

SIMULATION AND CLASSIFICATION OF POWER QUALITY EVENTS BASED ON WAVELET TRANSFORM ENERGY USING FEATURE SELECTION METHOD

J. Azarakhsh* and Z. Raisi

Faculty of Marine Engineering, Chabahar Maritime University, Chabahar, IRAN

Published online: 18 June 2016

ABSTRACT

In this paper mathematical modeling method has been used in order to simulating of power quality events. In regarding to excellent neural network function in classification and pattern recognition works, multilayer neural network has been used for events classification. The STFT and Discrete Wavelet Transform are used for feature extraction. Acquired features of this step are decreased using advance algorithm of feature selection. Then these extracted features were given to the neural network as input. The simulation results showed that the classification precision of data caused by the algorithm of feature selection is better in compare with the precision of total data classification.

Keywords: Power quality, neural network, Wavelet transform, Feature selection.

Author Correspondence, E-mail: j.azarakhsh@cmu.ac.ir

doi: <http://dx.doi.org/10.4314/jfas.v8i2s.152>

1. INTRODUCTION

The quick expansion of using electrical and sensitive electronical equipment in industry and house consumptions has transformed power supply with high quality to a critical subject. In order to providing the power with quality, the first matter is to determine the power quality events as fast as possible.



So far for classifying the power quality events researchers have used some methods such as: three dimensional method [1]. binary matrix method [2], wavelet transform based on method [3], learning machine method [4], fuzzy laws definition method [5], probable neural network method [6], and modular neural network method [7] which in the upon method S transform has been used for feature extraction and also this method has used from modular neural network.

In all power quality classification methods which have been mentioned, the mathematical model has been used for simulating of events. Also the number of events which has been recognized is in domain of 5 or 6 events. Also at the current paper, mathematical model has been used in order to simulating of events. Mathematical model will provide the possibility which we can simulate multiple events. In practical, possessing of multiple events will help that classifier is going to has more power of classification.

There are eight classified events in this paper and in comparison with other papers there are more separate events in number. After simulating the events, Wavelet Transform and STFT will be derivate from each signal. After processing the data of upon conversions, these data will be entered to neural network for classification. In order to gain better precision of classification we are going to use new feature called Range and wavelet transform energy which these feature have been extracted from wavelet transform of signals. The block diagram of Functional steps of neural classifier is showed in figure 1.

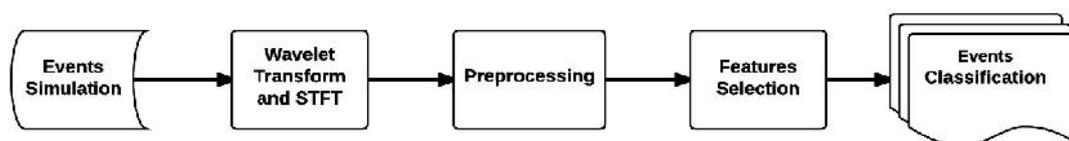


Fig.1. Functional steps of neural classifier

2. BACKGROUND STUDY AND LITERATURE REVIEW

This section presents a brief review of considered feature extraction methods includes: wavelet, variance and feature selection which are used in the proposed method and then reviews the basic concepts of the feed-forward back propagation neural network.

2.1. Wavelet Transform

Wavelet Transform is defined as one nonlinear operation used in analysis of non-stationary signals in order to extract relevant information related to variations in frequency and time domains. Wavelet is a waveform which has a limited period of time and also possesses average value of zero. With a simple comparison between wavelet and sinusoidal waveform which Fourier transform is based on this, we can declare that sinusoidal waveform does not have any limited period of time and also this is a smooth waveform on the other hand wavelet has a limited period of time and can be existed in non-concurrent shape [8].

Wavelet transform is consisted of breaking a signal into scaled and shifted forms from mother wavelet. With a simple survey to the sinusoidal waveform and wavelet signal we will be able to see that for simulating a signal with quick changes, using a non-concurrent wavelet is better than a sinusoidal smooth waveform. The Discrete Wavelet Transform is defined as:

$$c(a,b) = \int_R s(t) \frac{1}{\sqrt{a}} \mathbb{E} \left(\frac{t-b}{a} \right) dt \quad (1)$$

where, $s(t)$ represents a signal which the wavelet transform will be taken from and $\mathbb{E}(t)$ is the mother wavelet and also:

$$2^j, b = k2^j, (j, k) \in \mathbb{Z}^2 \quad (2)$$

where j is the level of signal decomposition and k shows the time separately. $c(a,b)$ is actually a coefficient which is going to show the level of similarity of transferred scaled mother wavelet with main signal. At any level of wavelet transform, A_j and D_j coefficients will be produced according to the following equations:

$$D_j(t) = \sum_{k \in \mathbb{Z}} c(j, k) \mathbb{E}_{j,k}(t) \quad (3)$$

$$\mathbb{E}_{j,k}(t) = 2^{-j/2} \mathbb{E}(2^{-j}t - k), j \in \mathbb{Z}, k \in \mathbb{Z} \quad (4)$$

$$A_J = \sum_{J > j} D_j \quad (5)$$

Where A_j is an approximation of main signal and D_j shows the detail of signal.

One of the most fundamental matter in wavelet transform is to choose the mother wavelet. One kind of wavelets is Daubechies wavelet which has ten models. Other mother wavelets are

Meyer, haar and Morlet. At the figure 2 you can see two kinds of different wavelets [8].

