

Online power quality disturbance detection by support vector machine in smart meter

Imtiaz PARVEZ¹, Maryamossadat AGHILI², Arif I. SARWAT¹,
Shahinur RAHMAN¹, Fahmida ALAM¹



Abstract Power quality assessment is an important performance measurement in smart grids. Utility companies are interested in power quality monitoring even in the low level distribution side such as smart meters. Addressing this issue, in this study, we propose segregation of the power disturbance from regular values using one-class support vector machine (OCSVM). To precisely detect the power disturbances of a voltage wave, some practical wavelet filters are applied. Considering the unlimited types of waveform abnormalities, OCSVM is picked as a semi-supervised machine learning algorithm which needs to be trained solely on a relatively large sample of normal data. This model is able to automatically detect the existence of any types of disturbances in real time, even unknown types which are not available in the training time. In the case of existence, the disturbances are further classified into

different types such as sag, swell, transients and unbalanced. Being light weighted and fast, the proposed technique can be integrated into smart grid devices such as smart meter in order to perform a real-time disturbance monitoring. The continuous monitoring of power quality in smart meters will give helpful insight for quality power transmission and management.

Keywords Machine learning, One-class support vector machine, Power quality, Disturbances, Smart grid, Smart meter

1 Introduction

An efficient, robust and smart power grid system is the key driving force for the development of the energy sector in the twenty-first century [1]. Smart grid plays a vital role to improve efficiency via reducing carbon emissions. In recent years, smart grid has become the emerging trend because of its flexible and steadfast energy distribution through a duplex communication system between the supplier control hub and the smart meters on the consumer end [2–4]. A smart meter is employed for monitoring power consumption with two-way communication and consumers can access detailed information about their usages and the quality of the service. These acquired data provide information to the customers regarding their overall power consumption, how they consume power, and also inform them in a way to reduce their consumption [5–9]. Similarly, the finer resolution of data is used for the supply to loads during peak time by the power companies. Hence, the ultimate aspiration of the smart meter is not only to increase effectiveness of the power management, but to integrate new power generation techniques into the

CrossCheck date: 12 September 2018

Received: 18 November 2017 / Accepted: 12 September 2018 /
Published online: 26 December 2018

© The Author(s) 2018

✉ Arif I. SARWAT
asarwat@fiu.edu

Imtiaz PARVEZ
iparv001@fiu.edu

Maryamossadat AGHILI
maghi001@fiu.edu

Shahinur RAHMAN
srahm026@fiu.edu

Fahmida ALAM
falam007@fiu.edu

¹ Department of Electrical and Computer Engineering, Florida International University, Miami, USA

² School of Computing and Information Sciences, Florida International University, Miami, USA

distribution level. Moreover, it reduces excessive power consumption and alerts people about it [10]. Now-a-days, advanced smart meters are equipped with monitoring of voltage disturbances, harmonics and power factor that assist power companies for better understanding of the power quality (PQ) [11].

The quality of power and its assurance have become a prime factor for the utility sector in the recent past. The equipments on the consumer ends are highly sensitive to numerous PQ problems. Apart from this, they have negative impact on the power supply system [12, 13]. Poor multifarious power system interruptions such as notches, transients, momentary disorder, voltage swell and sag, under-voltage, over-voltage, harmonic, etc., are caused due to poor quality of power supply [14, 15]. The majority of industrial electronic instruments are highly sensitive to PQ issues. Therefore, a cost effective solution is essential to keep them away from the malfunction and unnecessary expenses to install PQ monitoring system [16]. Simultaneously, it is necessary to know the root of the disturbance for both sides - utility and supply, before taking suitable mitigating action to facilitate electric PQ and energy efficient system. A continuous in-sit observation of PQ enables consumers to inform the distribution companies regarding the issues. Therefore, a real-time on-line quality monitoring and assurance is the only solution to correct various actions such as reduction of consumption, power factor improvement and load demand balance.

In this paper, we propose a combination of discrete wavelet transform (DWT) and two types of machine learning-based algorithms to extract features from the power signal and sort out PQ issues. From a voltage transformer, continuous data are given as input to wavelet transform (WT) filters, then the output feature set is fed to a model trained by one-class support vector machine (OCSVM). The disturbance is detected and tagged as abnormalities if any. In the case of disturbance detection, another multi-class support vector machine (SVM) will analyze the corrupted data to define the disturbance types.

In the literature, various solutions for this problem are presented which are mainly composed of two steps: ① extracting useful information from the waveform as a feature set; ② training a model based on the provided feature sets and labeled data for supervised pattern recognition. However, the driver events and various components of the distribution network along with scarcity of all kinds of abnormal data in the training period, hinder the effective applications of the proposed methods. Also for PQ detection at the distribution level, such as meter level, we need light but robust method due to the limited computational resources in smart meters. Based on the aforementioned

data driven pipeline for this problem, majority of the research papers defined and simulated the possible disturbance patterns and then trained their models based on the simulated data. While these methods can detect the abnormalities of the defined pattern, they fail to capture the unknown types of abnormalities. In this paper, with focus on the real power data, we redress this shortcoming by applying a semi-supervised technique.

Our proposed method applies a cascade two-level classification algorithm on a simulated data set of power voltage after carefully preprocessing the data. The contribution of our study is:

- 1) To develop real-time PQ monitoring system for smart meter level.
- 2) To detect and classify any type of disturbance even the novel form.
- 3) To provide a light but robust method for PQ assessment.

The rest of this paper is as follows: In Section 2, the literature review and preliminaries of our applied techniques have been described. The system model is presented in Section 3. The simulation results are illustrated in Section 4. Finally, a brief conclusion is included in Section 5.

2 Literature review

In the literature, several research directions can be found for PQ disturbance detection. One direction explores the efficient, accurate and high-speed techniques of feature extraction from signals using various methods such as Fourier transform (FT), S-transform (ST), WT, root mean square (RMS), fast Fourier transform (FFT), and fast dyadic Fourier transform (FDFT), etc. [17–20]. Another direction investigates the optimal sets of features out of all the extracted features from the signal [21, 22]. This direction performs PQ disturbance detection and segregation using diverse artificial intelligence (AI) and machine learning (ML) techniques such as fuzzy logic [23], neural network [24], SVM [25, 26], decision tree [27, 28], expert system [29–31], and hidden Markov model [32]. Most of the researchers not only considered the classification and anomaly detection performance but also gave importance on the computational efficiency and processing speed [27].

FT being one of the earliest techniques for signal analysis can detect the existence of a specific frequency in power waveforms. However, it is incapable of recognizing time evolving effect of non-stationary signals [33]. Short-time Fourier transform (STFT) was proposed afterwards to



solve the aforementioned problem of FT. Although STFT or windowed FT improves the FT, it suffers from fixed window width. Fixed window size cannot analyze low-frequency and high-frequency of transient signals at the same time which is important for disturbance detection [34]. Later, WT gained popularity among researchers because of its capability to analyze power system's non-stationary root in various disturbances [35, 36].

