**Huseyin ERISTI[1], Yakup DEMIR[2]**

Tunceli University (1), Firat University, Turkey (2)

# The Feature Selection based Power Quality Event Classification using Wavelet Transform and Logistic Model Tree

***Abstract***. *This paper presents a new power quality event classification technique using wavelet transform and logistic model tree. The proposed method uses the samples of three cycle duration of three line voltage of power quality events. The features of these samples are obtained by using the wavelet transform and a few different feature extraction techniques. The sequential forward selection method based a feature selection process is done to ensure good classification accuracy by selecting 20 better features from all 90 features generated from the wavelet transform coefficients. The obtained features are used to train a single logistic model tree. The feasibility of the proposed algorithm has been tested using real life power quality events. The result indicates that the feature selection based proposed method reliably classifies all types of power quality events with high accuracy.*

***Streszczenie.*** *W artykule zaproponowano nową metodę oceny jakości energii wykorzystującą transformatę falkową i logistyczny model drzewa. W metodzie analizuje się trzy cykle w trzech liniach napięcia. Możliwa jest klasyfikacja 90 zdarzeń i wybranie 20 typowych cech. (**Selekcja cech bazująca na klasyfikacji jakości energii z wykorzystaniem transformaty falkowej i modelu drzewa**)*

**Keywords:** Power quality events, Wavelet transform, Feature selection, Logistic model tree
**Słowa kluczowe:** jakość energii, transformata falkowa, selekcja cech

## Introduction

Recently, power quality (PQ) has become a significant issue for both utilities and customers due to the demand for clean power. The PQ problems can originate the consequences of the increasing use of solid state switching devices, nonlinear and power electronically switched loads, unbalanced power systems, lighting controls, computer and data processing equipment as well as industrial plant rectifiers and inverters. On the other hand, a number of causes of transients such as lightning strokes planned switching actions in the distribution or transmission system, self-clearing faults or faults cleared by current-limiting fuses can reveal the PQ problems on the power system. A PQ problem usually involves a disturbance in the voltage or current, such as voltage dips and fluctuations, momentary interruptions, harmonics and oscillatory transients causing failure or mal-operation of the electrical equipments. These failures might trip any protection device initiating a short interruption to the supplied power. Excess current produced by transients may lead to complete damage to system equipment during the transient period. Moreover, if such disturbances are not mitigated, they can lead to failures or malfunctions of various sensitive loads in power systems and may be very costly.

The classification of a PQ problem is an important issue for operating and protection of the power system because a large class of events is due to the normal operation of power systems, and these events should not cause nuisance tripping of protection equipment in the network. Therefore, there is a growing need to develop PQ monitoring techniques that can classify the potential sources of disturbances. The literature is rich in terms of proposals for the classification of PQ problems. In the literature, the proposed methods can be roughly divided into two groups. First group is called as classification of PQ disturbance [1-7]. In this group, the detected disturbances are classified in a number of typical classes such as voltage sags, voltage swells, and interruptions, etc. In second group of the classification called event classification, the underlying causes of disturbances such as faults, capacitor switching, and transformer energizing are classified [8-13].

The classical major steps for classification of both PQ events and PQ disturbances are feature extraction and classification that constitute a pattern recognition process. Feature extraction is generally called upon when there is a need to extract specific information from the raw data, which typically in power systems are the voltage and current waveforms. The feature extraction of signals can be performed by direct techniques, such as the RMS value [10] of the raw samples, or transformation techniques, such as the Fourier transform [1], the wavelet transform (WT) [4,9] and S-transform [2]. In the classification step, feature vectors that are obtained from the transformation process is applied the classifier algorithm, such as artificial neural network [6], support vector machines [9,10], expert system [11] and fuzzy-expert system [1]. In order to attain higher classification performance, a proper feature set is obtained from feature extraction step. As can be seen from above PQ event classification studies, the created feature sets have been also extracted using only one feature extraction technique. Moreover, these features have not able to provide additional information about the characteristics of signals. As a result, there is a big need to add the new step that is feature selection for classification of PQ events.

This paper investigates a new classification approach that selecting the efficiency characteristics of different feature sets for the classification of PQ events. In this approach, it is essential to select and adequate features that can recognize the main characteristics of signal and reduce its data size. The main contribution of this paper is to present a novel scheme combined of feature extraction and feature selection for obtaining adequate features of PQ events. For this purpose, firstly, the WT is applied to the three phase PQ event signals. Then, an optimal feature vector is created by using feature selection approach after the feature sets are obtained by different extraction techniques to the wavelet coefficients of all decomposition levels of the event signals. The features extracted from WT are given to the logistic model tree (LMT). From the experimental results, it is found that the proposed algorithm classifies the PQ events more effectively.