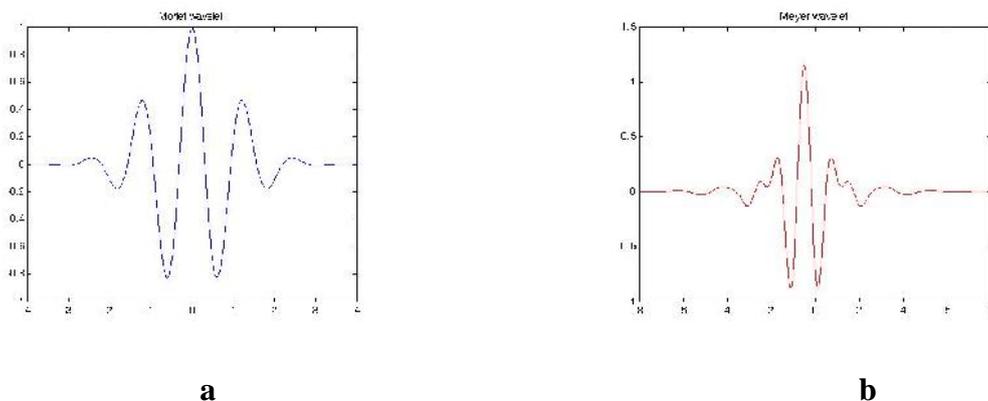


Fig.2. (a) Morlet wavelet, (b) Meyer wavelet

In this conversion there are two low and high pass filters which each of wavelets has its own special filter. For calculating the wavelet transform, regarding to the level of signal decomposition; we are going to use the detail of all decomposition level with signal approximation at the last level which this procedure will be seen in the following figure.

According to the figure 3 for a signal which has been analyzed for three levels, the energy which is belong to CD_1, CD_2, CD_3, CA_3 will be obtained by the following energy equation:

$$E_a = \sum_{n \in Z} S^2[n] \tag{6}$$

The energy which has been calculated by this method will be given to neural network as a feature.

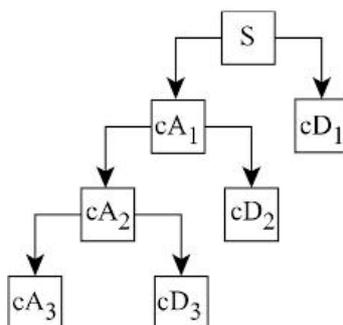


Fig.3. Wavelet decomposition tree with three levels of approximation and details coefficients

2.2. Variance

Variance is one of the statistical tools that we are going to use it for determining the scale of data dispersal. The following equation is used for calculating of variance in a vector with N samples:

$$\dagger = \sqrt{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2\right)} \quad (7)$$

where \bar{x} is the average of data and defined as :

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (8)$$

2.3. Feature Selection

For feature selection there are two methods called sequential forward selection and sequential backward selection. At sequential forward method first we select an empty matrix then we add the features to that one by one and we keep doing this until more increase of features will have no effect on regression increase. The linear equation of regression is:

$$y = S_0 + S_1x_1 + S_2x_2 + S_3x_1x_2 + S_4x_1^2 + S_5x_2^2 + V \quad (9)$$

In this equation, the inputs and fixed coefficients are used for y approximation. The procedure of sequential forward selection method is showed in figure 4.

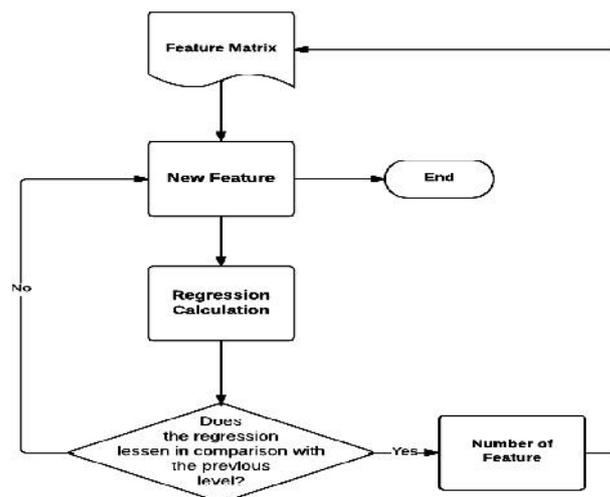


Fig.4. SFS Flowchart

In this paper in order to select the features, the sequential forward selection method has been used.

2.4. Neural Network

Back propagation network is a multilayer network with non-linear shift function which uses from a widrow-hoff rule. We use input vector and target vector for approximation of function and recognition of equation between input and output [8], [9].

Feed forward networks often have one or some hidden layer which include neurons with sigmoid shift function and use from a linear output layer. In these networks each of neurons are connected with the next layer by weights that these weights will be updated by two methods. In the first method by applying any new input, weights will be updated and at the second method inputs will be applied all together and weights are going to be regulated only once. In the figure 5 feed forward neural network construction will be seen with a hidden layer. The feed forward network can be used in order to approximation of functions, pattern recognition or pattern classification.

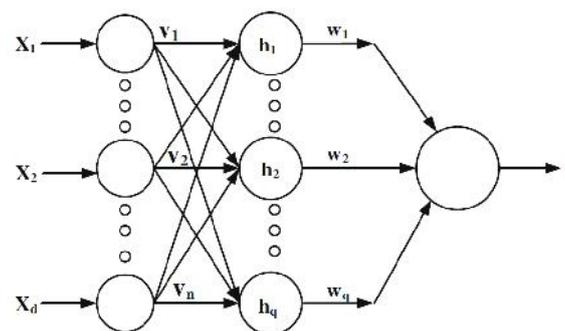


Fig.5. Feed forward neural network construction

A repetition of this algorithm is as the following equation:

$$X_{K+1} = X_K - r_K g_K \quad (10)$$

Where X_K is the current vector of weights of biases, g_K is current slope and r_k is the learning speed.

