] are each numbered with separate brackets [1][1]-[2]. Refer simply to the reference number, as in [2]. Do not use "Ref. [2]" or "reference [2]" except at the beginning of a sentence: "Reference [2] shows...."

Number footnotes separately with superscripts (Insert | Footnote). Place the actual footnote at the bottom of the column in which it is cited. Do not put footnotes in the reference list. Use letters for table footnotes.

Use Arabic numerals for figures and Roman numerals for tables. Appendix figures and tables should be numbered consecutively with the figures and tables appearing in the rest of the paper. They should not have their own numbering system.

G. Units

Metric units are preferred for use in IEEE publications in light of their global readership and the inherent convenience of these units in many fields. In particular, the use of the International System of Units is advocated. This system includes a subsystem of units based on the meter, kilogram, second, and ampere (MKSA). British units may be used as secondary units (in parentheses). An exception is when British units are used as identifiers in trade, such as 3.5-inch disk drive.

H. Math and Equations

Number equations consecutively with equation numbers in parentheses flush with the right margin, as in (1). First use the equation editor to create the equation. Then select the "Equation" markup style. Write the equation number in parentheses using Insert | Caption.

Use the Microsoft Equation Editor for all math objects in your paper (Insert | Object | Create New | Microsoft Equation or MathType Equation). "Float over text" should *not* be selected.

To make your equations more compact, you may use the slash (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Use parentheses to avoid ambiguities in denominators. Number equations consecutively with equation numbers in parentheses flush with the right margin, as in (1). Be sure that the symbols in your equation have been defined before the equation appears or immediately following. Italicize symbols (*T* might refer to temperature, but *T* is the unit Tesla).

Use Insert | Reference | Caption to number equations. Refer to "(1)," not "Eq. 1" or "equation (1)," except at the beginning of a sentence: "Equation (1) is ...". Punctuate equations when they are part of a sentence, as in

$$\int_0^{r_2} F(r, \varphi) dr d\varphi = [\sigma r_2 / (2\mu_0)] \quad (1)$$

$$\cdot \int_0^{\infty} \exp(-\lambda |z_j - z_i|) \lambda^{-1} J_1(\lambda r_2) J_0(\lambda r_i) d\lambda$$

Use two column tables to locate equations and their numbers properly in one line, as follows:

$$I_F = I_B = -I_C = A^2 I_{A1} + A I_{A2} + I_{A0} = \frac{-J\sqrt{3}E_A}{Z_1 + Z_2} \quad (2)$$

where I_F is the fault current. Be sure that the border is off.

Results and Discussion

The Results section should briefly present the experimental data in text, tables or figures. Tables and figures should not be described extensively in the text.

The Discussion should focus on the interpretation and the significance of the findings with concise objective comments that describe their relation to other work in the area. It should not repeat information in the results. The final paragraph should highlight the main conclusion(s), and provide some indication of the direction future research should take.

Conclusion

As the Conclusion section is the most important element of a manuscript, so it must be more expanded scientifically and contently at least half a page length.

Example:

In this study, a forecast model was developed to determine the generation of MSW in the municipalities of the CCS, Chiapas State, Mexico. A MLR was used to obtain the forecast model with social and demographic explanatory variables. Two forecast models were presented and analyzed, with variables that met the multicollinearity test. The most important variables to predict the rate of MSW generation in the study area were the population of each municipality (XPop), the population born in another municipality (XPbam) and the population density (XPd). XPop is the most influential explanatory variable of waste generation, particularly it is related in a positive way. XPbam is less related to waste generation. XPd is the variable that least influences waste generation prediction; in addition, it can present problems of correlation with other explanatory variables. Although other variables, such as daily per capita income (XDpi) and average schooling (XAs), are very important, they do not seem to have an effect on the response variable in this study. The user of this forecast model should use model 2, since it is the one with the highest parsimony (it uses fewer variables); R^2_{adj} , MAPE, MAD and RMSE values indicated high influence on the explained phenomenon and high forecasting capacity. Additionally, it is important to mention that when using the models proposed for forecasting purposes, it is necessary to make a

transformation in the explanatory and response variables (use inverse of natural logarithm). The inferences made on the municipalities of the study area showed that, except in some municipalities, the MSW generation rate usually presented a gradual increase with respect to population growth and with respect to the number of inhabitants that were born in another entity (migration). Finally, this study can be a solid basis for comparison for future research in the area of study. It is possible to use different mathematical models such as artificial neural network, principal component analysis, time-series analysis, etc., and compare the response variable or the predictors.

Author Contributions

Each author role in the research participation must be mentioned clearly.

Example:

A. Mahboobi, B. Bagheri, and C. Ahmdi designed the experiments. A. Mahboobi collected the data. A. Mahboobi carried out the data analysis. A. Mahboobi, B. Bagheri, and C. Ahmdi interpreted the results and wrote the manuscript.

Acknowledgment

The following is an example of an acknowledgment. (Please note that financial support should be acknowledged in the unnumbered footnote on the title page.)

The author gratefully acknowledges the IEEE I. X. Austan, A. H. Burgmeyer, C. J. Essel, and S. H. Gold for their work on the original version of this document.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

Define less common abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title unless they are unavoidable.