The remaining part of the paper is organized as follows. In Section 2, the real life PQ event data is explained. In Section 3, it is given several brief definitions of WT. Theory of feature extraction and selection is given in Section 4. A short review to LMT is presented in Section 5. Section 6 contains the experimental study of the proposed algorithm. Finally, conclusions are discussed in Section 7.

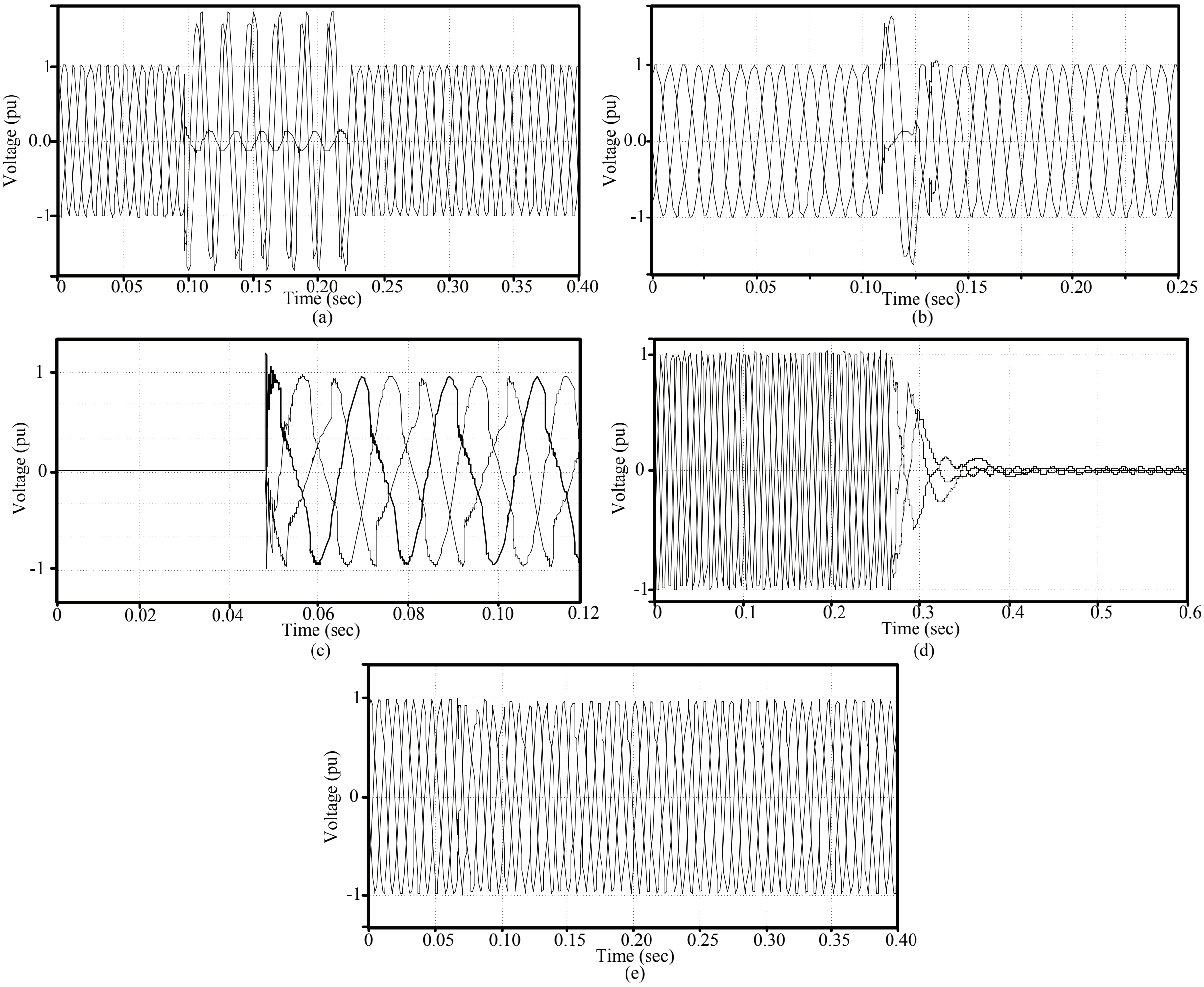

Fig. 1. The event waveforms and RMS values of: (a) fault, (b) Self-extinguishing fault, (c) Line energizing, (d) Non-fault interruption and (e) Transformer energizing.