3. THE PROPOSED SYSTEM

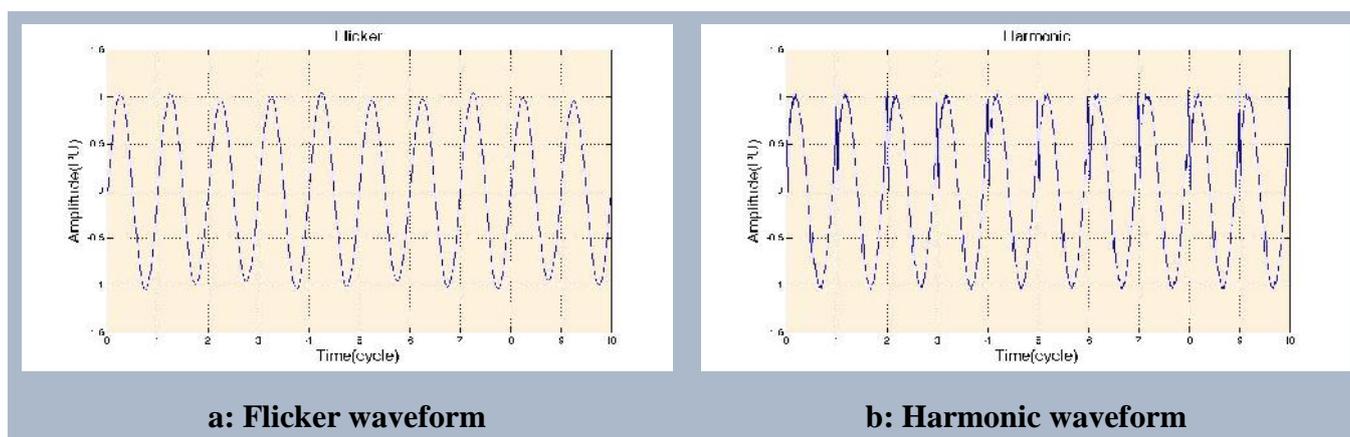
In this section the proposed method that used in this paper are described in continues.

3.1. Events Simulation

In this paper for events simulation' the mathematical equations are used [10]. The simulation of mathematical equation has been done by MATLAB. At the current paper 8 events includes: impulse, transient, flicker, harmonic, voltage sag, voltage swell, voltage notch and interrupt have been surveyed and classified. For the proper cover of events which may occur at any second with any domain in the power systems, 300 samples have been produced with different domain and phases for each of events which this work matters very much in the classification universality. Events have been simulated in 10 periods of alternation which every period includes 64 sampled points and maximally we possess 300*640 matrix for each of events. Sampling frequency of signals is 3.2 kHz. This frequency is selected regarding to Shannon sampling theory. This theory expresses that the sampling frequency should be at least two times bigger than the biggest existing frequency in signals.

3.2. Features Extraction

In the classification procedure, using the features which will produce the best result with the least calculation work is so important. In this step we use to conversions called STFT and Discrete Wavelet Transform in order to features extraction. In figure 6, we used a sample of simulated signals at MATLAB perimeter in order to applying wavelet transforms and STFT. Relevant differences to the wavelet transform of events, will allow using this transform for the events classification.