In [37], authors applied WT on the wave as the first step to remove the noise in the signal. Following this, parameters such as peak value and periods can be calculated by FFT, which helps to identify the disturbance types. Gaouda et al. exploited the WT combined with K-nearest neighbor (KNN) for signal feature extraction and classification [38]. Although this method performed well in the low level of noise, its accuracy degraded when the level of noise increases (i.e. noise level is greater than 0.5%). In [39], the authors proposed a less computationally expensive algorithm which extracted the features of the signal by two relatively simple and quick methods of discrete Fourier transform (DFT) and RMS. Based on a limited set of features and a rule-based decision tree, they detected and defined the types of nine categories of signal disturbances in real time.

One of the earliest approaches of PQ disturbance detection conducts point-to-point comparison of the signal in a cycle. Despite its simplicity, it has poor performance when the pattern of disturbance is repetitive in nature [40]. Ghosh and Lubkeman as the pioneer, used artificial neural network (ANN) for automatic waveform classification [41]. They applied two different variations of the neural networks in the unprocessed signal. Later in 2001, a rule-based system was introduced in [42] for disturbance classification. Although this method is simple to implement, its model can not be generalized easily, and with the growth of disturbance types and patterns, the increasing number of "If" and "Else" rules hinders the system efficiency and capacity.

In [26], the authors utilized SVM with radial basis function (RBF) kernel to detect the disturbance patterns in the three-phase simulated signals. The results show a promising performance with an acceptable accuracy. In [25], Axelberg et al. applied the SVM model for classifying the voltage disturbance on the real and synthetic data. Extracted features from the data such as minimum RMS voltage, harmonic components, symmetric components, RMS voltage at selected time instants, total harmonic distortion, and the duration of the disturbance composed an informative sample space [25, 43]. In [44], weighted SVM (WSVM), FT and WT coupled hand in hand to classify five disturbance categories. As both of selected features and tuning parameters influence the classification performance, Moravej et al. incorporated a two-stage feature selection by

the combination of SVM and digital signal processing (DSP) techniques [21]. Before applying the SVM on the feature set, they performed mutual information feature selection technique (MIFS) and correlation feature selection (CFS) for prominent features selection and redundancy elimination. Eristi et al., in a similar approach, added feature extraction and selection methods along with WT, aiming to reduce the feature space in order to reach higher accuracy and lower resource consumption [45]. However, all these proposed approaches are mostly for the transmission and generation levels of the grid considering possession of huge computational capacity and resources.

In order to analyze PQ at the distribution level, the utility company looks forward for techniques at the smart meter level. But the meters have limited memory, and the processing capacity requires the signal processing should be computationally light weight and fast enough. In this regard, very limited works have been found in the literature. Borges et al. proposed their technique which is executable inside the smart meter [27]. They extracted some of the features by using the FFT that are computationally low cost and routinely integrated in hardware and they captured the remaining features in time domain. For detecting the type of the disturbance, they exploited the decision tree and ANN and reported higher than 90% precision rates. All the processes were designed to be embedded in smart meter.

Reviewing above studies, we propose new method SVM for PQ detection in smart meter. Being computationally light weight and guaranteed global optimum value, our technique is more feasible in smart meter [46]. Our method not only detects the disturbances but also classifies them with higher precession compared to other methods such as Isolation Forest. Besides that, our method provides real-time PQ monitoring at meter level and can detect any kind of new disturbances.

2.1 PQ issues

In the power system, even a short period of disturbances might lead to a huge amount of power losses. Hence, PQ monitoring becomes one of the major services of the power companies. One way of guaranteeing the PQ is the precise monitoring of the waveform, capturing any distortion form and rooting out the cause of it.

Some common PQ interferences and their effects in the distribution sites are provided here to emphasize the importance of an online accurate disruption detection.

- 1) Unbalanced waveform: frequency is kept almost steady in large interconnected distribution networks, and changes are a rare incident. Therefore, frequency deviation is frequent on smaller networks, specially for

those supplied by on-site generators. Lower value of the inertia constant due to reduction of connected generators to the power system is the main reason behind it. It may cause damage of electrical equipment and have the worse impact on clock (motor driven) speed.

- 2) Transients: transients can be defined as unexpected changes of voltage or current from rated values. Its duration is very small, typically lasting from 200 μ s to 1 s. Lightning strikes, electrostatic discharges (ESD), poor grounding, load switching, and faulty wiring are the main reasons behind it. Transients can delete or change computer data and make hard to identify calculation errors. In severe cases, they can damage electronic instruments and hamper power system operations.
- 3) Voltage sag: a voltage sag is a short span drop of RMS voltage. Undervoltages are defined as voltage drop greater than two minutes duration. Common reasons of undervoltages and voltage sags are faults (short circuits) on the power system, weather factors, motor starting, inclusions of customer load and introduction of large loads in the power system. Sags can shut down computers and other sensitive equipments within a moment. Undervoltage conditions can hamper certain types of electrical instruments.
- 4) Voltage swell: voltage swell is momentary rise in voltage magnitude. Overvoltages are defined as voltage increase greater than two minutes duration. Overvoltages and voltage swells are typically generated by deviations of power line switching and large load. If voltage rise has too high value, it may destroy electrical instruments and shut down power systems. The consumer's voltage controlling device is unable to act as quickly as to give protection from all swells or sags.
- 5) Interruptions waveform: when voltage magnitude reduces to zero, interruptions occur in the power system. They are categorized as long-term, temporary or momentary. Momentary disruptions happen when utility system is being interrupted and automatically restored within a short duration (less than 2 s).

Researchers have considered diverse types of disturbances and they trained their models based on their predefined types. Uyar et al. [47], Koleva [48] and Kostadinov [49] defined 6 classes of disturbances: sag, swell, outage, harmonic, swell with harmonic and sag with harmonic. Sahani [50] composed 9 classes, including momentary interruption, sag, swell, harmonics, flicker, notch, spike, transient, and sag with harmonics. Khokhar [51] presented 6 more categories in addition to [52] which are swell with harmonics, interruption and harmonics, impulsive

transient, flicker with harmonics, flicker with swell, and flicker with sag.

Considering the fact that various types of disturbances can occur together and make several new disturbance categories, defining a closed set of disturbance has an obvious drawback of missing some undefined, and unknown pairs of anomalies [52]. This is one of the limitation of many proposed automatic disturbance detection and classification techniques. To address this issue, by considering the fact that any types of unknown disturbance can possibly happen together and results a new form of disturbance with a varying degree of noise, our model (OCSVM) is trained particularly on the normal dataset. Training a semi supervised model on the abundant set of normal data granted the model the ability of abnormality detection with almost 93% accuracy. After detection of the disturbance by OCSVM model, a complementary multi-class classification model will capture the correct label of the disturbances. This method will at least detect the disturbance even if it fails in capturing the appropriate type. This technique is highly advantageous in the uncertain and real environment while the types of abnormality cannot be accurately predicted prior to the application of the system or when enough samples of different abnormalities are not available for training the system.

2.2 WT

WT has emerged as a useful tool for presenting a signal in the time-frequency domain. It is superior than FT when the frequency content of a signal is non stationary. This transform provides both time and frequency information required for extracting transient information from non stationary signals. WT has multiple implementation such as DWT, continuous wavelet transform (CWT) and wavelet packet transform (WPT). Among all, DWT and WPT have been applied in real world problems.