**Real life PQ event data**

To test the performance of the proposed classification system, experiments were conducted using PQ events originated from real power networks. Real life PQ event data were obtained from the measurement results of the National Power Quality Monitoring System, which is designed to monitor the PQ of the Turkish Electricity Transmission System [14]. Real life PQ event data consists of three line voltage waveforms obtained from monitoring system of high voltage and medium power networks, with 50 Hz, in different cities in Turkey. There are 538 PQ events and the duration of each event is 150 periods, which including event periods. The sampling rate for the collection of PQ event data is 25.6 kHz. The final classes of PQ events are constructed as Table 1.

Table 1. PQ event types

| Event Type | Class | Size |
|---|---|---|
| Fault events | C1 | 234 |
| Self-extinguishing fault events | C2 | 90 |
| Line energizing events | C3 | 50 |
| Non-fault interruption events | C4 | 44 |
| Transformer energizing events | C5 | 120 |
| **Total** | | **538** |

**Wavelet transform theory**

The WT has been successfully applied to solve many PQ problems [3–9, 15–17]. Since WT can accurately represent any non-stationary waveforms while preserving both time and frequency information, it is used in this paper to extract the distinctive features of PQ events. The WT decomposes a signal into different scales of resolution, providing information about time and frequency domains.

The WT can be continuous or discrete. Discrete WT can be viewed as a subset of Continuous WT. In practical applications, the discrete WT is commonly used. The discrete WT is normally implemented by Mallat's algorithm [18]. Discrete WT uses the low-pass *h(k)* and the high-pass *g(k)* filters to divide the frequency-band of the input signal *f(k)* in respective low- and high-frequency components into octave bands [19]. The low-pass filter *h(k)* is determined from the scaling function. The high-pass filter *g(k)* is determined from both the wavelet and scaling functions. The wavelet and scaling functions are respectively given as,

(1) $$\psi(k) = \sqrt{2}\sum_{n} g(n)\phi(2k-n)$$

(2) $$\phi(k) = \sqrt{2}\sum_{n} h(n)\phi(2k-n)$$

where *n* is integers and represent the number of samples. While the low-pass filtering produces the approximations *Aj*, the high-pass filtering produces the details *Dj* of the decomposition. The relationship of the approximation coefficients and detail coefficients between two adjacent levels are given as,

(3) $$A_{j+1}(k) = \sum_{n} h(n-2k)A_j(n)$$

$$D_{j+1}(k) = \sum_n g(n-2k) A_j(n) \quad (4)$$

where *j* is frequency band level. The WT multiresolution analysis is based on decomposition of the original signal into different signals at various levels of resolution. First, the original signal is passed through the two filters producing the detail *D1* and approximate *A1* coefficients for *j=1*. After down-sampling by a factor of 2, the approximate coefficients *A1* are passed through the same filters again to obtain the coefficients for *j=2*. After another down-sampling, the approximate coefficients *A2* are then filtered again to obtain the next level of coefficients. This filtering operation continues in this way. A given signal *f(k)* is expanded in terms of its orthogonal basis of scaling and wavelet functions. In essence, it is represented by one set of scaling coefficients, and one or several sets of wavelet coefficients,

$$f(k) = \sum_n A_1(n)\phi(k-n) + \sum_n \sum_{j=1} D_j(n) 2^{-j/2} \psi\left(2^j k - n\right) \quad (5)$$

There are many wavelet functions named as mother wavelets. The choice of mother wavelet is important because different types of mother wavelets have different properties. Several popular wavelet functions are Haar, Morlet, Coiflet, Symlet and Daubechies wavelets. In Fig.2, the frequency band levels and frequency divisions of levels are shown for a signal which has 25.6 kHz sampling rate.

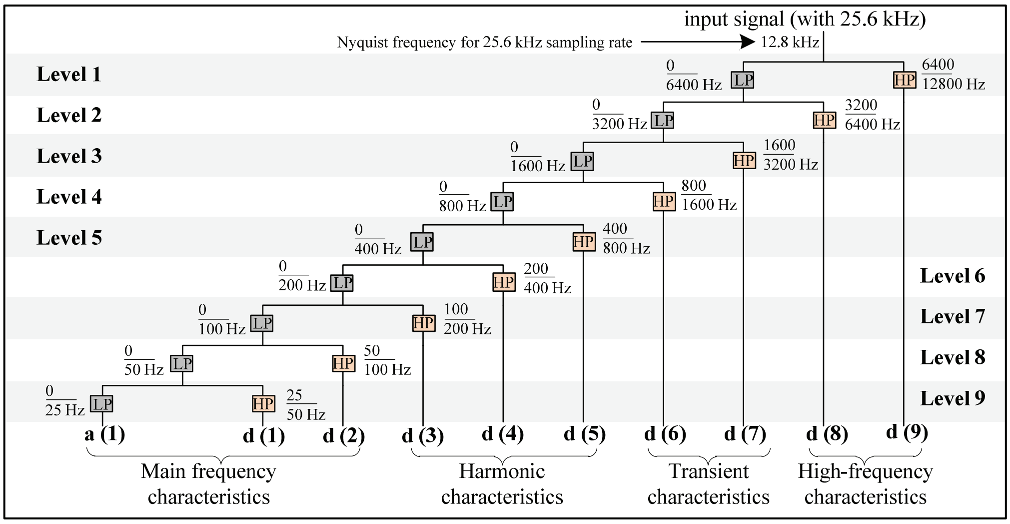


Fig. 2. Frequency division of WT filters for 25.6 kHz sampling rate.