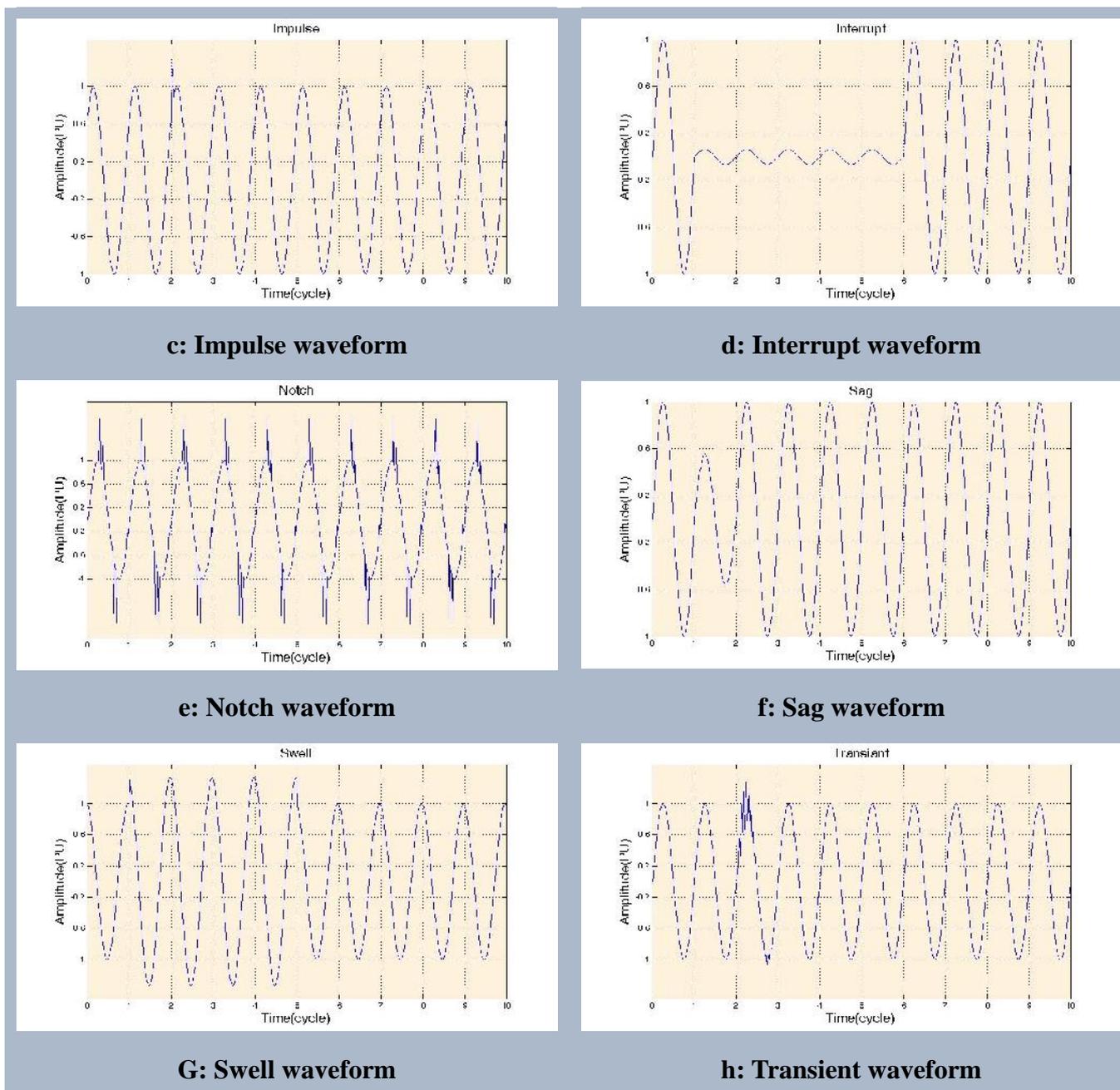
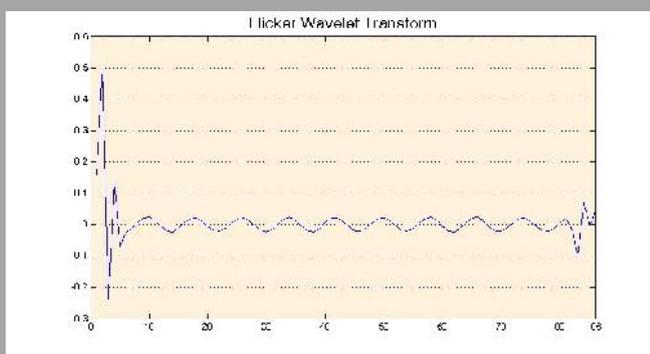
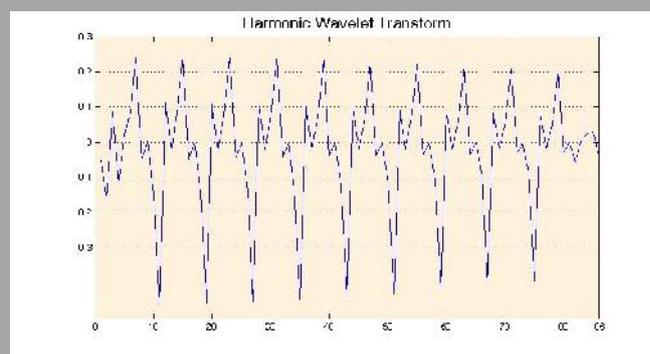


Fig.6. Power quality events

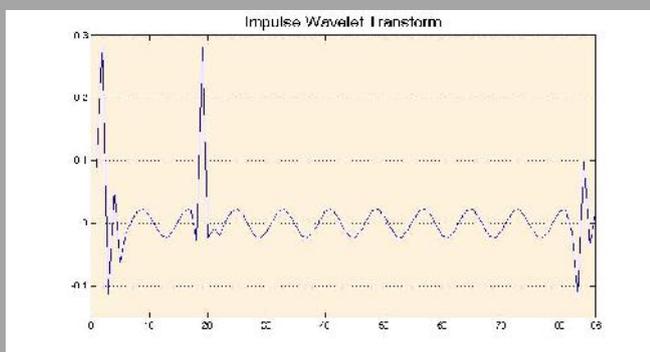
In figures 7 and 8, wavelet transform and STFT transform regarding to PQ events have been shown. As you can see in the figure7, wavelet transform will show initial and end time of an event so well. This wavelet transform feature is so important in classification work. In Discrete Wavelet Transform, db4 wavelet has been used as mother wavelet. The primary reason of selecting the wavelet transform for features extraction is that. This conversion will show the initial and end location of severe frequency changes which means the initial and sequel point of a power quality event so well.



