



Flexible Entropy Based Feature Selection and Multi Class SVM for Detection and Classification of Power Quality Disturbances

Aslam Shaik ^{1*} Ankireddy Srinivasulu Reddy ²

¹*Department of Electrical and Electronics Engineering,
Jawaharlal Nehru Technological University, Hyderabad, Telangana, India*

*Department of Electrical and Electronics Engineering,
CMR Engineering College, Hyderabad, Telangana, India*

* Corresponding author's Email: aslam.careers@gmail.com

Abstract: Power quality disturbances (PQD) are becoming more and more complicated due to the vast growth of power quality oriented applications particularly the power electronic equipment's. Since the power with less power quality consequences to a huge loss in the electronic devices, there is a need to detect the power quality disturbances more effectively. This paper proposes a new method for PQD detection and classification based on the entropy characteristics of PQD signals and a novel support vector machine algorithm (SVM). The proposed approach develops a flexible entropy based feature selection (FEFS) mechanism to extract the unique characteristics of all types of PQDs such that the detection system can detect more effectively with in less time. Further to increase the accuracy of individual PQDs, this paper proposed Multi Class SVM (MC-SVM). Various experiments are conducted over the proposed work and at every experiment; the performance is measured through Classification accuracy, False Alarm Rate and computation time. Further a comparative analysis reveals the outstanding performance of proposed approach when compared with conventional approaches.

Keywords: Power quality disturbances, Entropy, Mutual information, Support vector machine, Sag, Swell, Accuracy.

1. Introduction

In recent years, due to the rapid growth in the power electronic equipment components, solid state switching devices used in the public sectors and industrial sectors, an increased quality of power is demanded which makes the sensitive equipment more secure from typical power accidents. The quality of electrical supplies has gradually become an important issue for electric utilities and their customers. Since the occurrence of power quality disturbance results a great impact over the equipment, there is a need to design effective power disturbance detection mechanisms by which the safe and economical operations can be preserved more efficiently in the electric systems [1]. Hence, the study and analysis of issues in the power systems is significantly more important for the improvement of

power quality [2]. Particularly the exact detection and classification of power quality disturbances is a top priority in this direction and can also support power quality evaluation.

Generally the problems associated with power quality are voltage sag/swell with and without harmonics, interruption, oscillatory transient, pure harmonics, and flicker etc. Along with these problems, there exist some more problems, derived based on the deviations in the time, magnitude and frequency characteristics of electrical signals. To detect and classify the power quality disturbances, a complete knowledge about the characteristics of electrical signals is required [3]. So many approaches are developed in earlier to analyze the in depth nature of different electrical signals such that the perfect discrimination between the power quality problems can be obtained. Further the detection and

classification of power quality problems is carried out in three phases, (i) pre-processing, (ii) feature extraction and (iii) classification. Feature selection is always the key element among these processes as some essential features may be overlooked and some non-essential features may be inappropriately regarded. Any resulting combination of inappropriate attributes would add to the difficulty of classification when disturbance and noise exist simultaneously. In the feature extraction phase, extracting the optimal feature set is more important by which the classification accuracy increases significantly. But this feature extraction could not result in extra burden over the system. For this purpose, the obtained feature set needs to be optimal and also more informative. Further selection of an appropriate classifier also constitutes a major concern in this PQD detection regard.

This paper proposes a new Power Quality Disturbances detection and classification mechanism based on the Flexible Entropy oriented feature selection (FEFS) and Multi class Support Vector Machine (MC-SVM). The proposed FEFS extracts an optimal feature set by which the power quality problems can be detected more accurately. The proposed FEFS derives the mutual relationships between all the signals and accomplishes those relationships in the feature selection. This gives an optimal feature set with reduced dimensionality. Further the proposed MC-SVM helps in the accurate detection of every class of disturbances and reduces the false alarm rate. Various simulation experiments are accomplished over the proposed approach to prove the efficiency and the obtained results reveals an outstanding performance.

Rest of the paper is organized as follows: Section 2 describes the literature survey details. The details of power quality disturbances and their control parameters are illustrated in section 3. The details or proposed detection mechanism is illustrated in section 4. Experimental evaluation is described in section 5 and finally the conclusions and future scope are described in section 6.

2. Literature survey

There are several sources of disturbances in any real-time system and it is unquestionably essential to detect and classify the disturbances automatically to improve the quality and reliability of power. Different approaches are proposed in earlier to achieve maximum detection accuracy in the power quality disturbances detection. Based on the objective aimed to achieve, the earlier developed

approaches are categorized as feature extraction approaches and classification approaches.

2.1 Feature extraction approaches

In this category, the techniques extracts features from signals to understand the characteristics of different power quality problems. Different techniques such as wavelet transform (WT), short-time Fourier transform (STFT), Gabor-Winger transform, S-transform (ST) have been applied for detection and classification of PQ disturbances over the past years. STFT [4] provides the information about time-frequency characteristics of a disturbance signals, however it has fixed window size which results in the not-efficient representation about the behavior of transient signals. Further the multi-resolution analysis (MRA) [5] and its derivatives are accomplished over the electrical signals to analyze the characteristics of power quality disturbances in the resolution level but it is observed that the performance is quite less in the case of noise affected signals. To overcome this drawbacks, some more techniques like WT [6], S-Transform [7, 8], Kalman filter [9] Gabor-Winger Transform [10], Hilbert Haung Transform (HHT) [11], Parallel Computing [12], and some hybrid techniques [13 - 15] have been proposed for the detection of power quality disturbances in the noisy environments. Wavelet Transform decomposes a the PQD signal into low frequency and high frequency bands through which the more detailed information about the signal disturbance can be revealed. But when the signals are buried under harmonics or noise, Wavelet based methods becomes unreliable. Further the main drawback with WT and its subsequent is the spectral leakage by which the performance of PQD detection system becomes degraded [20].

Hilbert Haung Transform based on empirical mode decomposition is proposed in [16, 17], to analyze the power quality disturbances of the electrical signals by decomposing them into intrinsic mode functions (IMFs). Here the signal is decomposed into IMFs initially and then they are processed for analysis through Hilbert transform. However the HHT cannot reveal the frequency characteristics of the signal, resulting in an increased burden. S-Transform is derivative of WT and inherits the STFT and WT since it is considered as the WT with phase correction or STFT with variable sized window. Since the ST [18, 19], has an ability to analyze the signal even under noisy environments, there has been a wide usage is observed in the detection of Power quality disturbances. However,

the high computational demand of ST constraints its application to real time applications. Though all these transform techniques can alleviate the characteristics of power quality problems effectively, considering all the feature results in an unnecessary computational complexity, which was not focused in earlier approaches.