**Feature Extraction and Selection**

In this study, the PQ event signals were analyzed with a 9-level multiresolution decomposition of the WT. The db4 wavelet which is widely used in electromagnetic transient analysis [20] was selected as the wavelet function. The detail coefficients belonging to each levels and 9th final level approximation coefficient level was used to extract the features. For each phase of a PQ event, statistical features like mean, standard deviation, skewness, kurtosis, RMS, energy, Shannon-entropy, log-energy entropy and norm entropy of the detail and approximation coefficients are calculated by using the following equations,

$$\text{Mean } \mu_{ki} = \frac{1}{N}\sum_{j=1}^{N} C_{ij} \quad (6)$$

$$\text{Standart deviation } \sigma_{ki} = \left(\frac{1}{N}\sum_{j=1}^{N}\left(C_{ij}-\mu_i\right)^2\right)^{\frac{1}{2}} \quad (7)$$

$$\text{Skewness } SK_{ki} = \sqrt{\frac{1}{6N}}\sum_{j=1}^{N}\left(\frac{C_{ij}-\mu_i}{\sigma_i}\right)^3 \quad (8)$$

$$\text{Kurtosis } KRT_{ki} = \sqrt{\frac{N}{24}}\left(\frac{1}{N}\sum_{j=1}^{N}\left(\frac{C_{ij}-\mu_i}{\sigma_i}\right)^4 - 3\right) \quad (9)$$

$$\text{RMS } rms_{ki} = \sqrt{\frac{1}{N}\sum_{j=1}^{N} C_{ij}^2} \quad (10)$$

$$\text{Energy } E_{ki} = \sum_{j=1}^{N}\left|C_{ij}\right|^2 \quad (11)$$

$$\text{Shannon entropy } SE_{ki} = -\sum_{j=1}^{N} C_{ij}^2 \log\left(C_{ij}^2\right) \quad (12)$$

$$\text{Log energy entropy } LOE_{ki} = \sum_{j=1}^{N} \log\left(C_{ij}^2\right) \quad (13)$$

$$\text{Norm entropy } NE_{ki} = \sum_{j=1}^{N}\left(C_{ij}\right)^P \quad 1 \le P \quad (14)$$

where *i = 1, 2, ... , l* is decomposition level and *N* is number of the coefficients of detail or approximation at each decomposition level. Here *C* is detail coefficients of each levels and 9th final level approximation coefficient level. *k* represents phase-*1*, *-2* and *-3*. Thus, a ten-dimensional feature vector is constructed for 9-level wavelet decomposition. In the end of this feature extraction process, three feature vectors are obtained for per feature extraction technique. In this study, a single feature vector is created for each PQ event as a result of applying the standard deviation technique to these vectors as follows:

$$SF_i = \left(\frac{1}{N}\sum_{k=1}^{3}\left(feat_{ki} - \frac{1}{N}\sum_{k=1}^{3} feat_{ki}\right)^2\right)^{\frac{1}{2}} \quad (15)$$

where $feat_{ki}$ is statistical feature obtained for $k$ phase and $i$-level. $SF$ represents single feature vector of each feature extraction technique. Thus in the present case the decomposition of the signal up to 9th level yields 10 different nodes corresponding to different frequency sub-bands. Selecting the above mentioned features for each level we will obtain the feature vector of length 90. The feature vector is denoted by,

$$\begin{aligned} \text{Features} = [&\mu_{D_1}\ \mu_{D_2} \ldots \mu_{D_9}\ \mu_{A_9} \ldots \\ &\sigma_{D_1}\ \sigma_{D_2} \ldots \sigma_{D_9}\ \sigma_{A_9} \ldots \\ &SK_{D_1}\ SK_{D_2} \ldots SK_{D_9}\ SK_{A_9} \ldots \\ &KRT_{D_1}\ KRT_{D_2} \ldots KRT_{D_9}\ KRT_{A_9} \ldots \\ &rms_{D_1}\ rms_{D_2} \ldots rms_{D_9}\ rms_{A_9} \ldots \\ &E_{D_1}\ E_{D_2} \ldots E_{D_9}\ E_{A_9} \ldots \\ &SE_{D_1}\ SE_{D_2} \ldots SE_{D_9}\ SE_{A_9} \ldots \\ &LOE_{D_1}\ LOE_{D_2} \ldots LOE_{D_9}\ LOE_{A_9} \ldots \\ &NE_{D_1}\ NE_{D_2} \ldots NE_{D_9}\ NE_{A_9}\ ] \end{aligned} \tag{16}$$

where $D$ is feature of detail coefficients and $A$ is feature of approximation coefficients.