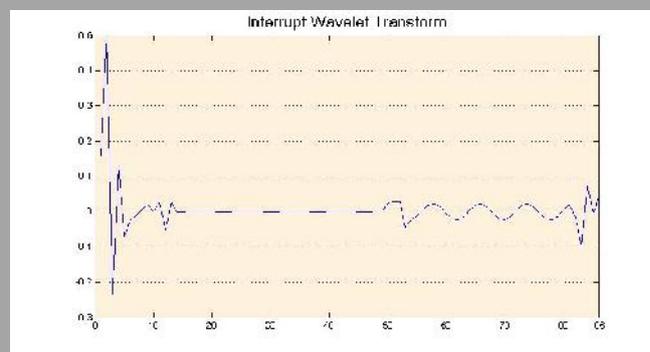
A: Flicker wavelet transform



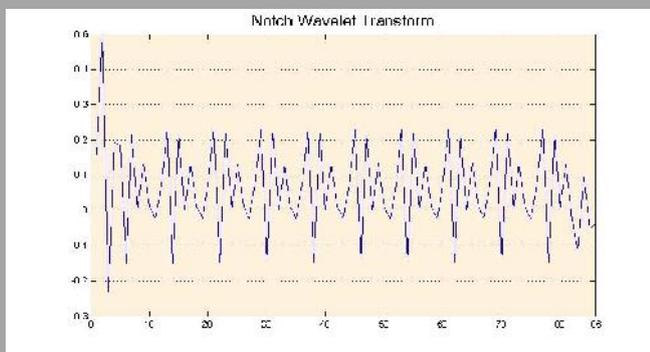
B: Harmonic wavelet transform



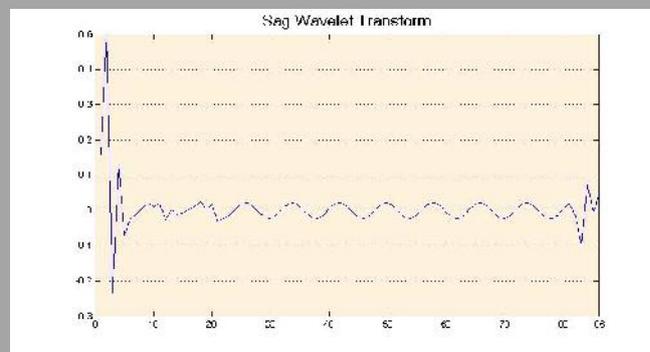
C: Impulse wavelet transform



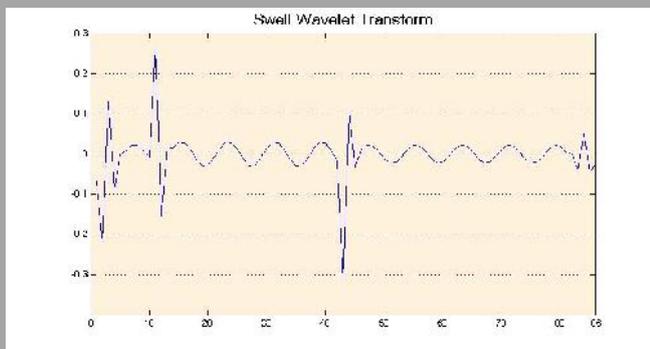
D: Interrupt wavelet transform



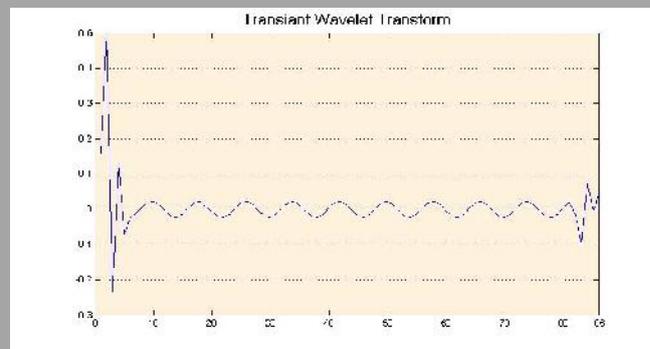
E: Notch wavelet transform



F: Sag wavelet transform

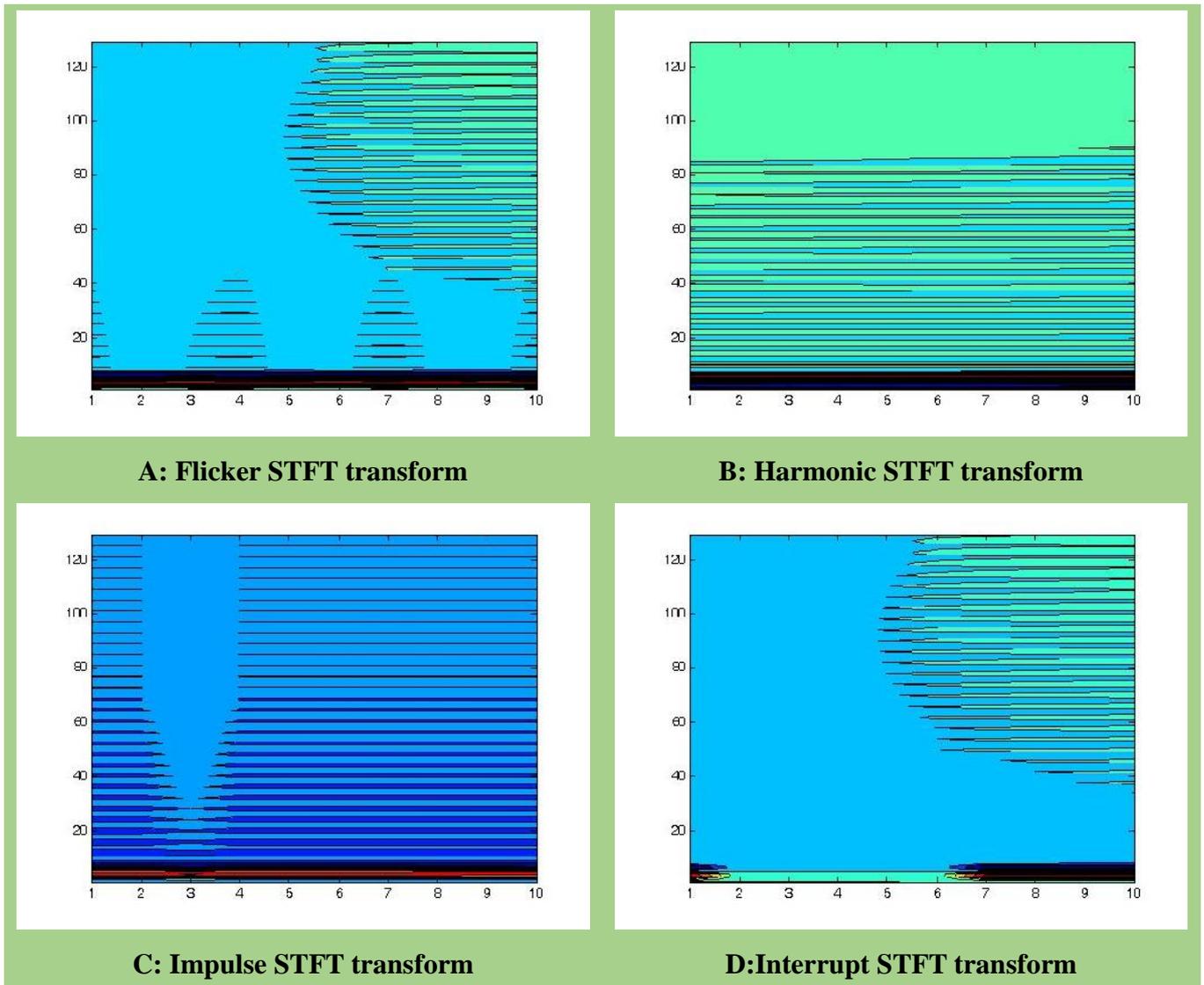


G: Swell wavelet transform



H: Transient wavelet transform

Fig.7. Wavelet transform for PQ events



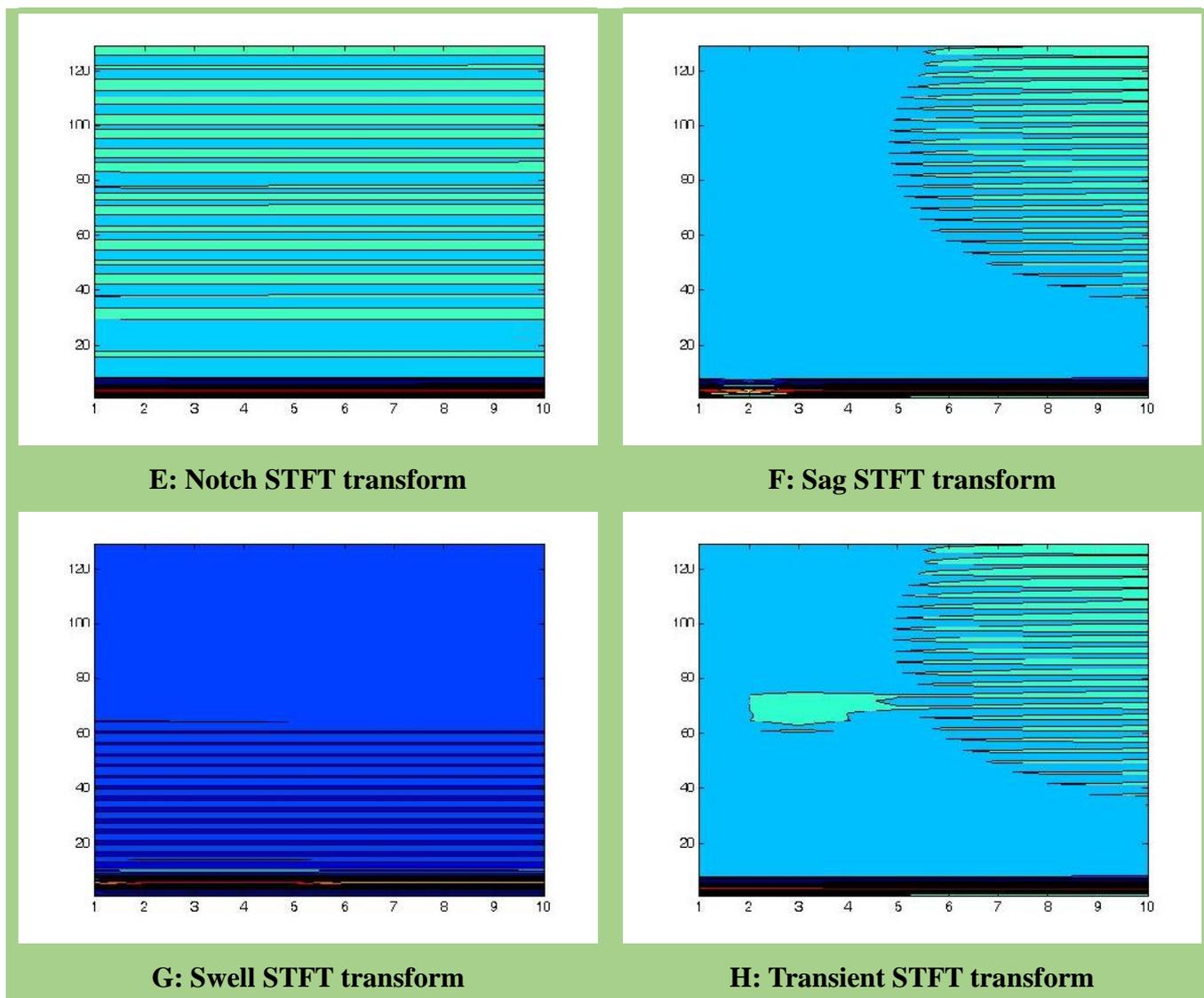


Fig.8. STFT for PQ events

The selection’s reason of db4 wavelet as the mother wavelet is that, use generality from this wavelet and its good performance in classification of power quality events. We have Separate Fourier transform as:

$$v^n [K] = \sum_{i=0}^{N-1} v[i + (n-1)N] e^{-j(2\pi ki / N)} \tag{11}$$

Where, $v[i]$ shows input signal ($i=0, 1, 2... L-1$); L is input signal length and N is the plurality of samples in a time window. We use n in order to show the number of alternation. In figure 8, Contour diagram in MATLAB has been used in order to representing STFT transform.