In our study, we apply DWT which is one of the most practical types of WT. DWT is mostly used for decomposing a time series signal $S(t)$ into components such as detailed coefficients and approximation coefficients [53]. The low pass and high pass filter yield approximation coefficients and detailed coefficients, respectively. The output of low pass filter is further decomposed into level 2 components, consisting of approximation coefficients and detailed coefficients which is shown in Fig. 1.

2.3 Signal features extraction and selection

Signal features extraction and selection are two important steps for signal classification. Effective features set can heavily impact the performance of the classifier. Feature selection reappears a subset of the original features while



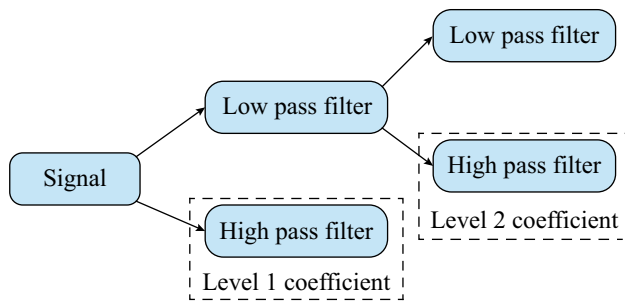


Fig. 1 Two-level filter analysis in DWT

feature extraction produces new feature from signal having native features.

Each signal carries many native features of basically two types – irrelevant and redundant. Removing those does not cause information loss [54]. Irrelevant and redundant features are two separate ideas. One pertinent feature can be considered as redundant in the existence of another feature with which it is powerfully connected [55]. Though wavelet and multi-resolution analysis (MRA) of signal extract the significant information, they are inefficient and sometimes misleading to apply the classifier on the large feature set. The optimized distinct features must be extracted and selected in a way to reduce feature vector's dimension and maximize classification performance.

Many features such as entropy (Ent), energy (E), mean value (M), standard deviation (Sd), and RMS were widely used in several research papers as the most informative and discriminating features. By carefully evaluating those research studies [14, 52, 56–58], the feature set provided by probabilistic neural network based artificial bee colony (PNN-ABC) optimal feature collection algorithm by Khokhar et al. [51] is exploited in our study. Proposed favorable features are [E (d1), Kurtosis (KT)(d2), RMS (d3), Skewness (SK)(d4), SK (d5), E (d6), RMS (d7), Ent (d8), KT (d8)] which extracted from Daubechies mother wavelet at level 4 and in different levels of decompositions. These features are calculated for all the 3 phases of voltage signal. Calculating a group of attributes for each phase of the voltage, all three separate sets are concatenated linearly and composed a row vector corresponding to a 3 phase voltage signal.

Let us consider a signal $S(t)$ fed into DWT for disturbance detection and classification. To detect disturbance events in the signal, recognizing factor (R_f) [6] can be calculated as:

$$R_f = \sqrt{\frac{\sum_{y=1}^M 2^y \sum_{x=1}^{M_{d_y}} d_y^2(x)}{2^M \sum_{x=1}^{M_{c_M}} c_M^2(x)}} \quad (1)$$

where M , M_{d_y} and d_y are the highest decomposition number, detail coefficient numbers at level y and the coefficients of approximation at y level, respectively; M_{c_M} and c_M are detail coefficient numbers and the coefficients of approximation at M level, respectively.

If $R_f > 1\%$, classification will be done by calculating wavelet coefficients. For $R_f < 1\%$, further calculation will be in halted state and it prevents unnecessary calculation.

The signal $S(t)$ to be resolved into M parts uses DWT as follows:

$$S(t) = S_{c_1}(t) + S_{d_1}(t) + S_{d_2}(t) + \dots + S_{d_M}(t) \quad (2)$$

where $S_{c_1}(t), S_{d_1}(t), S_{d_2}(t), \dots, S_{d_M}(t)$ are the decomposed components of $S(t)$. We can further express $S(t)$ as:

$$S(t) = \sum_i c_1(i) \phi(t-i) + \sum_i \sum_{j=1}^M d_j(i) 2^{j/2} \Psi(2^j t - i) \quad (3)$$

where j is the level of resolution; d_j is the detail coefficient in level j ; $\phi, \Psi \in \mathbf{R}$. According to MRA, a set of nested subspaces V_j and W_j are calculated as follows:

$$\begin{cases} V_M \supset V_{M-1} \supset \dots \supset V_j \supset \dots \supset V_2 \supset V_1 \\ V_{j+1} = V_j \oplus W_j \\ V_j \cap W_j = \emptyset \end{cases} \quad (4)$$

where a summation of two subspaces is marked as \oplus .

The input signal $S(t)$ is resolved into corresponding subsets in accordance with subsets V_1 and W_j , respectively:

$$S_{c_1}(t) = \sum_i c_1(i) \phi(t-i) \quad (5)$$

$$S_{d_j}(t) = \sum_i d_j(i) 2^{j/2} \Psi(2^j t - i) \quad (6)$$

The average value from the signal $S(t)$ detail coefficients at decomposition j level ($|\bar{S}_{d_j}|$) is:

$$|\bar{S}_{d_j}| = \frac{1}{2^j M_{d_j}} \sum_i |d_j(i)| \quad (7)$$

where M_{d_j} is the number of detail coefficients at j level.

Accordingly, the average value of the input signal at individual decomposition levels calculated from detail coefficients and approximation coefficients is as below:

$$\bar{S}(t) = [|\bar{S}_{c_1}(t)|, |\bar{S}_{d_1}(t)|, \dots, |\bar{S}_{d_M}(t)|] \quad (8)$$

The standard deviation (SD) of detail coefficients' absolute values at j level ($\sigma_{S_{dj}}(t)$) is:

$$\sigma_{S_{dj}}(t) = \sqrt{\frac{1}{2jN_{d_j}} \sum_i |d_j(i)| - \bar{S}_{d_j}} \quad (9)$$

The SD of detail coefficients absolute values in individual decomposition level is:

$$\sigma_{S_{dj}}(t) = [\sigma_{S_{c_1}}(t), \sigma_{S_{d_2}}(t), \dots, \sigma_{S_{d_M}}(t)] \quad (10)$$

After sampling at a rate of 20 kHz, the feature vector (wavelet network (WN) input) [6] is:

$$\mathbf{x}_1 = \frac{\sigma_{S_{d_{1,2,3}}}(t)}{|\bar{S}_{d_{1,2,3}}(t)|} \quad (11)$$

where \mathbf{x}_1 is the ratio of the standard deviation measured from input signal $S(t)$ detail coefficients at decomposition 1, 2 and 3 levels ($\sigma_{S_{d_{1,2,3}}}(t)$) to the average value collected from input signal $S(t)$ detail coefficients at identical decomposition levels ($|\bar{S}_{d_{1,2,3}}(t)|$). This determines change of detail coefficients at 1, 2 and 3 levels, from average values without concerning about dimension of those coefficients. Moreover, this technique both diminishes the dimension of WN by normalizing the details data and maintains significant characteristics of the input signals. In the same way, vectors \mathbf{x}_2 , \mathbf{x}_3 and \mathbf{x}_4 are defined as:

$$\mathbf{x}_2 = \frac{\sigma_{S_{d_{4,5,6}}}(t)}{|\bar{S}_{d_{4,5,6}}(t)|} \quad (12)$$

$$\mathbf{x}_3 = \frac{\sigma_{S_{d_{7,8}}}(t)}{|\bar{S}_{d_{7,8}}(t)|} \quad (13)$$

$$\mathbf{x}_4 = \frac{\sigma_{S_{d_{9,10,11,12}}}(t)}{|\bar{S}_{d_{9,10,11,12}}(t)|} \quad (14)$$