2.2 Classification approaches

In this category, the main objective is to achieve maximum detection accuracy. Different techniques such as Artificial Neural Networks (ANN) [21, 23], Support Vector Machine (SVM) [22, 24], K-nearest Neighbor (K-NN) [25], Fuzzy Logic (FL) [26], Adaboost Classifier [27], Decision Tree (DT) [28], etc. have been applied for detection and classification of PQ disturbances over the past years. Kanirajan [21] proposed power quality disturbance detection technique by combining wavelet transform with radial basis function neural network (RBFNN). The obtained results are compared with generalized regressive neural network, feed forward neural network, learning vector quantization and probabilistic neural network techniques and it is shown that the accuracy is improved. Adaptive feature extraction technique that combines the EMD and Hilbert Spectral Analysis is combined with probabilistic neural network in [23] to detect the multiple power quality problems.

Further the Moravej [24] proposed a PQ detection technique based on the support vector machine algorithm based on the inherent characteristics of signals. In [25], Dasgupta developed a new method based on the K-nearest neighbor classifier by measuring the correlation to detect and classify the transmission line faults. The noise immune S-transform is combined with fuzzy expert system in [26] for assigning a certainty factor for every classification rule thereby to improve the robustness of the PQ detection system in the presence of noise. Another approach for power quality detection is proposed by considering the rule-based S-Transform as a feature extraction and an Adaptive Boost (AdaBoost) as a classifier in [27]. By considering the advantages of ANN and decision tree, a hybrid power quality detection framework is proposed in [28]. This approach is applied on multiple PQ disturbances such as harmonics with swell, sag, interruption and flicker, observed according to the recommended practice in IEEE [30]. Further a Decision making strategy combined with S-Transform is developed in [31] to identify the power quality disturbances. Due to the unstructured decision making strategy, this method comprised

with more computational complexity. Combining the Wavelet Transform with Probabilistic Neural Network (PNN) and an adaptive iterative approach is proposed in [32] to achieve increased detection accuracy in the classification of power quality disturbances. Combining the Most efficient SVM with Empirical Mode decomposition, a novel fault classification technique is proposed in [33]. EMD is applied to extract the features of power disturbance signals and SVM is accomplished to classify the disturbances based on the frequency characteristics of Intrinsic Mode Functions. A multiple SVM model is introduced for classifying the fault condition among ten power system faults. Algorithm is validated using MATLAB/ SIMULINK environment. However all these approaches have their own advantages and disadvantages like some methods are robust for only some power quality problems. For example, the decision tree algorithm achieves better classification accuracy in the case of normal power quality problems but won't work on the harmonics based power quality problems. Similarly the convergence time of neural network related approaches is observed to be high. To overcome the problems with conventional approaches, this paper proposed a new method combining the Multi-Class SVM with entropy based feature selection. Since the Entropy based feature not only reveals the linear dependency between the signals, but also provides the non-linear dependencies between the power disturbances signals by which the classification accuracy will be increased more.

3. Power quality disturbance signals

A pure sine wave with frequency 50 Hz and magnitude at 1.0 p.u. as well as eight other PQ disturbance signals such as sag, swell, harmonics, interruption, sag with harmonics, swell with harmonics, flicker, and oscillatory transients are generated through Matlab simulation software. The respective formulae of the power quality disturbances and their control parameters are represented in Table 1. The total recording time of the signal is 0.4 s. The reference frequency of pure sine wave is 50 Hz. The PQ disturbance signal generation models and their control parameters are shown in Table 1. Using the PQ disturbance equations mentioned in Table 1, the power quality disturbances are generated through MATLAB software and represented in Fig. 1.

Table 1. Power quality disturbance signals and its controlling parameters

PQD	Class	Formula	Parameter Variations
Normal	C1	$s(t) = \sin(\omega_n t)$	$\omega_n = 2\pi \times 50 \text{ rad/sec}$
Swell	C2	$s(t) = [1 + \alpha(u(t - t_1) - u(t - t_2))] \sin(\omega_n t)$	$0.1 \leq \alpha \leq 0.8, T \leq t_2 - t_1 \leq 9T$
Sag	C3	$s(t) = [1 - \alpha(u(t - t_1) - u(t - t_2))] \sin(\omega_n t)$	$0.1 \leq \alpha \leq 0.9, T \leq t_2 - t_1 \leq 9T$
Flicker	C4	$s(t) = [1 + \alpha \sin(2\pi\beta t)] \sin(\omega_n t)$	$0.1 \leq \alpha \leq 0.2, 5\text{Hz} \leq \beta \leq 20\text{Hz}$
Interruption	C5	$s(t) = [1 - \alpha(u(t - t_1) - u(t - t_2))] \sin(\omega_n t)$	$0.9 \leq \alpha \leq 1, T \leq t_2 - t_1 \leq 9T$
Oscillatory transient	C6	$s(t) = \sin(\omega_n t) + \alpha \exp(-(t - t_1)/\tau)(u(t - t_2)) \sin(2\pi f_n t)$	$0.1 \leq \alpha \leq 0.8, 0.5T \leq t_2 - t_1 \leq 3T, 300 \text{ Hz} \leq f_n \leq 900\text{Hz}, 8\text{ms} \leq \tau \leq 4\text{m}$
Harmonics	C7	$s(t) = \alpha_1 \sin(\omega_n t) + \alpha_3 \sin(3\omega_n t) + \alpha_5 \sin(5\omega_n t) + \alpha_7 \sin(7\omega_n t)$	$0.05 \leq \alpha_3, \alpha_5, \alpha_7 \leq 0.15, \sum \alpha_i^2 = 1$
Swell with harmonics	C8	$s(t) = [1 + \alpha(u(t - t_1) - u(t - t_2))] \times [\alpha_1 \sin(\omega_n t) + \alpha_3 \sin(3\omega_n t) + \alpha_5 \sin(5\omega_n t) + \alpha_7 \sin(7\omega_n t)]$	$0.1 \leq \alpha \leq 0.8, T \leq t_2 - t_1 \leq 9T, 0.05 \leq \alpha_3, \alpha_5, \alpha_7 \leq 0.15, \sum \alpha_i^2 = 1$
Sag with harmonics	C9	$s(t) = [1 - \alpha(u(t - t_1) - u(t - t_2))] \times [\alpha_1 \sin(\omega_n t) + \alpha_3 \sin(3\omega_n t) + \alpha_5 \sin(5\omega_n t) + \alpha_7 \sin(7\omega_n t)]$	$0.1 \leq \alpha \leq 0.9, T \leq t_2 - t_1 \leq 9T, 0.05 \leq \alpha_3, \alpha_5, \alpha_7 \leq 0.15, \sum \alpha_i^2 = 1$