Extracted features for classification are defined as:

$$\alpha_n = \text{angle}(v^n[1]) - \text{angle}(v^1[1]) \quad (12)$$

$$THD_n = \frac{\sqrt{\sum_{k=2}^{\text{int}(N/2)} \{abs(v^n[k])\}^2}}{v^1[1]} \quad (13)$$

$$TS_n = \frac{\sqrt{\sum_{k=2}^{10} \{abs(v^n[k])\}^2}}{v^1[1]} \quad (14)$$

$$OS_n = \text{root}(v_{miss}^s) \quad (15)$$

$$RN = \text{root}(v_{rms}^s - \text{mean}(v_{rms}^s)) \quad (16)$$

$$N_n = \text{peak}(abs(WC^s)) \quad (17)$$

$$ED_j = \sum_{m=1}^1 CD_j^2[m] \quad (18)$$

$$\text{Range} = \text{Max}(WC^s) - \text{Min}(WC^s) \quad (19)$$

Where, wc^n is the wavelet coefficients regarding to a period and wc^s is one array of $wc^n[k]$ which $k=1, 2, \dots, L_e$ and L_e is the length of wc^n .

The terms v_{rms}^n and v_{miss}^s are defined as follow: $V_{miss}^s = V_{miss}[i]$, $i=0,1,2,\dots,L-1$, is the signal's length. In these equations j shows the signal analysis level and a shows the matrix length of wavelet coefficients.

$$v_{miss}[i] = v[i] - \frac{2}{N} \times abs(v^1[1]) \times \cos\{\text{angle}(v^1[1]) + 2f[i-1]/N\} \quad (20)$$

Where v_{rms}^s is an array which has been formed from v_{rms}^n elements, which n will be varied from 1 to 10.

v_{rms}^n is defined as:

$$v_{rms}^n = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} v^2[i + (n-1)N]} \quad (21)$$

Range is a new feature which has not used in other papers. This feature will provide the range of wavelet coefficients changes which is relevant to each signal. The relevant equation of this

feature is defined in equation 18. Figure 9 shows the difference of Range feature for transient and impulse event.

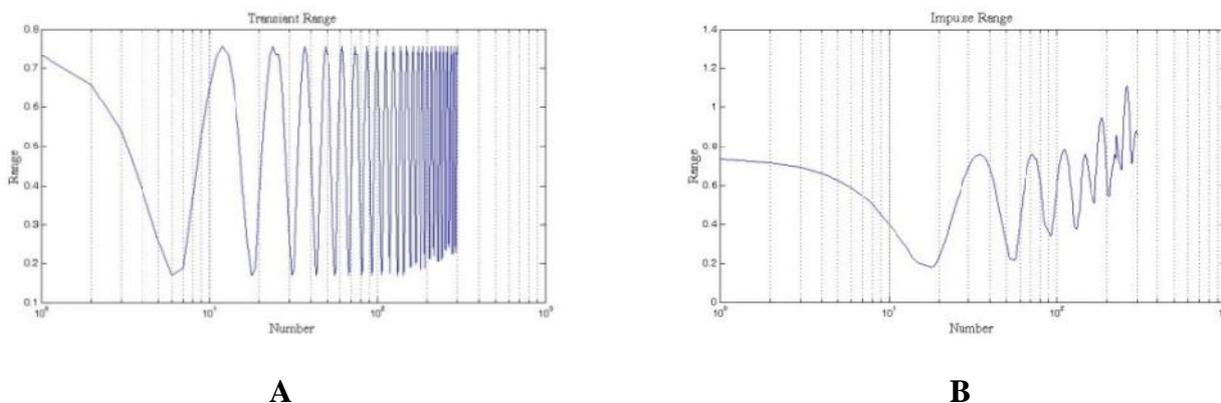


Fig. 9. Range feature in (A) transient (B) impulse

Since in power quality recognition operation, recognition’s speed is so important beside precision, Range feature will cause the increase of recognition’s speed with less calculation work in comparison with other features.

3.3. Classification by Neural Network

After extracting the desired features, first process operation will be conducted on these data in two steps. In the first step, we omit the lines which have fixed quantities and at the second step we substitute initial data with their differences with their variance. Initial process cause that the applied data to the neural network will have a standard range.

$$\begin{bmatrix} b_{1,1} & b_{1,2} & \cdot & \cdot & b_{1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{m,1} & b_{m,1} & \cdot & \cdot & b_{m,n} \end{bmatrix} \longrightarrow \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & \cdot & a_{1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m,1} & a_{m,1} & \cdot & \cdot & a_{m,n} \end{bmatrix}$$

Fig.10. Data transformation

As we can see at figure 10, first we will transform A matrix to B matrix by a transform. In this transform we have:

$$b_{i,j} = a_{i,j} - \dagger_i \tag{22}$$

Where σ_i is variance of I^{th} row.