The feature selection methods realize the selection of the best subset of the input feature set. In this study, in order to less the large number of PQ event features, the sequential forward selection (SFS) method are tried for the selection of useful and elimination of useless feature identifiers among the described 20 dimensional feature vector. SFS method, proposed by Maril and Green [21], performs a heuristic-guided Depth-First search on feature space, starting with the empty set. At each step, all features not yet included in subset of selected features are individually incorporated in the subset and a criterion value computed. The feature that yields the best value is then included in the new subset.

**LMT Classifier**

The LMT has the advantage of having a fast learning process. Another important advantage of LMT is that it does not need adjustment of different sensitive parameters. As the structure of LMT is simple and learning efficiency is very high, it is suitable for signal classification problems. Moreover, the LMT classifier can be effectively used in problems with a high dimensional input vector, and it can be built with high level of accuracy using little data preparation [2].

LMT algorithm is a combination of a tree structure and logistic regression functions to produce a single decision tree. LMT employs the LogitBoost algorithm for building the logistic regression functions at the nodes of a tree and uses the well-known CART algorithm for pruning. LogitBoost is used to select the most relevant attributes in the data when performing logistic regression by performing a simple regression in each iteration and stopping before convergence to the maximum likelihood solution. The benefit obtained from using LogitBoost is that a separate smoothing process is not required [22-24].

A LMT consists of a tree structure that is made up of a set of inner or non-terminal nodes $N$ and a set of leaves or terminal nodes $T$. Let $S$ denote the whole instance space, spanned by all attributes that are present in the data. Then the tree structure gives a disjoint subdivision of S into regions St, and every region is represented by a leaf in the tree:

$$S = \bigcup_{t \in T} S_t,\ \ S_t \cap S_{t'} = \varnothing \ \ for \ \ t \neq t' \tag{17}$$

Unlike ordinary decision trees, the leaves $t \in T$ have an associated logistic regression function *ft* instead of just a class label. The regression function *ft* takes into account a subset $Vt \subseteq V$ of all attributes present in the data (where we assume that nominal attributes have been binarized for the purpose of regression), and models the class membership probabilities as,

$$\Pr(G = j \mid X = x) = \frac{e^{F_j(x)}}{\sum_{k=1}^{J} e^{F_k(x)}} \tag{18}$$

where,

$$F_j(x) = \alpha_0^j + \sum_{v \in V_t} \alpha_v^j \cdot v, \tag{19}$$

or equivalently,

$$F_j(x) = \alpha_0^j + \sum_{k=1}^{m} \alpha_{v_k}^j \cdot v_k \quad if\ \alpha_{v_k}^j = 0 \ \ for \ \ v_k \notin V_t. \tag{20}$$

The model represented by the whole LMT is then given by,

$$f(x) = \sum_{t \in T} f_t(x) \cdot I(x \in S_t) \tag{21}$$

where $I(x \in S_t)$ *is* 1 *if* $x \in S_t$ and 0 otherwise.

***Proposed Classification Algorithm***

In this work, the proposed algorithm based on the feature selection process is designed to identify five types of PQ events including fault, self-extinguishing fault, line energizing, non-fault interruption and transformer energizing events. The proposed algorithm for PQ event classification is divided into the four stages: pre-processing, feature extraction, feature selection and classification shown as Fig. 3.

*Pre-processing stage*

In the pre-processing stage of the proposed algorithm, the PQ event signals are firstly segmented as three cycles in the form of one cycle of the pre-event segment and two cycles of the event segment. Then, the normalization step is fulfilled for per PQ event. In the normalization step, the event voltage waveform is converted to a relative scale, per unit (pu), by dividing the input signal by the nominal RMS voltage.

*Feature extraction stage*

In this stage, the distinctive features for per event are extracted by using WT and feature extraction methods. The wavelet coefficients, detail and approximation, for each phase of event are obtained by applying a 9-level multiresolution analysis. These coefficients include effective feature information for event type. Nine feature extraction methods mentioned above are applied to the detail coefficients belonging to each levels and 9th final level approximation coefficient level. Thus, the three feature vectors are obtained for per event. Finally, a single feature vector is extracted by applying the standard deviation technique to three feature vectors. By means of feature extraction stage, both distinctive features are obtained and the dimensionality of the feature space is lessened.