4. Proposed approach

This work proposes a new PQD detection and classification framework based on the entropy features of PQD signals and multi-class Support vector machine. Since the SVM is a binary classifier, this work applied multi-class SVM to support multiple power quality problems detection. AT every phase, the SVM classifier classifies the total input data into two classes, for instance the first mode SVM classifies the signal into normal and disturbance signals. The complete details are illustrated in the following subsections.

4.1 Entropy based feature selection

Feature selection plays a more vital role in the Power quality disturbances detection and classification system. The main objective of this feature selection is to get an optimal feature set from every PQD signal such that they can boost up the detection performance followed by a reduced computational burden over the system. To be able to detect and classify the various types of power quality disturbances in the power systems, it is necessary to perform further pre-processing over the original signals. The feature extraction phase accomplishes to extract the feature of original signals by applying some pre-processing techniques. To do this, there is a need to study the unique characteristics of every PQD signal, due to various types of disturbances in the modern AC power systems. Further simply the feature extraction can be defined as a process that transforms the original signal into a new form so that the required information can be extracted.

In this paper, a new feature selection approach is proposed based on the entropy characteristics of PQD signals. In this section, the feature selection is carried out based on the mutual information. There are so many techniques like Euclidean distance, Distribution law distance and correlative measures etc. to discover the spatial relationships between the features of different classes. Among those, the Mutual Information (MI) [29] is one promising measure in the realm of variable dependence estimation. Fortunately, the MI measure not only derives linearly dependent variables but also non-linearly dependent ones. Hence this metric is chosen as a base of the proposed feature selection mechanism.

MI is a feature selection metric that derives the relationship between two features. The highest value of MI indicates that the two features are dependent and the lowest value indicates that the two features are statistically independent. Let's consider two PQD signals X and Y as $X = \{x_1, x_2, x_3, \dots, x_n\}$ and $Y = \{y_1, y_2, y_3, \dots, y_n\}$, where n is number of sample in X and Y. The MI between X and Y can be evaluated as,

$$MI(X; Y) = H(X) + H(Y) - H(X, Y) \tag{1}$$

Where $H(X)$ and $H(Y)$ are the entropies of X and Y. Mathematically the entropies of X and Y are formulated as

$$H(X) = \sum_{i=1}^n p(x_i) \log(p(x_i)) \tag{2}$$

and

$$H(Y) = \sum_{i=1}^n p(y_i) \log(p(y_i)) \tag{3}$$

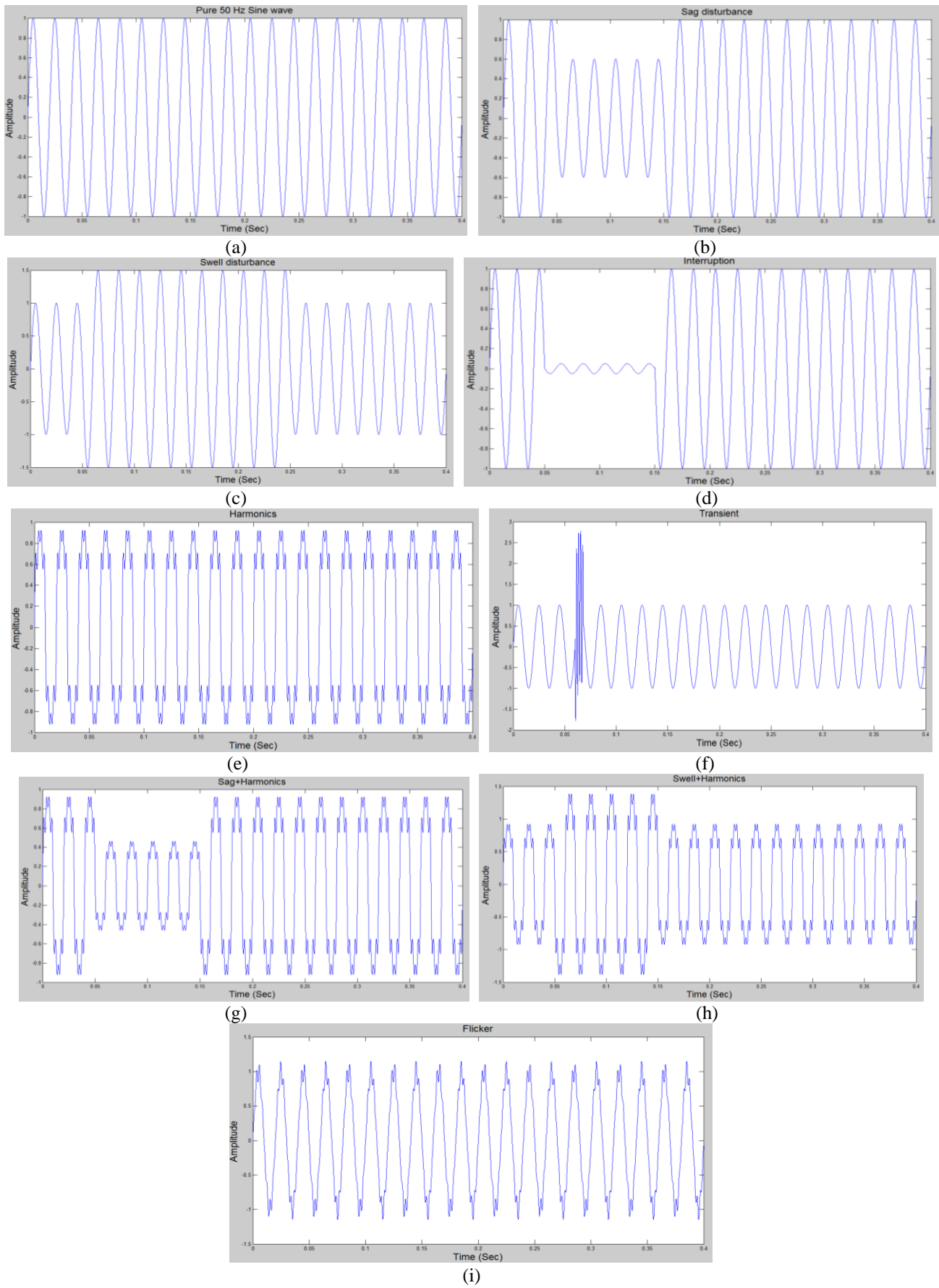


Figure.1 Power quality disturbances: (a) Normal, (b) Sag, (c) Swell, (d) Interruption, (e) Harmonics, (f) Oscillatory Transient, (g) Sag with harmonics, (h) Swell with harmonics, and (i) Flicker

Algorithm 1: Entropy based Feature Selection

Step 1. Initialization: Set $F \leftarrow$ "Initial Set of all features", Set $S \leftarrow$ "Empty Set".