The \mathbf{x}_5 , \mathbf{x}_6 , \mathbf{x}_7 and \mathbf{x}_8 inputs are calculated as follows:

$$\begin{cases} \mathbf{x}_5 = \max |S_{d_2}| \\ \mathbf{x}_6 = \max |S_{d_5}| \\ \mathbf{x}_7 = \max |S_{d_8}| \\ \mathbf{x}_8 = \max |S_{d_{12}}| \end{cases} \quad (15)$$

All the vector inputs (\mathbf{x}_1 to \mathbf{x}_8) are extracted from the distorted waveforms. Following this, all the 8 elements constitute the feature vector for the input of WN. \mathbf{x}_1 is greater for transient disturbances than other PQ disturbances. High values of \mathbf{x}_4 and \mathbf{x}_2 are responsible for voltage flicker and harmonic distortions respectively. During the voltage swell and sag, \mathbf{x}_3 and \mathbf{x}_5 are higher than other elements. \mathbf{x}_5 and \mathbf{x}_7 are responsible for the voltage interrupt and notching, respectively. WN can detect DC

offset by evaluating \mathbf{x}_8 . Therefore, different features related to PQ can be extracted by calculating \mathbf{x}_1 to \mathbf{x}_8 .

2.4 OCSVM in a nutshell

SVM is a machine learning algorithm based on modern statistical learning theory [59–62]. It separates two classes by constructing a hyper surface in the input space. In the input space, input is mapped to higher dimensional feature space by non-linear mapping. In this section, we will explain the multi-class SVM followed by OCSVM.

Let us consider, a data space $\Psi = (\mathbf{x}_i, y_i)$, $i = \{1, 2, \dots, n\}$ where $\mathbf{x}_i \in \mathbf{R}^n$ is input data and $y_i \in \{-1, +1\}$ is corresponding output pattern in the dedicating class membership. For simplicity, we define the input and projected data as \mathbf{x} and \mathbf{y} . SVM first projects the input vector \mathbf{x} to a higher dimensional space \mathcal{H} by a non-linear operator $\Phi(\cdot) : \mathbf{R}^n \rightarrow \mathcal{H}$ where the data projection is linearly separable.

The non-linear SVM classification is expressed as (16), where \mathbf{w} is the hyper plane direction; b is the offset scalar:

$$\Omega(\mathbf{x}) = \mathbf{w}^T \Phi(\mathbf{x}) + b \quad \mathbf{w} \in \mathcal{H}, b \in \mathbf{R} \quad (16)$$

which is linear in consideration of projected data $\Phi(\mathbf{x})$ and non-linear in consideration of original data \mathbf{x} .

SVM tries to maximize the margin in hyperplane. Slack variables (ξ_i) are proposed to permit some data to lie within margin in order to protect SVM from over fitting with turbulent data (or introducing soft margin). So the objective function which includes minimization of $\|\mathbf{w}\|$ can be written as:

$$\min_{\mathbf{w}, b, \xi_i} \left(\frac{\|\mathbf{w}\|}{2} + C \sum_{i=1}^n \xi_i \right) \quad (17)$$

subject to:

$$\begin{cases} y_i(\mathbf{w}^T \Phi(\mathbf{x}_i) + b) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases} \quad (18)$$

where C is the regularization parameter (usually greater than 0) that regulates the trade-off of enlarging the margin and number of trained data within that margin (thus reducing the training errors); ξ_i ($i = 0, 1, \dots, n$) is the slack variable; n is the number of input data.

To minimize the objective function of (17) using Lagrange multipliers technique, the necessary condition for \mathbf{w} is:

$$\mathbf{w} = \sum_{i=1}^n \gamma_i y_i \Phi(\mathbf{x}_i) \quad (19)$$

where $\gamma_i > 0$ is Lagrange multiplier corresponding to the constraints in (18). γ_i can be solved from (17) and written as:



$$\max \mathbf{w}(\gamma_i) = \sum_{i=1}^n \gamma_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \gamma_i \gamma_j y_i y_j k(\mathbf{x}_i, \mathbf{x}_j) \quad (20)$$

subject to:

$$\begin{cases} 0 \leq \gamma_i \leq C \\ \sum_{i=1}^n \gamma_i y_i = 0 \end{cases} \quad (21)$$

This $k(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x})^T \Phi(\mathbf{y})$ is known as kernel function. It determines the mapping of input vector to high dimensional feature space.

Gaussian RBF kernel for multi-class SVM is:

$$k(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{\sigma}\right) \quad (22)$$

where $\sigma \in \mathbf{R}$ is the width of RBF function.

OCSVM, which is a variation of the SVM, detects the abnormal data within a class [61, 62]. OCSVM maps the input vector to feature dimension according to the kernel function and separates them from origin keeping high margin. It penalizes the outliers by employing slack variables ξ in the objective function and controls carefully the trade off between empirical risk and regularization of penalty.

The quadratic programming minimization function is:

$$\min_{\mathbf{w}, \xi_i, \rho} \left(\frac{1}{2} \|\mathbf{w}\|^2 + \frac{1}{vn} \sum_{i=1}^n \xi_i - \rho \right) \quad (23)$$

subject to:

$$\begin{cases} \mathbf{w} \cdot \Phi(\mathbf{x}_i) \geq \rho - \xi_i \\ \xi_i \geq 0 \end{cases} \quad (24)$$

where $v \in (0, 1]$ is a prior fixed constant; ρ is the resolved value that indicates whether a stated point falls within the considered high density area.

Then the resultant decision function $f_{w,p}^m(\mathbf{x})$ takes the form:

$$f_{w,p}^m(\mathbf{x}) = \text{sgn}((\mathbf{w}^*)^T \Phi(\mathbf{x}) - \rho^*) \quad (25)$$

where ρ^* and \mathbf{w}^* are values of \mathbf{w} and ρ solving from (23).

In OCSVM, v characterizes the solution instead of C (smoothness operation) in that:

- 1) It determines a top boundary limit on the fraction of outliers.
- 2) It finds a lower boundary on the number of trained instances considered as support vector.

Due to high significance of v , OCSVM is also termed as v -SVM.

3 System model

Considering the fact that abnormality patterns in controlled experiment are different from the reality and that the real abnormal patterns are not easily accessible for training purpose, a functional model should be trained on the available data with less assumption about the types of abnormalities. As it is required to run the model in a smart meter in real time, it should take the least possible computation time to learn the pattern of normal and abnormal signals from a small set of samples.