*Feature selection stage*

Feature selection is an important step in many pattern classification problems. It is applied to select a subset of PQ event features obtained from feature extraction stage. For this purpose, the best subset of PQ event features is firstly searched using SFS method. Then, selected better 20 features from all 90 features are applied as input to LMT classifier.

*Classification stage*

In the classification stage, the PQ event is classified using the LMT technique. The features selected from SFS method are given to the LMT for training and are subsequently tested for an effective classification. Besides, in order to evaluate the performance of the LMT, back-propagation neural network (BPNN) which is a well-known classifier method is used.

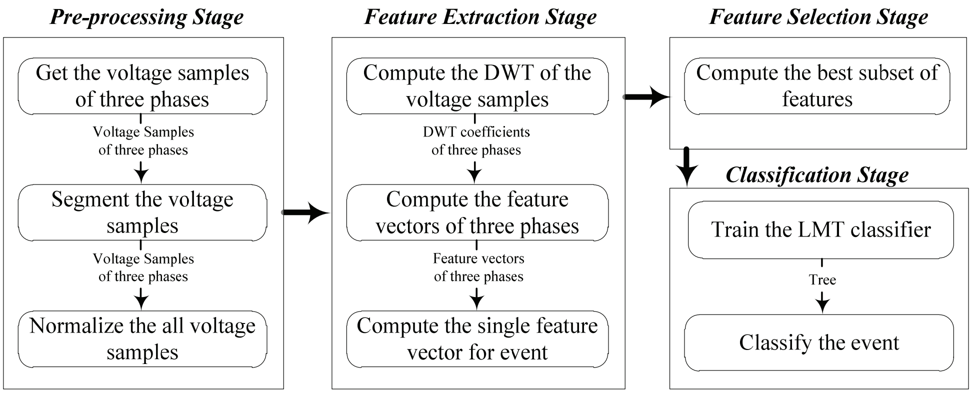


Fig.3. Flowchart of the proposed automatic event classification algorithm.

*Results of the Feature Selection Stage*

Due to obtain the optimal feature set, SFS method is applied to the 90 features obtained from the feature extraction stage. The feature selection process is fulfilled in the training stage of the algorithm and applied to the training event data. In table 2, optimally selected 20 features are tabulated.

Table 2. The first best 20 ones of the selected features.

| Feature names | | | |
|---|---|---|---|
| 1 | 7th level of Mean | 11 | 1st level of energy |
| 2 | 8th level of Mean | 12 | 2nd level of energy |
| 3 | 4th level of Std.Deviation | 13 | 5th level of energy |
| 4 | 8th level of Std.Deviation | 14 | 9th level of energy |
| 5 | 3rd level of Skewness | 15 | App. level of energy |
| 6 | 4th level of Skewness | 16 | App. level of Shan. entropy |
| 7 | 7th level of Skewness | 17 | 7th level of Log en.entropy |
| 8 | 9th level of Skewness | 18 | 8th level of Log en. entropy |
| 9 | 4th level of Kurtosis | 19 | 9th level of Log en. entropy |
| 10 | 5th level of Kurtosis | 20 | 7th level of Norm entropy |

*Results of the Classification Stage*

In the proposed algorithm, after realizing the feature extraction step using the WT, the feature selection step is implemented by using the SFS technique. Therefore, distinctive features of the PQ event signals are applied to the input of the LMT classifier. 538 PQ events with 20-dimensional optimal feature sets are used for evaluating of the proposed automatic recognition and classification algorithm. Fifty percent of these events were used for training the algorithm. The remaining PQ event signals were used for testing the algorithm.

The classification results of the generated proposed method are illustrated in Table 3. In this table, the correct classification results are tabulated at diagonal elements. The misclassification results are tabulated at non-diagonal elements.

Table 3. Classification results of the proposed algorithm.

| **True Class** | **$C_1$** | **$C_2$** | **$C_3$** | **$C_4$** | **$C_5$** | **Accuracy (%)** |
|---|---|---|---|---|---|---|
| **$C_1$** | 111 | 1 | 0 | 2 | 2 | 95.7 |
| **$C_2$** | 2 | 50 | 0 | 0 | 0 | 96.1 |
| **$C_3$** | 1 | 0 | 18 | 0 | 0 | 94.7 |
| **$C_4$** | 0 | 0 | 0 | 20 | 1 | 95.2 |
| **$C_5$** | 2 | 0 | 0 | 0 | 59 | 96.7 |
| Overall success rate (%) : 95.9 | | | | | | |

Table 4 gives the classification results obtaining from feature vector of the each feature extraction method. Classification results show that norm entropy technique has higher classification accuracy with regard to the other techniques.