Step 2. Compute MI with the output class C : for every feature $f \in F$ Compute $MI(C; f)$.

Step 3. Find the first feature f that maximizes $MI(C; f)$: set $F \leftarrow F \setminus \{f\}$; Set $S \leftarrow \{f\}$.

Step 4. Find the further features f as one that maximizes $MI(C; f) - \beta \sum_{s \in S} MI(f; s)$;

Set $F \leftarrow F \setminus \{f\}$; Set $S \leftarrow S \cup \{f\}$.

Step 5. Output the set S containing the selected features.

Further the term $H(X, Y)$ denotes the joint entropy and formulated as

$$H(X, Y) = \sum_{i=1}^n \sum_{j=1}^n p(x_i, y_j) \log(p(x_i, y_j)) \quad (4)$$

Simply to derive the dependency of a variable X on Y (or vice versa) known as Mutual Information and formulated as

$$MI(X; Y) = \sum_{i=1}^n \sum_{j=1}^n p(x_i, y_j) \log\left(\frac{p(x_i, y_j)}{p(x_i)p(y_j)}\right) \quad (5)$$

Where $p(x_i, y_j)$ is the joint probability density function and $p(x_i)$ and $p(y_j)$ are the marginal density functions. In the feature selection process, the feature is considered as relevant to a particular class, if it contains important information about class C ; otherwise it is considered as irrelevant or redundant feature. Hence the relevance or irrelevance between the feature and a class C is derived through the MI, as $MI(C; f_i)$, where f_i is the feature that is relevant or irrelevant to class C . In this context, the feature with high predictive power are the feature that have larger mutual information $MI(C; f_i)$. On the contrary, in the case of $MI(C; f_i)$ is equal to zero, the feature f_i and class C are proven to be independent to each other. This means the feature f_i contributes the redundancy to the classification. The Mutual Information based feature selection algorithm is formulated as Algorithm 1.

Here the parameter β regulates the relative importance of the MI between the candidate feature and the already selected features with respect to the MI in the output class. Here the importance of β is formulated in three ways. If $\beta = 0$, only the mutual information with output class is considered for each feature selection. If β increases, this measure subtracts a certain quantity proportional to the mutual Information with respect to the already selected features. However the selection of appropriate values for the parameter β is a complex issue in the PQD detection.

To remove the burden of setting an appropriate value for β , a new version of entropy based feature

selection process is outlined in this paper. This new feature selection approach is an enhancement to the feature selection involved in the step 4 of algorithm 1. The process of feature selection carried out through step 4 is replaced with the equation shown represented in Eq. (6). Eq. (6) shows a new formulation of the feature selection process involved, which is intended to select a feature from an initial input feature set that maximizes $MI(C; f)$ and minimizes the average of minimum redundancy (MR) simultaneously.

$$G_{MI} = \operatorname{argmax}_{f_i \in F} \left(MI(C; f_i) - \frac{1}{|S|} \sum_{f_s \in S} MR \right) \quad (6)$$

Where $MI(C; f)$ is the amount of information that feature f_i carries about the class C . MR, in Eq. (6), is the relative minimum redundancy of feature f_i against feature f_s and is defined by Eq. (7).

$$MR = \frac{MI(f_i, f_s)}{MI(C; f_i)} \quad (7)$$

Where $f_i \in F$ and $f_s \in S$. In the case of $MI(C; f_i) = 0$, feature f_i can be removed without computing Eq. (6). If f_i and f_s are relatively highly dependent with respect to $MI(C; f_i)$, the feature f_i contributes to the redundancy. Thus to reduce the number of features that needs to be observed, a numerical threshold (Th=0) value is applied over the G_{MI} in Eq. (6) such that the obtained G_{MI} can be have the following cases;

Case 1: $G_{MI} = 0$. In this case, the current feature f_i subjects to the redundancy for the particular class C . Since the feature f_i cannot approximate any information about the class C , it is subjected to removal from the set S . This signifies the feature f_i 's irrelevancy or unimportance to the class C .

Case 2: $G_{MI} \neq 0$. In this case, the current feature f_i is relevant or important to the class C because it can provide some extra information to the classification after the selection of feature subset S . Thus the current feature f_i is added to the subset S .

Algorithm 2: Flexible Entropy based Feature Selection (FEFS)

- Step 1.** Initialization: Set $F \leftarrow$ “Initial Set of all features”, Set $S \leftarrow$ “Empty Set”.
- Step 2.** Compute MI with the output class C : for every feature $f_i \in F$ Compute $MI(C; f_i)$.
- Step 3.** Find the first feature f that maximizes $MI(C; f_i)$: set $F \leftarrow F \setminus \{f_i\}$; Set $S \leftarrow \{f_s\}$.
- Step 4.** Find the further features f as one that maximizes $G_{MI} = \operatorname{argmax}_{f_i \in F} \left(MI(C; f_i) - \frac{1}{|S|} \sum_{f_s \in S} MR \right)$;
- Set $F \leftarrow F \setminus \{f_i\}$;
- Step 5:** If $G_{MI} \neq 0$, Set $S \leftarrow S \cup \{f_s\}$.
- Step 6.** Output the set S containing the selected features.