To create the mentioned model, signal samples are simulated with the Simulink toolbox in MATLAB with the help of different circuits [63], mimicking the real condition in the distribution network. The sampling time for all the samples is set to 5 μs , and 500 samples of normal and 500 samples of abnormal waveforms, equally spread in different five categories with 0.2 s length are generated. However, the entire sample set is not used for training and testing in each experiment. The input voltage and current to the circuits randomly change between 0.05 higher or lower than the defined standard in all the circuits resembling the real power change in distribution circuits. We attempt to keep a standard setting for all the parameters in all the circuits to avoid the effects of confounding variables. The DWT is exploited to extract the most informative features out of each signal. Having 3 phase signal, we process each phase separately extracting the following features: RMS, Ent, E, average value (I), KT, standard deviation (δ_r), SK, range (RG), detail value (D), and approximation (A) coefficients. All these are being measured using MATLAB built in wavelet decomposition function. The process flow is provided in Fig. 2.

The processed features are fed into OCSVM for disturbance detection. If disturbance is appeared, the data is further processed by multi-class SVM for getting details disturbance classifications. The process is described in Fig. 3.

4 Simulation result

Having the same set of features for each signal phase, they all are concatenated linearly to compose a sample set of feature vectors. Creating the sample space, OCSVM is

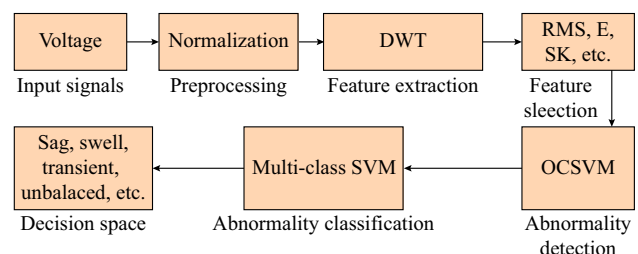


Fig. 2 Process flow of disturbance detection by SVM

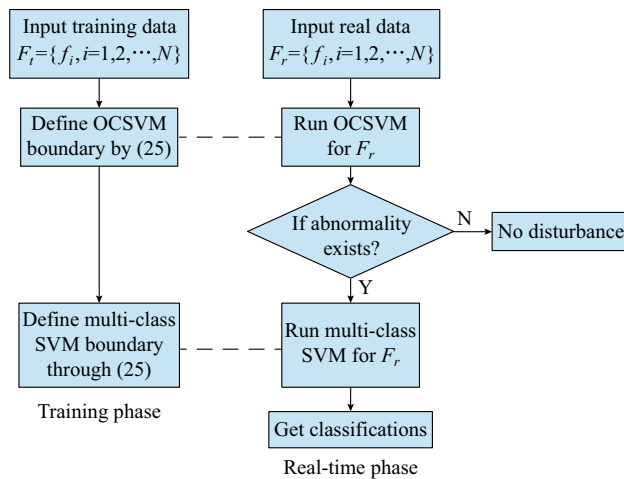


Fig. 3 Disturbance detection and classification

trained exclusively with the 300 samples. Another 200 samples are kept for testing phase. To define the hyper parameters of ν and γ , grid search is performed and the results are shown in Fig. 4. The figure represents the confusion matrix for different parameters setting, representing the numbers of true positive (TP), false positive (FP), true negative (TN) and false negative (FN) by adjusting different values for ν and γ . Considering the fact that we aim at detecting all the abnormalities with the cost of some false alarms, we picked $\nu = 0.01$ and $\gamma = 0.1$ with the highest number of TN equal to 199 and TP equal to 188, which means that among all 200 abnormal samples and 200 normal samples, OCSVM with (Gaussian) radial basis function kernel is able to detect 188 normals and 199 abnormal samples correctly with 13 false alarms. In this case, the average accuracy becomes $(188/200 + 199/200)/2 \times 100\% \approx 97\%$.

Principle component analysis (PCA) is applied on the dataset to reduce the dimensionality and provide results representable in two dimensional (2D) space. The OCSVM boundaries with RBF kernel and two different ν and γ on a sample set of data are illustrated in Fig. 5. The training samples are green spots in the center of the figure surrounded by decision boundaries. The blue clouds show nearest vectors to the defined plane. The closer it gets, it becomes darker and it shows the higher risk area. The pink

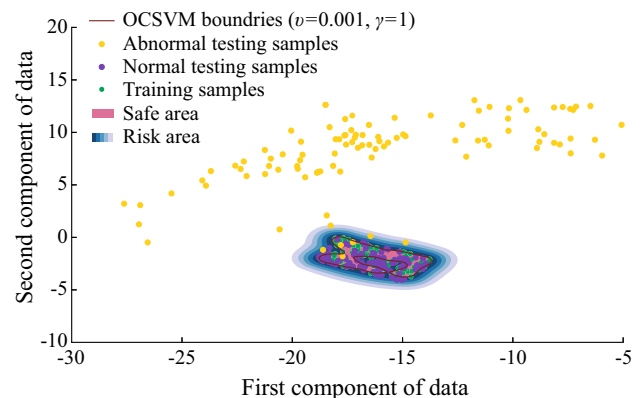
	$\gamma=0.0001$	$\gamma=0.001$	$\gamma=0.01$	$\gamma=0.1$	$\gamma=1$	
$\nu=0.0001$	0 0 200 200	0 0 200 200	195 0 21 184	194 0 18 188	74 194 129 197	Interpretation TN: Actual normal, detected normal FP: Actual normal, detected abnormal FN: Actual abnormal, detected normal TP: Actual abnormal, detected abnormal
$\nu=0.001$	0 0 200 200	198 0 23 179	198 0 21 181	198 0 14 188	135 0 68 197	
$\nu=0.01$	195 0 27 178	197 0 25 178	197 0 21 182	199 0 13 188	172 0 31 197	
$\nu=0.1$	176 0 39 185	178 0 37 185	175 0 38 187	171 0 40 189	166 0 37 197	
$\nu=1$	0 0 200 200	0 0 200 200	0 0 200 200	0 0 200 200	0 0 200 200	

Fig. 4 Grid search result of hyper parameters

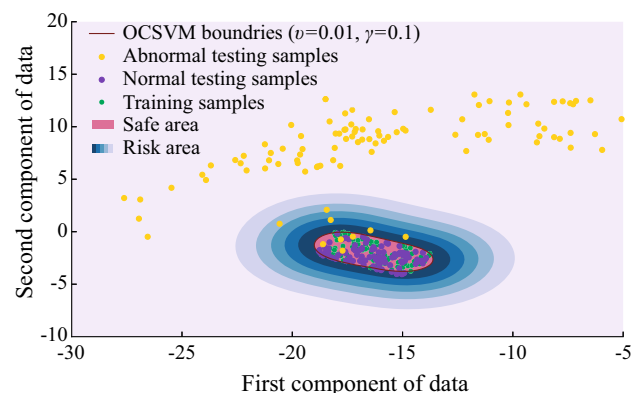
area surrounded by red boundary is the safe area, and all the samples which fall in this area is flagged as normal. For $\nu = 0.001$ and $\gamma = 1$, the accuracy is 80% (160/200). On the other hand, accuracy as high as 93% (187/200) is found for $\nu = 0.01$ and $\gamma = 0.1$. While representation of the entire samples in the 2D space reduces the interpret-ability of the figure, a subset of the sample data is selected and represented here to clarify the effects of hyper parameter in decision boundaries and classification performance.