Table 4. Classification results obtaining from each feature extraction method.

| Feature extraction methods | Classification accuracy (%) | | | | | Overall success rate (%) |
|---|---|---|---|---|---|---|
| | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | |
| Mean | 87.1 | 73.1 | 94.7 | 85.7 | 73.8 | 81.8 |
| Std. deviation | 87.9 | 75.0 | 78.9 | 66.7 | 80.3 | 81.4 |
| Skewness | 72.4 | 23.1 | 26.3 | 71.4 | 85.2 | 62.5 |
| Kurtosis | 86.2 | 1.9 | 0 | 14.3 | 86.9 | 58.4 |
| RMS | 78.4 | 71.2 | 84.2 | 85.7 | 90.2 | 80.7 |
| Energy | 87.1 | 78.8 | 84.2 | 100 | 70.5 | 82.5 |
| Shannon-entropy | 77.6 | 51.9 | 89.5 | 52.4 | 75.4 | 71.0 |
| Log-en. entropy | 91.4 | 3.8 | 47.4 | 0 | 65.6 | 58.4 |
| Norm entropy | 87.1 | 88.5 | 94.7 | 61.9 | 88.5 | 86.2 |

Table 5. Classification results obtaining from using the BPNN classifier.

| True Class | C1 | C2 | C3 | C4 | C5 | Accuracy (%) |
|---|---|---|---|---|---|---|
| C1 | 109 | 1 | 0 | 0 | 6 | 94.0 |
| C2 | 8 | 43 | 0 | 0 | 1 | 82.7 |
| C3 | 0 | 0 | 19 | 0 | 0 | 100 |
| C4 | 0 | 0 | 0 | 20 | 1 | 95.2 |
| C5 | 3 | 0 | 0 | 0 | 58 | 95.1 |
| Overall success rate (%) : 92.6 | | | | | | |

A comparison with BPNN was made in order to evaluate the proposed algorithm well. The network was trained with same training samples. The classification results with BPNN were shown as Table 5. The results in Table 5 show that, compared with BPNN, identification of PQ events with LMT classifier has a much higher correct ratio.

**Comments to the Proposed Algorithm and the Results**

The most important advantage of the proposed algorithm based WT and LMT is the reduction of data size as well indicating and recognizing the main characteristics of signal without losing its distinguishing characteristics. It can reduce memory space, shorten preprocessing needs, the network size and increase computation speed for the classification of PQ events. Furthermore, another advantage of the proposed system is to have a single feature vector for a three-phase event as different from the similar studies in the literature.

In this study, proposed algorithm is evaluated using the PQ events originated from the real power network. The events include both noise with different level and different characteristics such as incidence angle and amplitude. The overall classification accuracy for proposed algorithm are 95.9% and the classification results for feature extraction techniques are 81.8%, 81.4%, 62.5%, 58.4%, 80.7%, 82.5%, 71.0%, 58.4, respectively. The classification result with BPNN is 92.6% for selected features. Obtained results indicate that the proposed automatic event classification algorithm is robust and has ability to distinguish different power quality event classes easily.

**Conclusions**

Recently, a significant amount of work in literature has been done towards the development of methods for automatic classification of power quality disturbances. These works define classes based on disturbance types (e.g., dip, interruption, and transient), rather than classes based on their underlying causes. Such works are important for the development of methods but has, as yet, limited practical value. Of more practical needs are tools for classification based on underlying causes (e.g., voltage sag due to fault or transformer energizing, voltage swell due to fault).In this paper, a feature selection and LMT based approach for classifying of underlying causes of PQ disturbances has been presented. It is based on a reduced and simple set of features extracted from the WT in order to advance accuracy, simplicity and reliability. In proposed algorithm, feature selection algorithm based on SFS starts with a relatively large feature vector and selects the more useful feature subsets using several feature selection methods. The experiments are repeated for LMT classifiers. The results indicate that proposed algorithm can detect and classify different PQ events correctly.

REFERENCES


[1] Liao, Y. and Lee, J. B., A fuzzy-expert system for classifying power quality disturbances, *Electrical Power and Energy Systems,* (2004), 26(3), 199–205.

[2] Moravej, Z., Abdoos, A. A. and Pazoki, M., New combined s-transform and logistic model tree technique for recognition and classification of power quality disturbances, *Electric Power Components and Systems*, (2011), 39(1), 80–98.

[3] Erişti, H., Uçar, A. and Demir, Y., Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines, *Electric Power Systems Research,* (2010), 80(7), 743–752.

[4] Gaouda, A. M., Salama, M. M. A., Sultan, M. R. and Chikhani, A. Y., Power quality detection and classification using wavelet multiresolution signal decomposition, *IEEE Transactions on Power Delivery,* (1999), 14(4), 1469–1476.