4.2 Multi-class SVM (MC-SVM)

Since there exists multiple types of disturbances in the power systems, Multi-class classification is required. SVM is one of the most popular non-linear classifier which gives optimal performance in the PQD detection. At a time SVM classifies only two classes. When the SVM is applied on the multiple way, the multiple types of disturbances were detected. Binary tree, one-against-all and one-against-one are the most popular techniques in SVM multiclass classification. The one-against-one SVM classification technique requires $k(k-1)/2$ two-class SVM classifiers where everyone is trained on data for two classes. The one-against-all SVM classification technique requires ‘ k ’, two-class SVM classifiers to perform the detection. Finally, the binary tree SVM technique requires $k-1$, two class SVM classifiers for a test of k classes. Thus, the binary tree SVM classification technique is used here to perform PQD detection. Based on the properties of various power quality disturbances types, multiple SVM classifiers are implemented here to detect different PQDs such as Sag, Swell, harmonics, interruption, sag with harmonics, swell with harmonics, flicker, and oscillatory transients. Initially the first SVM (SVM1), SVM1 is trained to classify the total signals into normal state and disturbances state. During this classification, the SVM produces only +1 and -1 as the output. At SVM1, if the signal representing the normal state, the output of SVM1 is +1 otherwise it is -1. Further, the obtained samples of disturbance state are given as input to SVM2. The further two class SVMs are trained to classify the further disturbances. According to the binary tree multiclass SVM, the number of two class SVMs required for this PQD detection is 8. The number of classes need to be classified, $k=9$ (1.normal, 2.Sag, 3.Swell, 4.Transient, 5.Interruption, 6.Flicker, 7.Harmonics, 8.Sag with harmonics and 9.swell with harmonics). Hence the number of two class SVMs required is $k-1=9-1=8$. The binary tree multi-class SVM model accomplished in this paper is represented in Fig. 2.

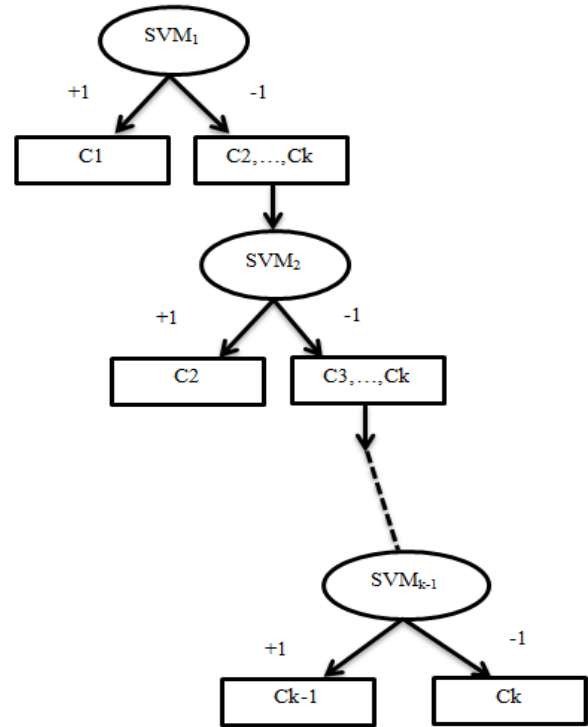


Figure.2 The scheme of PQD detection based on binary tree multi class SVM

After extracting the features, the SVM is accomplished for training of the dataset. At the optimal solution, decision function using SVM is modeled as

$$f(t) = \operatorname{sgn}(\sum_{i=1}^p (\alpha_i - \hat{\alpha}_i) K(t_i, t_j) + b) \quad (8)$$

Where α_i and $\hat{\alpha}_i$ are the Lagrange multiplier coefficient for the i^{th} sample, $K(t_i, t_j)$ is the kernel function and b is an arbitrary constant.

Here the proposed approach accomplished the most popular RBF kernel as a kernel function to perform the decision making. Since the SVM with RBF kernel is declared as an effective combination from earlier studies, the RBF kernel is used as a kernel function. The mathematical formulation of RBF kernel is given as

Table 2. Signals generated with varying control parameters

PQD	Varying parameters	Total signals
Normal	$\omega_n = 2\pi \times 50$ rad/sec	10
Swell	$T=0.4, 9 \times T=3.6, \alpha = 0.1, 0.2, 0.3, \dots, 0.8$	$33 \times 8=264$
Sag	$T=0.4, 9 \times T=3.6, \alpha = 0.1, 0.2, 0.3, \dots, 0.9$	$33 \times 9=297$
Flicker	$\alpha = 0.1, 0.2, \beta = 5, 6, 7, \dots, 20$	$16 \times 2=32$
Interruption	$T=0.4, 9 \times T=3.6, \alpha = 0.9$ and 1.0	$33 \times 2=66$
Oscillatory transient	$T=0.5 \times T$ to $3 \times T=0.2$ to 1.2, $f_n = 300, 400, \dots, 900$, $\alpha = 0.1, 0.2, 0.3, \dots, 0.8$	$6 \times 8 \times 6=288$
Harmonics	$\alpha_3, \alpha_5, \alpha_7=0.05, 0.1, 0.15$	$3 \times 3 \times 3=27$
Swell with harmonics	$\alpha_3, \alpha_5, \alpha_7=0.05, 0.1, 0.15, \alpha = 0.1, 0.2, 0.3, \dots, 0.8$	$27 \times 8=216$
Sag with Harmonics	$\alpha_3, \alpha_5, \alpha_7=0.05, 0.1, 0.15, \alpha = 0.1, 0.2, 0.3, \dots, 0.9$	$27 \times 9=243$
Total		1444

$$K(t_i, t_j) = \exp\left(\frac{-\|t_i - t_j\|^2}{\sigma^2}\right), \sigma \in R \tag{9}$$

According to the functional theory, as long as the function $K(t_i, t_j)$ satisfies Mercer’s condition, it can be denoted as a positive definite kernel.

5. Performance evaluation

To simulate the proposed PQD framework, MATLAB software was used. Different signals are generated according to the formulae shown in Table 1 by varying the control parameters for both normal and disturbances. To alleviate the effect of variations in control parameters, totally 1444 different cases are generated through the parametric equations. This is performed for every class of event by randomly changing the various control parameters. Depth, starting time and duration of disturbances are the parameters used to vary the classes of events. Here the depth of event is defined as the change in the amplitude of a signal. Further the starting time defines the time at which the disturbance starts and the duration of event defines the period of time up to which the disturbance exists [30].