Receiver operating characteristic curve (ROC) which is used for evaluating classifier output quality, is shown in Fig. 6. TP and FP rates are represented on the y-axis and x-axis, respectively. The red dashed line indicates that if we do not use any binary classification algorithm and just randomly label the samples, we will tag them correctly by 50% chance. Therefore, the ROC will be 50%. The more we cover the area, the better and more accurate the classifier is. The blue line shows the diagnostic ability of the proposed classifier, while discrimination threshold is varied. In the figure the top left corner of the graph is the ideal point where a zero FP rate, and a one TP rate are found.

Following the provided process schematic (Fig. 2), after abnormality detection, the type of the abnormality is



(a) FP rate is 37/200, FN rate is 3/200



(b) FP rate is 9/200, FN rate is 4/200

Fig. 5 Effects of hyper parameter adjustment on decision boundary and algorithm performance



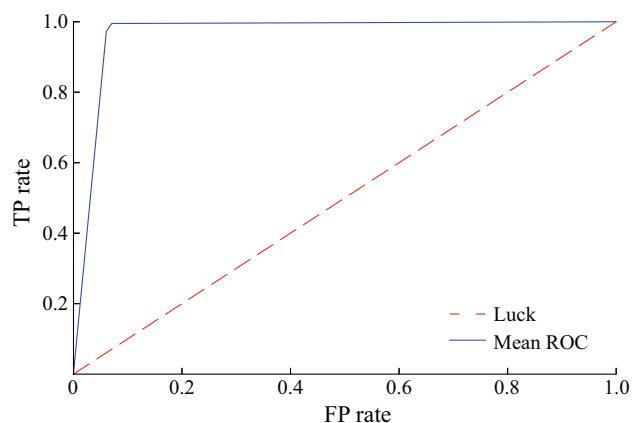


Fig. 6 ROC of OCSVM

defined in the next step by a multi-class SVM classification algorithm. This two-step approach increases the robustness of the model specially in the detection phase when an unknown disturbance appears.

To the best of our knowledge, the majority of the research in the field has simulated their own dataset with an arbitrary assumption about the size of the sample data and the disturbance pattern complexity as well as the simulation parameter settings. Considering the fact that the accuracy and F-measure can vary dramatically based on the complexity of the underlying input dataset and knowing that there is no access either to earlier research dataset or to any other public dataset in the scope [64], there is no actual baseline available to compare the current result with it. Furthermore, considering the fact that this kind of dataset is unbalanced (abnormal samples are not as abundant as normal samples), classification accuracy alone is not informative enough. We use Precision, Recall, Confusion Matrix and F-measure metrics to have a better understanding of how the method performs [65]. Precision is the ratio of TP number to TP and FP numbers, and Recall is the ratio of TP number to TP and FN numbers.

After applying anomaly detection by OCSVM, types of the anomaly should be detected by a multi-class classification algorithms. Considering the fact that in the training time the abnormal samples are scarce, to achieve a realistic result multiple algorithms are trained on a relatively small training dataset in which the number of abnormality samples are at most 10 in each class. The accuracy of the algorithms in an unbalanced testing dataset are shown in the Table 1. It is noted that the F-measure (F1-score or F-score) is a measure of an algorithm's accuracy and is defined as the weighted harmonic mean of the precision and recall of the test. F-measure can be calculated in multiple way and one of them is F1-macro, as shown in Table 1. The outcomes obviate the superiority of the multi-class SVM and random forest algorithms for this multi-class classification task.

Table 1 Classification result with a small training set

Algorithm	Accuracy	F1-macro
Nearest neighbors	0.66	0.60
Linear SVM	0.68	0.67
Gaussian	0.61	0.61
Decision tree	0.69	0.59
Random forest	0.73	0.67
Neural network	0.25	0.17

Table 2 Classification result with a large training set

Algorithm	Accuracy	F1-macro
Nearest neighbors	0.86	0.79
Linear SVM	0.89	0.84
Gaussian	0.65	0.34
Decision tree	0.90	0.88
Random forest	0.90	0.82
Neural network	0.84	0.74
Onevsall_svm	0.90	0.86

To attest the effect of the training size on the accuracy and F-measure, the same experiment is repeated on a larger dataset composed of 20 abnormal samples in each classes. The result shown in Table 2 clarifies the direct relation of the training size and classification accuracy. This comparison is done to demonstrate that different algorithms and techniques can not be compared unless they are applied in the same training and testing set.

In Table 2, the algorithm onevsall_svm or one-vs-rest is a classification strategy, in which a single classifier per class is trained, with the samples of that class as positive samples and all other samples as negatives. Repeating one-vs-all strategy on a multi-class data can discriminate the data into more than two classes. This technique is reducing the problem of multi-class classification to multiple binary classification problems.

5 Conclusion

PQ reports by consumer meter is important addition to smart grid as seen by utility company. Following this, in this study, we propose machine learning based disturbance detection. For segregation between regular data and abnormal data, we propose one-class version of SVM. Further, for categorization of disturbances, we propose multi-class SVM. OCSVM detects disturbances with 93% accuracy. On the other hand, multi-class SVM can classify the detected disturbances with accuracy as high as 90% depending on the training data set. The outcome of SVM

will be reported to utility company back office. This will help to get insight into the PQ issue at lowest distribution level and maintain good PQ.