[5] Pires, V. F., Amaral, T. G. and Martins, J. F., Power quality disturbances classification using the 3-D space representation and PCA based neuro-fuzzy approach, *Expert Systems with Applications,* (2011), 38(9), 11911–11917.

[6] Gaing, Z. L., Wavelet-based neural network for power disturbance recognition and classification, *IEEE Transactions on Power Delivery,* (2004), 19(4), 1560–1568.

[7] Ekici, S., Classification of power system disturbances using support vector machines, *Expert Systems with Applications,* (2009), 36(6), 9859–9868.

[8] Hong, Y. Y. and Wang, C. W., Switching detection/classification using discrete wavelet transform and self-organizing mapping network, *IEEE Transactions on Power Delivery*, (2005), 20(2), 1662–1668.

[9] Erişti, H. and Demir, Y., A new algorithm for automatic classification of power quality events based on wavelet transform and SVM, *Expert Systems with Applications,* (2010), 37(6), 4094-4102.

[10] Axelberg, P. G. V., Irene, Y. H. G. and Bollen, M. H. J. Support vector machine for classification of voltage disturbances, *IEEE Transactions on Power Delivery,* (2007), 22(3), 1297–1303.

[11] Styvaktakis, E., Bollen, M. H. J. and Gu, I. Y. H. Expert system for classification and analysis of power system events, *IEEE Transactions on Power Delivery,* (2002), 17(2), 423–428.

[12] Bollen, H. J., Gu, I. Y. H., Axelberg, P. G. V. and Styvaktakis, E. Classification of underlying causes of power quality disturbances: deterministic versus statistical methods, *EURASIP Journal on Advances in Signal Processing*, (2007), article id: 79747, 17 pp.

[13] Santoso, S., Lamoree, J., Grady, W. M., Powers, E. J. and Bhatt, S. C., A scalable PQ event identification system, *IEEE Transactions on Power Delivery*, (2000), 15(2), 738–743.

[14] Demirci, T., Kalaycıoğlu, A., Küçük, D., Salor, Ö., Güder, M., Pakhuylu, S., Atalık, T., İnan, T., Çadırcı, I., Akkaya, Y., Bilgen, S. and Ermiş, M. , Nationwide real-time monitoring system for electrical quantities and power quality of the electricity transmission system, *IET Generation, Transmission & Distribution*, 2011, 5(5), 540-550.

[15] Morsi, W. G. and El-Hawary, M. E., Power quality evaluation in smart grids considering modern distortion in electric power systems. *Power Systems Research,* (2011), 81(5), 1117–1123.

[16] Wang, M. H. and Tseng, Y. F., A novel analytic method of power quality using extension genetic algorithm and wavelet transform, *Expert Systems with Applications,* (2011) 38(10), 12491–12496.

[17] Morsi, W. G. and El-Hawary, M. E., A new reactive, distortion and non-active power measurement method for nonstationary waveforms using wavelet packet transform, *Electric Power Systems Research,* (2009), 79(10), 1408–1415.

[18]Mallat, S. G., A theory for multiresolution signal decomposition: The wavelet representation *IEEE Transaction on Pattern Analysis and Machine Intelligence*, (1989), 11(7), 674–693.

[19] Daubechies, I. (1992). *Ten lectures on wavelets*. Philadelphia, USA: CBMSNSF Regional Conference Series, SIAM.

[20]Chen, S. and Zhu, H. Y., Wavelet transform for processing power quality disturbances, *EURASIP Journal on Advances in Signal Processing*, (2007), article ID: 47695, 20 pp.

[21] Marill T. and Green D. M., On the effectiveness of receptors in recognition systems, *IEEE Transactions on Information Theory,* (1963), 9(1), 11–17.

[22] Landwehr, N., Hall, M. and Frank, E., Logistic model trees, *Machine Learning,* (2005), *59*(1/2), 161–205.

[23] Hosmer D. W. and Lemeshow, S., Applied logistic regression, (2000), 2nd Edition, Wiley-Interscience, New York.

[24] Friedman, J., Hastie, T. and Tibshirani, R., Additive logistic regression: A statistical view of boosting, *Annals of Statistics*, (2000), 32(2), 337–374.



***Authors**: Assist. Prof. Dr. Huseyin Eristi, Tunceli University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Tunceli/Turkey, E-mail: huseyineristi@gmail.com, heristi@tunceli.edu.tr; Prof. Dr. Yakup Demir, Firat University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Elazig/Turkey, E-mail: ydemir@firat.edu.tr*

*The correspondence address is:*
*e-mail: huseyineristi@gmail.com*