To obtain a sag, swell and interruption disturbance with varying amplitudes, the control parameter α is varied from 0.1 to 0.9, 0.1 to 0.8 and from 0.9 to 1 respectively. Further the time periods t_1 and t_2 are varied to obtain the disturbances with different starting times and durations. The varying values of t_1 results in the disturbance with different starting times and different durations while the varying values of t_2 results in the disturbance with different durations only. Simultaneously to obtain different harmonics, the value of $\alpha_3, \alpha_5, \alpha_7$ are varied from 0.05 to 0.15 and also considering the constraint, $\sum \alpha_i^2 = 1$. To further attain the sag with

harmonics and swell with harmonics, the respective control parameters are varied according to their range. The oscillatory transient disturbances with different amplitudes, starting times and durations is obtained by varying the control parameters α, τ, t_1 and t_2 . Finally different flicker disturbances are obtained by varying the control parameters α and β . Here the variations in the control parameter α results in the flicker disturbances with different amplitudes and the variations in the control parameter β results in the flicker disturbances with starting times and durations. The total signals generated by varying the control parameters for all disturbances are formulated in Table 2.

All the generated samples are divided into two groups of train and test. Out of the generated 1444 signals, 1083 signals are used for training and 361 signals are used for testing. Initially the complete train set is subjected to the feature extraction followed by training through SVM. Further the test set is subjected to testing process and the classified results are formulated in the following Table 3. To check the efficiency of proposed approach, a further simulation carried out in the presence of noise. In this simulation, all the disturbance signals are subjected to noise addition and then processed for testing. The obtained classification results in this case are represented in the following Table 4.

For the obtained confusion matrices in Table 3 and Table 4, the accuracy and false alarm rate are measured according to the following formula.

$$\text{False Alarm Rate} = \frac{FP}{TN+FP} \tag{10}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{11}$$

Where
TP= True Positives

Table 3. Classification result of testing signals (Confusion matrix)

	C1	C2	C3	C4	C5	C6	C7	C8	C9	Total
C1	3	0	0	0	0	0	0	0	0	3
C2	0	62	1	0	1	0	0	1	0	65
C3	0	1	71	0	1	0	0	0	1	74
C4	1	0	0	7	0	0	0	0	0	8
C5	0	1	0	0	16	0	0	0	0	17
C6	1	0	0	1	0	70	0	0	0	72
C7	0	0	0	0	0	0	7	0	0	7
C8	0	1	0	0	0	0	0	52	1	54
C9	0	0	1	0	0	0	0	0	60	61
Total										361
Overall Average Accuracy										98.65%
False Alarm Rate (FAR)										0.0135

Table 4. Classification result of testing signals in the presence of noise (Confusion matrix)

	C1	C2	C3	C4	C5	C6	C7	C8	C9	Total
C1	3	0	0	0	0	0	0	0	0	3
C2	0	60	1	0	2	0	0	2	0	65
C3	0	3	69	0	1	0	0	0	1	74
C4	2	0	0	5	1	0	0	0	0	8
C5	0	1	1	0	14	0	0	1	0	17
C6	1	0	0	2	0	68	0	0	1	72
C7	0	0	0	0	0	0	7	0	0	7
C8	0	1	1	0	0	0	0	51	1	54
C9	0	1	1	0	0	0	0	0	59	61
Total										361
Overall Average Accuracy										95.23%
False Alarm Rate (FAR)										0.0538

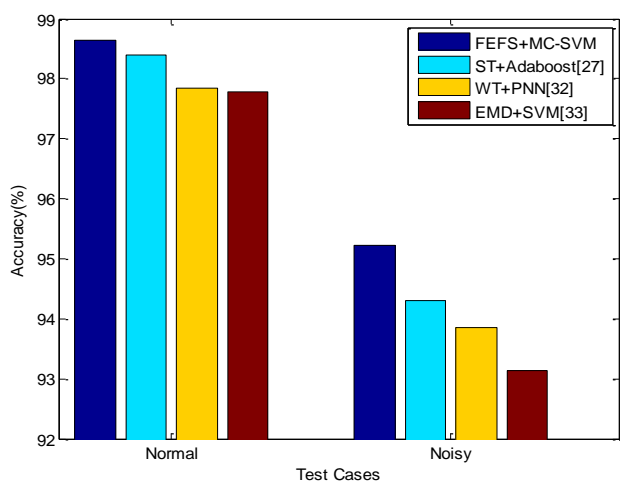


Figure.3 Accuracy comparison

TN = True Negatives
 FP= False Positives
 FN = False Negatives

By substituting the obtained TPs, TNs, FPs and FNs from the confusion matrix into the above specified formulae, the Overall Average Accuracy achieved is 98.65% and the the False Alarm Rate is 0.0135for normal test case (table.3). Similarly the

Overall Average Accuracy achieved is 95.23% and the the False Alarm Rate is 0.0538 for the noisy test case (table.4). Further the proposed FEFS+MC-SVM is compared with the conventional approaches with respect to the Accuracy, False Alarm Rate and Computational time and are described below.

Fig. 3 illustrates the comparative analysis between the proposed approach and conventional approaches with respect to the classification accuracy. The simulation is carried out in two cases, one is under normal and another is in the presence of external noise. From Fig. 3, it can be observed that the accuracy of proposed approach is high compared to the conventional S-Transform (ST)+Adaboost [27], Wavelet Transform (WT)+Probabilistic Neural Network (PNN) [32] and Empirical Mode Decomposition (EMD)+Support Vector Machine (SVM) [33] approaches. Since the proposed approach evaluates the mutual entropies between the signals, the signals which needs to be classified should have a perfect discrimination with all the other signals. And the derived mutual relationship also relates the signals both linearly and non-linearly by which any type of disturbance can be classified