Example:

<i>MS</i>	Multispectral
<i>SMF</i>	Spectral Matched Filter
<i>SAM</i>	Spectral Angle Mapper
<i>MSD</i>	Matched Subspace Detector
<i>OSP</i>	Orthogonal Subspace Projection

<i>CEM</i>	Constrained Energy Minimization
<i>ASD</i>	Adaptive Subspace Detector
<i>STD</i>	Sparsity Based Target Detector
<i>KSAM</i>	Kernel Based SAM
<i>DTD</i>	Difference Based Target Detection
<i>AP-CR</i>	Attribute Profile Based Collaborative Representation
<i>ROC</i>	Receiver Operating Characteristic
<i>MS</i>	Multispectral
<i>SMF</i>	Spectral Matched Filter
<i>SAM</i>	Spectral Angle Mapper
<i>MSD</i>	Matched Subspace Detector
<i>OSP</i>	Orthogonal Subspace Projection
<i>CEM</i>	Constrained Energy Minimization
<i>ASD</i>	Adaptive Subspace Detector
<i>STD</i>	Sparsity Based Target Detector
<i>KSAM</i>	Kernel Based SAM
<i>DTD</i>	Difference Based Target Detection
<i>AP-CR</i>	Attribute Profile Based Collaborative Representation
<i>ROC</i>	Receiver Operating Characteristic
<i>MS</i>	Multispectral
<i>SMF</i>	Spectral Matched Filter
<i>SAM</i>	Spectral Angle Mapper
<i>MSD</i>	Matched Subspace Detector
<i>OSP</i>	Orthogonal Subspace Projection
<i>CEM</i>	Constrained Energy Minimization
<i>ASD</i>	Adaptive Subspace Detector
<i>STD</i>	Sparsity Based Target Detector
<i>KSAM</i>	Kernel Based SAM

References

References are important to the reader; therefore, each citation must be complete and correct. There is no editorial check on references; therefore, an incomplete or wrong reference will be published unless caught by a reviewer or discussor and will detract from the authority and value of the paper. References should be readily available publications. List only one reference per reference number. If a reference is available from two sources, each should be listed as a separate reference. Give all authors' names; do not use *et al.*

Samples of the correct formats for various types of references are given below.

Periodicals:

- [1] J. F. Fuller, E. F. Fuchs, K. J. Roesler, "Influence of harmonics on power distribution system protection," *IEEE Trans. Power Deliv.*, 3(2): 549-557, 1988.

Books:

- [2] E. Clarke, *Circuit Analysis of AC Power Systems*, vol. I. New York: Wiley: 81, 1950.

Technical Reports:

- [3] E. E. Reber, R. L. Mitchell, C. J. Carter, "Oxygen absorption in the Earth's atmosphere," Aerospace Corp., Los Angeles, CA, Tech. Rep. TR-0200 (4230-46)-3, Nov. 1968.
- [4] S. L. Talleen. (1996, Apr.). The Intranet Architecture: Managing information in the new paradigm. Amdahl Corp., Sunnyvale, CA.

Papers Presented at Conferences (Unpublished):

- [5] D. Ebehard, E. Voges, "Digital single sideband detection for interferometric sensors," presented at the 2nd Int. Conf. Optical Fiber Sensors, Stuttgart, Germany, 1984.
- [6] Process Corp., Framingham, MA. Intranets: Internet technologies deployed behind the firewall for corporate productivity. Presented at INET96 Annu. Meeting.

Papers from Conference Proceedings (Published):

- [7] J. L. Alquerque, J. C. Praca, "The Brazilian power system and the challenge of the Amazon transmission," in Proc. IEEE Power Engineering Society Transmission and Distribution Conf.: 315-320, 1991.

Dissertations:

- [8] S. Hwang, "Frequency domain system identification of helicopter rotor dynamics incorporating models with time periodic coefficients," Ph.D. dissertation, Dept. Aerosp. Eng., Univ. Maryland, College Park, 1997.

Standards:

- [9] IEEE Guide for Application of Power Apparatus Bushings, IEEE Standard C57.19.100-1995, Aug. 1995.

Patents:

- [10] G. Brandli and M. Dick, "Alternating current fed power supply," U.S. Patent 4 084 217, Nov. 4, 1978.

Biographies

A technical biography for each author may be included, but without any title, as it is seen herein. It should begin with the author's name (as it appears in the byline). A photograph and an electronic file of the photo should also be included for each author. The photo should be black and white, glossy, and 3.0 centimeters (1.18 inches) wide by 3.8 centimeters (1.5 inches) high. The head and shoulders should be centered, and the photo should be flush with the left margin. The following is an example of the text of a technical biography:



Nikola Tesla (M'1888, F'17) was born in Smiljan in the Austro-Hungarian Empire, on July 9, 1856. He graduated from the Austrian Polytechnic School, Graz, and studied at the University of Prague. His employment experience included the American Telephone Company, Budapest, the Edison Machine Works, Westinghouse Electric Company, and Nikola Tesla Laboratories. His special fields of interest included high frequency. Tesla received honorary degrees from institutions of higher learning including Columbia University, Yale University, University of Belgrade, and the University of Zagreb. He received the Elliott Cresson Medal of the Franklin Institute and the Edison Medal of the IEEE. In 1956, the term "tesla" (T) was adopted as the unit of magnetic flux density in the MKSA system. In 1975, the Power Engineering Society established the Nikola Tesla Award in his honor. Tesla died on January 7, 1943.

Copyrights

©2020 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

**How to cite this paper:**

F. Author, S. Author, T. Author, "Instructions and formatting rules for authors of journal of electrical and computer engineering innovations, JECEI," J. Electr. Comput. Eng. Innovations, x(x): xxx-xxx, xxxx.

DOI:

URL: <http://jecei.srttu.edu/journal/authors.note>