Acknowledgements This research was supported in part through U.S. National Science Foundation (No. 1553494). Imtiaz PARVEZ and Maryamossadat AGHILI contributed equally to this work.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- [1] Ekanayake J, Jenkins N, Liyanage K et al (2012) Smart grid: technology and applications. Wiley, New Jersey
- [2] Qiu RC, Hu Z, Chen Z et al (2011) Cognitive radio network for the smart grid: experimental system architecture, control algorithms, security, and microgrid testbed. *IEEE Trans Smart Grid* 2(4):724–740
- [3] Parvez I, Khan, T, Sarwat A et al (2017) LAA-LTE and WiFi based smart grid metering infrastructure in 3.5 GHz band. In: *Proceedings of IEEE region 10 humanitarian technology conference (R10-HTC)*, Dhaka, Bangladesh, 21–23 December 2017, pp 151–155
- [4] Parvez I, Jamei M, Sundararajan A et al (2014) RSS based loop-free compass routing protocol for data communication in advanced metering infrastructure (AMI) of smart grid. In: *Proceedings of 2014 IEEE symposium on computational intelligence applications in smart grid (CIASG)*, Orlando, USA, 9–12 December 2014, 6 pp
- [5] Hao X, Wang Y, Wu C et al (2012) Smart meter deployment optimization for efficient electrical appliance state monitoring. In: *Proceedings of 2012 IEEE third international conference on smart grid communications*, Tainan, China, 5–8 November 2012, pp 25–30
- [6] Masoum MAS, Jamali S, Ghaffarzadeh N (2010) Detection and classification of power quality disturbances using discrete wavelet transform and wavelet networks. *IET Sci Meas Technol* 4(4):193–205
- [7] Parvez I, Sarwat A, Wei L et al (2016) Securing metering infrastructure of smart grid: a machine learning and localization based key management approach. *Energies* 9(9):1–18
- [8] Parvez I, Islam A, Kaleem F (2014) A key management-based two-level encryption method for AMI. In: *Proceedings of IEEE power and energy society general meeting*, National Harbor, USA, 27–31 July 2014, 5 pp
- [9] Sarwat A, Sundararajan A, Parvez I (2017) Trends and future directions of research for smart grid IoT sensor networks. In: *Proceedings of international symposium on sensor networks, systems and security*, Lakeland, USA, 31 August–2 September 2017, pp 45–61
- [10] Fang X, Misra S, Xue G et al (2012) Smart grid-the new and improved power grid: a survey. *IEEE Commun Surveys Tutorials* 14(4):944–980
- [11] Morsi WG, El-Hawary ME (2011) Power quality evaluation in smart grids considering modern distortion in electric power systems. *Electr Power Syst Res* 81(5):1117–1123
- [12] Sarwat A, Sundararajan A, Parvez I et al (2018) Toward a smart city of interdependent critical infrastructure networks. In: Amini M (eds) *Sustainable interdependent networks. Studies in systems, decision and control*, vol 145. Springer, Cham, pp 21–45
- [13] Fuchs EFH, Fuchs H, Masoum M (2008) Power quality of electric machines and power systems. In: *Proceedings of the 8th IASTED international conference*, Corfu, Greece, 23–25 June 2008, pp 35–40
- [14] Santoso S, Powers EJ, Grady WM et al (1996) Power quality assessment via wavelet transform analysis. *IEEE Trans Power Deliv* 11(2):924–930
- [15] Singh B, Al-Haddad K, Chandra A (1999) A review of active filters for power quality improvement. *IEEE Trans Ind Electron* 46(5):960–971
- [16] Daponte P, Di Penta M, Mercurio G (2004) TransientMeter: a distributed measurement system for power quality monitoring. *IEEE Trans Power Deliv* 19(2):456–463
- [17] Mallet Y, Coomans D, Kautsky J et al (1997) Classification using adaptive wavelets for feature extraction. *IEEE Trans Pattern Anal Mach Intell* 19(10):1058–1066
- [18] Learned RE, Willsky AS (1995) A wavelet packet approach to transient signal classification. *Appl Comput Harmonic Anal* 2(3):265–278
- [19] Granados-Lieberman D, Romero-Troncoso RJ, Osornio-Rios RA et al (2011) Techniques and methodologies for power quality analysis and disturbances classification in power systems: a review. *IET Gener Transm Distrib* 5(4):519–529
- [20] Styvaktakis E (2002) Automating power quality analysis. Dissertation, Chalmers University of Technology
- [21] Moravej Z, Banihashemi SA, Velayati MH (2009) Power quality events classification and recognition using a novel support vector algorithm. *Energy Conv Manag* 50(12):3071–3077
- [22] Panigrahi BK, Pandi VR (2009) Optimal feature selection for classification of power quality disturbances using wavelet packet-based fuzzy k-nearest neighbour algorithm. *IET Gener Transm Distrib* 3(3):296–306
- [23] Kezunovic M, Liao Y (2001) A new method for classification and characterization of voltage sags. *Electr Power Syst Res* 58(1):27–35
- [24] Monedero I, Leon C, Ropero J et al (2007) Classification of electrical disturbances in real time using neural networks. *IEEE Trans Power Deliv* 22(3):1288–1296
- [25] Axelberg PG, Gu YH, Bollen MH (2007) Support vector machine for classification of voltage disturbances. *IEEE Trans Power Deliv* 22(3):1297–1303
- [26] Janik P, Lobos T (2006) Automated classification of power-quality disturbances using SVM and RBF networks. *IEEE Trans Power Deliv* 21(3):1663–1669
- [27] Borges FA, Fernandes RA, Silva IN et al (2016) Feature extraction and power quality disturbances classification using smart meters signals. *IEEE Trans Ind Inform* 12(2):824–833
- [28] Abdel-Galil TK, Kamel M, Youssef AM et al (2004) Power quality disturbance classification using the inductive inference approach. *IEEE Trans Power Deliv* 19(4):1812–1818
- [29] Kezunovic M, Liao Y (2002) A novel software implementation concept for power quality study. *IEEE Trans Power Deliv* 17(2):544–549
- [30] Dash PK, Mishra S, Salama MA et al (2000) Classification of power system disturbances using a fuzzy expert system and a Fourier linear combiner. *IEEE Trans Power Deliv* 15(2):472–477
- [31] Styvaktakis E, Bollen MH, Gu IY (2002) Expert system for classification and analysis of power system events. *IEEE Trans Power Deliv* 17(2):423–428