After this step we will do the feature selection by sequential forward method. Implementation of feature selection will cause the decrease of inputs to neural network which this is going to decrease the calculation time considerably and it causes the increase of recognition speed. By using this method, we decrease features from 56 instances to 32 features. These features will be given to the neural network as input which this is a 32×2400 matrix that shows, 300 samples (300×8) have been provided from each event. We need an aim matrix in order to instruct the neural network, regarding to having 8 events for classification, also this is an 8×2400 matrix. Neural network which performs classification work, is a multilayer network with 2 hidden layers and an output layer. In first layer TANSIG function, at second layer LOGSIG function and at the output layer PURELIN function have been used. For achieving the number of proper layers, we used test and fault method as it gives the best output from the point of classification precision. After the neural network instruction and achieving the desired result, we compared the classification result with the mode which data have been given to the neural network directly. For the neural network training, 270 instructional samples and 15 samples for the test have been used; also 15 remaining samples have been used for the validation. We use from the random selection method in order to create each of sample groups.

4. RESULTS

We gave data produced by mathematical modeling to the neural network after initial process and proper feature selection which you can see events classification results in table 2. On the condition for classification, we use data which the feature selection has not been performed on them, classification precision will reach to 98.25 percent. Classification precision of selected data by the feature selection method is 99.17 percent which this possesses suitable precision in comparison with similar researches. After surveying the effect of features selection on the increase of neural network precision, now it's time to compare the achieved results of neural network classification with other methods. Based on table 2 the least of classification precision is for interrupt and after that voltage sag and this is for appearance similarity of

these two events.

Table 1. The results produced by synthetic data classification using neural network

	impulse	interrupt	swell	sag	notch	transient	harmonic	flicker	accuracy%
impulse	300	0	0	0	0	0	0	0	100
interrupt	0	300	0	16	0	0	0	0	94.67
swell	0	0	300	0	0	0	0	0	100
sag	0	4	0	296	0	0	0	0	98.67
notch	0	0	0	0	300	0	0	0	100
transient	0	0	0	0	0	300	0	0	100
harmonic	0	0	0	0	0	0	300	0	100
flicker	0	0	0	0	0	0	0	300	100
overall									99.17

4.1. Results' Comparison

In Table 2 you can see results' comparison produced by events classification after features selection with unprocessed. In this paper based on conducted studies we tried regarding to incidents which is going to occur on the real power network, performed simulation and classification would be closer to reality.

Table 2. Result's comparison of 2 data models

	Results of selective data%	Results of unprocessed data%
impulse	100	93.67
interrupt	94.67	96.33
swell	100	99.33
sag	98.67	97
notch	100	100
transient	100	99.67
harmonic	100	100
flicker	100	100
overall	99.17	98.25

In table 3 the results of this paper with similar works have been compared while the number of classified data in this paper is more than other works.

Table 3. Results' comparison of paper with similar works

	Paper results	Ref. [6]	Ref. [7]	Ref. [10]
impulse	100	-	-	95
interrupt	94.67	100	-	100
swell	100	98	98	100
sag	98.67	100	98	100
notch	100	-	-	100
transient	100	-	73	98.5
harmonic	100	98	100	97.5
flicker	100	100	100	97
Overall	99.17	-	96.29	98.5

Regarding to the performed comparison, you can see that achieved results of this paper have higher precision in comparison with similar methods. The results of this table have been achieved by data classification produced by mathematical model. Since in paper [6], also some of Synthetic events have been classified; we skipped bringing total precision word.

5. CONCLUSION

Achieved results show that neural network of this paper possesses high precision in comparison with similar works. For the improvement of classification results, we can use from other algorithms which exist for the feature selection as the classification precision will be increased. One of the feature selection methods is sequential backward method which the features selection can be done by this method.

6. REFERENCES

- [1] Pires V, Amaral T, and Martins J, "Power quality disturbances classification using the 3-D space representation and PCA based neuro-fuzzy approach," *Expert Syst. Appl.*, 2011.
- [2] Nguyen T and Y Liao, "Power quality disturbance classification utilizing S-transform and binary feature matrix method," *Electr. Power Syst. Res.*, 2009.
- [3] Liao Y and Lee J, "A fuzzy-expert system for classifying power quality disturbances," *Int. J. Electr. power energy ...*, 2004.
- [4] Moravej Z, Banihashemi S, and Velayati M, "Power quality events classification and recognition using a novel support vector algorithm," *Energy Convers. ...*, 2009.
- [5] Salem M, Mohamed A, and Samad S, "Rule based system for power quality disturbance classification incorporating S-transform features," *Expert Syst. Appl.*, 2010.
- [6] Moravej Z, Abdoos A, and Pazoki M, "An intelligent method for detection and classification of power quality events," 2012.
- [7] Cesar D "Automatic power quality disturbances detection and classification based on discrete wavelet transform and artificial intelligence," *Distrib. Conf.*, 2006.

-
- [8] Mostafa K, *Neural Network in MATLAB*. Tehran, 2007.
- [9] Guide M. U. S., “The Mathwork. Inc.” p. 5,333, 2010.
- [10] Hooshmand R and Enshae A, “Detection and classification of single and combined power quality disturbances using fuzzy systems oriented by particle swarm optimization algorithm,” *Electr. Power Syst. Res.*, 2010.

How to cite this article:

Azarakhsh.J and Raisi.Z., Simulation and Classification of Power Quality Events Based On Wavelet Transform Energy Using Feature Selection Method. J. Fundam. Appl. Sci., 2016, 8(X), 970-987.