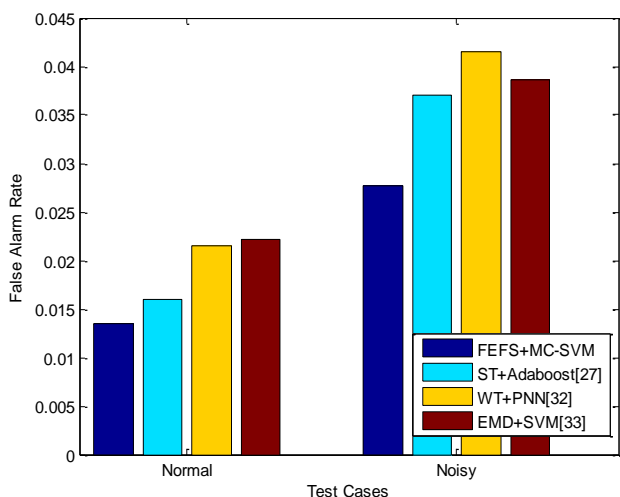


Figure.4 False alarm rate comparison

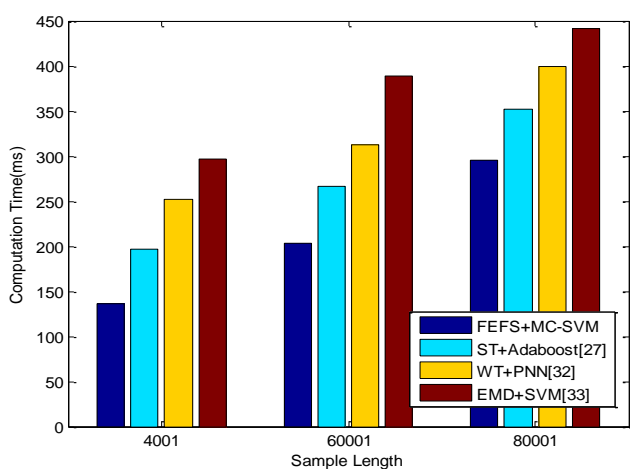


Figure.5 Computation time comparison

exactly. In the case on noise presence also, the proposed approach has shown an excellent performance.

According to Fig. 3, the accuracy achieved for normal test case is more compared to the noisy test case. In both cases, the proposed FEFS+MC-SVM approach achieved a better accuracy compared to the conventional approaches. Since the conventional approached accomplished the transform domain techniques such as S-transform, Wavelet transform and the EMD to extract the required feature set from the signals. Though these techniques reveals more information about the characteristics of disturbance type, the computational complexity is observed to be high. Thus the proposed approach succeeded in the performance enhancement with respect to the accuracy within the less computational time.

Fig. 4 illustrates the comparative analysis between the proposed approach and conventional approaches with respect to the False Alarm Rate (FAR). The FAR of the proposed approach is observed to be less compared to the conventional approaches. As much as the accuracy increases, the

FAR decreases accordingly. From the above Fig. 4, the proposed approach has obtained a less FAR compared to the conventional approaches. Due to the evaluation of both linear and non-linear dependencies between the signals, the proposed FEFS technique can provide much discrimination between the disturbances by which the classifier can classify the disturbance accordingly. This helps in the reduction of false alarm rate. In both normal and noisy cases, the proposed FEFS+MC-SVM is achieved less FAR compared to the conventional ST+Adaboost, WT+PNN and EMD+SVM. According to the Fig. 4, in the noisy test case, the conventional EMD+SVM is observed to have less FAR compared to the WT+PNN, due to the efficiency of SVM in the classification aspects.

The main aim of the proposed PQD detection and classification approach is to achieve an increased detection accuracy followed by the reduced computation time. Unlike the conventional approaches which applied transform techniques like Wavelet transform, S-Transform, the proposed approach works directly over the disturbance signals to reduce the computation time. Further the conventional approaches constitutes more computational time in the classification due to the time-frequency analysis, but the computation time of proposed approach is observed to be less because the proposed approach optimized the entropy based feature selection by adopting a flexible process in the feature selection. This mechanism further reduces the feature set by which the time taken for training as well as for testing reduces significantly. As shown in Fig. 5, the computational time is observed to be increasing for increment in the length of samples. But the increment due to the proposed FEFS+MC-SVM is less compared to the conventional approaches.

6. Conclusion and future scope

A new PQD detection and classification mechanism is developed in this paper based on the Flexible entropy based feature selection and multi-class Support vector machine. The proposed work developed a modified entropy based feature selection to extract an optimal feature set and further the SVM is accomplished in hierarchical fashion to achieve increased individual class accuracy. The extracted optimal features also help in the reduction of computation time over the system. Simulation carried out under two cases, such as normal and noisy and for every case, the performance is measured through the performance metrics like accuracy, false alarm rate and computation time.

The accuracy improvement through the proposed method in the detection of PQDs is obtained as 0.25% and 0.80% and 0.87% from the conventional ST + Adaboost, WT+PNN and EMD+SVM in the normal test case. In the case of noisy test case, it is approximately 0.93%, 1.38% and 2.09% respectively. Further False alarm rate is reduced to 0.0059, 0.0109 and 0.0148 from the conventional ST + Adaboost, WT+PNN and EMD+SVM approaches, respectively. Along with these subjective evaluations, the proposed approach is also achieved a reduced computational time and the reductions are 60.8040 ms, 116.54 ms and 160.23 ms from the conventional approaches ST + Adaboost, WT+PNN and EMD+SVM, respectively.

To achieve a further more accurate results followed by reduced computational overhead, this work can be extended by adopting a two stage detection framework by combining the linear and non-linear approaches. A new feature selection which gives the proper linear relations and one more for non-linear relations can be accomplished to achieve an increased performance in different types of disturbances.