- [32] Chung J, Powers EJ, Grady WM et al (2002) Power disturbance classifier using a rule-based method and wavelet packet-based hidden Markov model. *IEEE Tran Power Deliv* 17(1):233–241
- [33] Allen JB, Rabiner LR (1977) A unified approach to short-time Fourier analysis and synthesis. *Proc IEEE* 65(11):1558–1564
- [34] Dokur Z, Olmez T, Yazgan E (1999) Comparison of discrete wavelet and Fourier transforms for ECG beat classification. *Electron Lett* 35(18):1502–1504
- [35] Burrus CS, Gopinath RA, Guo H et al (1998) Introduction to wavelets and wavelet transforms: a primer, vol 1. Prentice Hall, New Jersey
- [36] Rioul O, Vetterli M (1991) Wavelets and signal processing. *IEEE Signal Process Mag* 8(4):14–38
- [37] Wang C, Gao H, Zhu T (2006) A new method for detection and identification of power quality disturbance. In: Proceedings of 2006 IEEE power and energy society power systems conference and exposition, Atlanta, USA, 29 October–1 November 2006, pp 1556–1561
- [38] Gaouda AM, Kanoun SH, Salama MMA (2001) On-line disturbance classification using nearest neighbor rule. *Electr Power Syst Res* 57(1):1–8
- [39] Zhang M, Li K, Hu Y (2011) A real-time classification method of power quality disturbances. *Electr Power Syst Res* 81(2):660–666
- [40] McEachern A (1988) Handbook of power signatures. Basic Measuring Instruments, Foster City
- [41] Ghosh AK, Lubkeman DL (1995) The classification of power system disturbance waveforms using a neural network approach. *IEEE Trans Power Deliv* 10(1):109–115
- [42] Santoso S, Lamoree J, Grady WM et al (2000) A scalable PQ event identification system. *IEEE Trans Power Deliv* 15(2):738–743
- [43] Bollen MH, Gu IY (2006) Signal processing of power quality disturbances, vol 30. Wiley, New Jersey
- [44] Hu GS, Zhu FF, Ren Z (2008) Power quality disturbance identification using wavelet packet energy entropy and weighted support vector machines. *Expert Syst Appl* 35(1–2):143–149
- [45] Eriti H, Uar A, Demir Y (2010) Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines. *Electr Power Syst Res* 80(7):743–752
- [46] Guo Y, De Jong K, Liu F et al (2012) A comparison of artificial neural networks and support vector machines on land cover classification. In: Proceedings of 6th international symposium on intelligence computation and applications, Wuhan, China, 27–28 October 2012, pp 531–539
- [47] Uyar M, Yildirim S, Gencoglu MT (2008) An effective wavelet-based feature extraction method for classification of power quality disturbance signals. *Electr Power Syst Res* 78(10):1747–1755
- [48] Koleva L, Taskovski D, Milchevski A et al (2012) Application of near perfect reconstruction filter banks in power quality disturbances classification methods. In: Proceedings of 2012 IEEE international workshop on applied measurements for power systems (AMPS), Aachen, Germany, 26–28 September 2012, pp 1–5
- [49] Kostadinov D, Taskovski D (2012) Automatic voltage disturbance detection and classification using wavelets and multiclass logistic regression. In: Proceedings of 2012 IEEE international instrumentation and measurement technology conference (I2MTC), Graz, Austria, 13–16 May 2012, pp 103–106
- [50] Sahani M, Mishra S, Ipsita A et al (2016) Detection and classification of power quality event using wavelet transform and weighted extreme learning machine. In: Proceedings of 2016 international conference on circuit, power and computing technologies (ICCPCT), Nagercoil, India, 18–19 March 2016, pp 1–6
- [51] Khokhar S, Zin AAM, Memon AP et al (2017) A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network. *Measurement* 95:246–259
- [52] Naik C, Hafiz F, Swain A et al (2016) Classification of power quality events using wavelet packet transform and extreme learning machine. In: Proceedings of 2016 IEEE 2nd annual southern power electronics conference (SPEC), Auckland, New Zealand, 5–8 December 2016, pp 1–6
- [53] Akansu AN, Haddad PA, Haddad RA (2001) Multiresolution signal decomposition: transforms, subbands, and wavelets. Academic Press, Cambridge
- [54] Bermingham ML, Pong-Wong R, Spiliopoulou A et al (2015) Application of high-dimensional feature selection: evaluation for genomic prediction in man. *Sci Rep* 5(10312):1–12
- [55] Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. *J Mach Learn Res* 3(6):1157–1182
- [56] Abdelgayed TS, Morsi WG, Sidhu TS (2018) A new approach for fault classification in microgrids using optimal wavelet functions matching pursuit. *IEEE Trans Smart Grid* 9(5):4838–4846
- [57] Mishra DP, Samantaray SR, Joos G (2016) A combined wavelet and data-mining based intelligent protection scheme for microgrid. *IEEE Trans Smart Grid* 7(5):2295–2304
- [58] Nath S, Dey A, Chakrabarti A (2009) Detection of power quality disturbances using wavelet transform. *World Acad Sci Eng Technol* 49:869–873
- [59] Sriyananda MGS, Parvez I, Gven I et al (2016) Multi-armed bandit for LTE-U and WiFi coexistence in unlicensed bands. In: Proceedings of 2016 IEEE wireless communications and networking conference (WCNC), Doha, Qatar, 3–6 April 2016, pp 1–6
- [60] Parvez I, Sriyananda MGS, Gven I et al (2016) CBRS spectrum sharing between LTE-U and WiFi: a multiarmed bandit approach. *Mobile Inf Syst* 2016:1–12
- [61] Vapnik V (1998) Statistical learning theory, vol 3. Wiley, New York
- [62] Schlkopf B (2001) Statistical learning and kernel methods. In: Della Riccia G (eds) Data fusion and perception. International center for mechanical sciences (courses and lectures), vol 431. Springer, Vienna, pp 3–24
- [63] Dhote PV, Deshmukh BT, Kushare BE (2015) Generation of power quality disturbances using MATLAB-Simulink. In: Proceedings of 2015 international conference on computation of power, energy, information and communication (ICCPEIC), Chennai, India, 22–23 April 2015, pp 301–305
- [64] Medjroubi W, Miller UP, Scharf M et al (2017) Open data in power grid modelling: new approaches towards transparent grid models. *Energy Rep* 3:14–21
- [65] Pedregosa F, Varoquaux G, Gramfort A et al (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* 12(10):2825–2830

Imtiaz PARVEZ received the B.Sc. degree in electrical and electronics engineering from Bangladesh University of Engineering and Technology, and the M.Sc. degree from University of North Texas. Currently, he is a Ph.D. candidate at Department of Electrical and Computer Engineering in Florida International University (FIU). He served as specialist engineer in Robi Axiata Ltd. in Bangladesh. He also worked for T mobile (through 3S network Inc.) in 2013. His research interests include 5G, LTE, wifi, cyber physical systems security and artificial intelligence with application to smart grid and IoT.

Maryamossadat AGHILI is a Ph.D. candidate at the School of Computing and Information Sciences (SCIS), FIU, Miami, USA. She received her B.Sc. degree in computer engineering from Amirkabir

University of Technology (AUT), Iran in 2007 and earned her M.Sc. in industrial engineering afterward in 2010. She has over 8 years of research and development experiences in computer science, ERP, MIS, and related fields. Her research interests include machine learning, data mining, image and signal processing, computer vision, anomaly detection and deep learning. She has published several research papers and currently works as a research assistant in computer vision laboratory at Florida International University under supervision of Dr. Malek Adjouadi.

Arif I. SARWAT received his Masters' degree from University of Florida in 2005 and Ph.D. degree from the University of South Florida in 2010. Currently, he is an Associate Professor and Director of FPL-FIU Solar Research Facility in the Department of Electrical and Computer Engineering at FIU, where he leads the Energy Power & Sustainability group. His research interests include smart grids, PlugIn Hybrid and Electric Vehicle (PHEV & EV Systems), high penetration renewable systems, grid resiliency, large-scale data analysis, advanced metering infrastructure, smart city infrastructure and cyber security. Dr. Sarwat is the recipient of the NSF CAREER award in 2016, and recipient of multiple federal and industry research awards. He was the author/co-author of conference best paper awards at the resilience week in 2017, and winner of best paper award in 2016 from Journal of Modern Power Systems and Clean Energy. Dr. Sarwat received Faculty Award for Excellence in Research & Creative Activities in 2016, College of Engineering & Computing Worlds Ahead Performance in 2016, and FIU TOP Scholar Award in 2015. Dr. Sarwat worked at Siemens for more than nine years,

winning three recognition awards. He is chair of IEEE Miami Section Vehicular Technology (VT) and Communication since 2012.

Shahinur RAHMAN is a Graduate Research Assistant, pursuing her Ph.D. in electrical and computer engineering at FIU. She received her B.Sc. and M.Sc. degrees in electrical engineering from Bangladesh University of Engineering & Technology (BUET) in 2014 and 2017 respectively. Her research area includes power quality and protection issues of photovoltaic integration into the smart grid at distribution level.

Fahmida ALAM is pursuing Ph.D. while working as a Graduate Research Assistant in Integrated Nanosystems Research Laboratory in the Department of Electrical and Computer Engineering of FIU, Miami, Florida, USA. She received her B.S. degree in engineering from Chittagong University of Engineering and Technology (CUET). She worked as a lecturer in the Department of Electrical and Electronic Engineering of University of Information Technology and Science from August 2010 to December 2014. Her research interest includes the design and micro-fabrication of point of care biosensors, for in-vitro transdermal investigation of bio-analytes using electrochemical sensing, development of miniaturized electrochemical sensing device in wearable platform for specific detection of bio-analytes in the human body.