References

- [1] D. G. Lieberman, R. J. R. Troncoso, and R. A. O. Rios, "Techniques and methodologies for power quality analysis and disturbances classification in power systems: a review", *IET Generation Transmission and Distribution*, Vol.5, No.4, pp. 519–529, 2011.
- [2] M. H. J. Bollen, "What is power quality?", *Electrical Power Systems Research*, Vol. 66, No. 1, pp. 5–14, 2003.
- [3] T. Lin and A. Domijan, "On power quality indices and real time measurement", *IEEE Transactions on Power Delivery*, Vol. 20, No. 4, pp. 2552–2562, 2005.
- [4] L. Coppola, Q. Liu, and S. Buso, "Wavelet transform as an alternative to the short-time Fourier transform for the study of conducted noise in power electronics", *IEEE Transactions on Industrial Electronics*, Vol.55, No.2, pp. 880–887, 2008.
- [5] J. G. M. S. Decanini, M. S. T. Neto, and F. C. V. Malange, "Detection and classification of voltage disturbances using a fuzzy-ARTMAP-wavelet network", *Electrical Power Systems Research*, Vol.81, No.12, pp. 2057–2065, 2011.
- [6] J. Barros, M. Apraiz, and R. I. Diego, "A virtual measurement instrument for electrical power quality analysis using wavelets", *Measurement*, Vol.42, No.2, pp. 298–307, 2009.
- [7] T. Nguyen and Y. Liao, "Power quality disturbance classification utilizing S-transform and binary feature matrix method", *Electrical Power Systems Research*, Vol.79, No.4, pp. 569–575, 2009.
- [8] M. Uyar, S. Yildirim, and M. T. Gencoglu, "An expert system based on S-transform and neural network for automatic classification of power quality disturbances", *Expert Systems Applications*, Vol.36, No.3, pp. 5962–5975, 2009.
- [9] A. A. Abdelsalam, A. A. Eldesouky, and A. A. Sallam, "Characterization of power quality disturbances using hybrid technique of linear Kalman filter and fuzzy expert system", *Electrical Power Systems Research*, Vol.83, No.1, pp. 41–50, 2012.
- [10] S. H. Cho, G. Jang, and S. H. Kwon, "Time-frequency analysis of power quality disturbances via the Gabor–Wigner transform", *IEEE Transactions on Power Delivery*, Vol.25, No.1, pp. 494–499, 2010.
- [11] N. Senroy, S. Suryanarayanan, and P. F. Ribeiro, "An improved Hilbert–Huang method for analysis of time-varying waveforms in power quality", *IEEE Transactions on Power Systems*, Vol.22, No.4, pp. 1843–1850, 2007.
- [12] B. V. Krishna and K. Baskaran, "Parallel computing for efficient time–frequency feature extraction of power quality disturbances", *IET Signal Processing*, Vol.7, No.4, pp. 312–326, 2013.
- [13] S. He, K. Li, and M. Zhang, "A Real-Time Power Quality Disturbances Classification Using Hybrid Method Based on S-Transform and Dynamics," *IEEE Transactions on Instrumentation and Measurement*, Vol. 62, No. 9, pp. 2465–2475, 2013.
- [14] P. K. Ray, N. Kishor, and S. R. Mohanty, "Islanding and power quality disturbance detection in grid-connected hybrid power system using wavelet and-transform", *IEEE Transactions on Smart Grid*, Vol.3, No.3, pp. 1082–1094, 2012.
- [15] S. R. Mohanty, N. Kishor, and P. K. Ray, "Comparative study of advanced signal processing techniques for islanding detection in a hybrid distributed generation system", *IEEE Transactions on Sustainable Energy*, Vol.6, No.1, pp. 122–131, 2015.
- [16] R. Yan and R. X. Gao, "A tour of the Hilbert–Huang transform: An empirical tool for signal analysis," *IEEE Instrumentation Measurement Magazine*, Vol. 10, No. 5, pp. 40–45, 2007.

- [17] Y. Bao and J. Xie, “Contrast Studt on power quality detection using EMD and EEMD”, In: *Proc. of International Conf. on Consumer Electronics and Communication Networks*, pp.2074-2077. 2011.
- [18] N. C. F. Tse, J. Y. C. Chan, W. H. Lau, and L. L. Lai, “Hybrid wavelet and Hilbert transform with frequency-shifting decomposition for power quality analysis”, *IEEE Transactions on Instrumentation Measurement*, Vol. 61, No. 12, pp. 3225–3233, 2012.
- [19] P. K. Dash, B. K. Panigrahi, and G. Panda, “Power quality analysis using S-transform,” *IEEE Transactions on Power Delivery*, Vol. 18, No. 2, pp. 406–411, 2003.
- [20] C. C. Liao, H. T. Yang, and H.H. Chang, “Denoising techniques with a spatial noise-suppression method for wavelet-based power quality monitoring”, *IEEE Transactions on Instrumentation Measurement*, Vol. 60, No. 6, pp. 1986–1996, 2011.
- [21] P. Kanirajan and V. S. Kumar, “Power quality disturbance detection and classification using wavelet and RBFNN”, *Applied Soft Computing*, Vol.35, No.1, pp.470–481, 2015.
- [22] S. Dalai, D. Dey, and B. Chatterjee, “Cross-spectrum analysis-based scheme for multiple power quality disturbance sensing device”, *IEEE Sensors Journal*, Vol.15, No.7, pp. 3989–3997, 2015.
- [23] R. Kumar, B. Singh, and D. T. Shahani, “Recognition of single-stage and multiple power quality events using Hilbert–Huang transform and probabilistic neural network”, *Electrical Power Components and Systems*, Vol.43, No.6, pp. 607–619, 2015.
- [24] Z. Moravej, S. A. Banihashemi, and M. H. Velayati, “Power quality events classification and recognition using a novel support vector algorithm”, *Energy Conversations and Management*, Vol.50, No.3, pp. 3071–3077, 2009.
- [25] A. Dasgupta, S. Debnath, and A. Das, “Transmission line fault detection and classification using cross-correlation and K-nearest neighbor”, *International Journal of Knowledge Based Intelligent. Engineering Systems*, Vol.19, No.10, pp. 183–189, 2015.
- [26] H. Behera, P. Dash, and B. Biswal, “Power quality time series data mining using S transform and fuzzy expert system”, *Applied Soft Computing*, Vol.10, No.3, pp. 945–955, 2010.
- [27] M. V. Reddy and R. Sodhi, “A rule-based S-transform and AdaBoost based approach for power quality assessment”, *Electrical Power Systems Research*, Vol.134, No.11, pp. 66–79, 2016.
- [28] R. Kumar, B. Singh, and T. Shahani, “Recognition of power-quality disturbances using S-transform-based ANN classifier and rule-based decision tree”, *IEEE Transactions on Industrial Applications*, pp. 1249–1258, 2015
- [29] R. Battiti, “Using mutual information for selecting features in supervised neural net learning”, *IEEE Transactions on Neural Networks*, Vol.5, No.4, pp.537-550, 1994.
- [30] “IEEE recommended practice for monitoring electric power quality”, 2009.
- [31] A. Enshae and P. Enshae, “A new S-Transform based method for Identification of power quality Disturbances”, *Arab Journal of Science Engineering*, Vol.43, No.6, pp. 2817-2832, 2017.
- [32] S. Khokar and A. A. M. Zin, “A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network”, *Measurement*, Vol.95, No.1, pp. 246-259, 2017.
- [33] O. P. Mahela and A. G. Shaik, “Recognition of power quality disturbances using S-transform based ruled decision tree and fuzzy C-means clustering classifiers”, *Applied Soft Computing*, Vol.59, No.3, pp.243-257, 2017.
- [34] N. R. Babu and B. J. Mohan, “Fault Classification in power systems using EMD and SVM”, *Ain Shams Engineering Journal*, Vol.8, No.2, pp.103-111, 2017.