An Algorithm for Power Quality Events Core Vector Machine Based Classification

H. Adaminejad[1], I.shayegani[2], M. ohammadi[3], E. Farjah[4]

Received :2015/10/20 Accepted: 2015/12/5

Abstract

Since it is essential to deliver smoothed sinusoidal voltage to the customers, diagnosing power quality (PQ) events has played important role in power delivery and conversion. This diagnostic scheme should be accurate to classify PQ events from other events in power system. Also it should be fast enough to rapidly mitigate PQ events. In this paper, an algorithm based on Core Vector Machine (CVM) has been introduced to classify power quality events. Feature selection method has also been applied to increase the accuracy of the classifier's algorithm. Some features have been selected among several others extracted by wavelet transform. In addition, eight different classes are simulated due to the corresponding equations used in previous studies. Evaluating the performance of the algorithm, different indices have been used to assess the operation of the classifier's algorithm. Simulations results show the robust capability of the proposed algorithm to classify the PQ events.

Keywords: Classifier Algorithm, Core Vector Machine Algorithm, Power Quality Events, Power Quality Classification.

Nomenclature

α_i *Lagrangian multipliers*

b *Bias vector*

f *Decision function*

R *Radius of bias vector*

C *Center of bias vector*

T_{K_i} *Number of cases in class* K_i *and classified correctly by the trained CVM*

F_{K_i} *Number of cases of other classes classified in class* K_i

$F_{K_{oi}}$ *Number of cases of other classes classified in class* K_i *by the trained CVM*

K *Number of the classes*

ω *Weight coefficients vector*

T *Denotes transpose operator*

ϕ *Mapping function*

j *Frequency band level*

K_i *Class index*

1. Introduction

Three-phase voltages in an ideal power system have dispatched amplitude with a constant frequency and distinct phase sequence, whereas in real power system these features may vary due to a variety of reasons, e.g. unbalanced loads, distant faults, inverter loads, lightning and switching. Power quality events occur when any contingencies influence particular features of voltages in power system. Nowadays the demand of a pure sinusoid voltage is alleviating due to incremental sensitive loads. The operation of these loads may be interrupted by the changes in the quality of voltage of the power system, so it is important to identify these events and determine a rapid and accurate solution for each case to eliminate the hazardous effects of these events.

Many researches have been focused on PQ classification and they have offered different methods and algorithms to classify PQ events. An expert system based on S-Transform and the Neural Network (NN) algorithm has been used for automatic classification [1]. The formulation of nine different classes of PQ events is simulated and classified by Artificial Neural Network (ANN). Nevertheless, in ANN the learning time is relatively lengthy. Thus, ANN is not applicable for some applications in which higher speed is important. Wavelet transform can also be used for the feature extraction stage [2]. Energy of the signal has been calculated as an

1. School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran .
2. School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran.
3. School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran, m.mohammadi@shirazu.ac.ir.
4. School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran.

[1] Corresponding author: Department of Power and Control Engineering, School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran, Tel: 0098-7136133278, e-mail address: m.mohammadi@shirazu.ac.ir

input feature of the classifier and in order to classify events Support Vector Machine (SVM) has been conducted. SVM is a fast algorithm with high accuracy; however in applications with huge data set, the training process may be time-consuming. Wavelet transform can also be applied to the sample data, and for events classification the Decision Tree (DT) algorithm is used [3]. Although DT is a functional classifier with an acceptable speed, it has a drawback for some special applications, such as on-line PQ events classification is very time-consuming. DT can also be applied to classify events using the signals energies and special rules have been defined for DT algorithm [4]. This algorithm is very difficult and slow. Moreover, all the rules need to be changed when a new class is classified. Signal segmentation has also been implemented and signal variation has been considered as the input feature of the classifier [5]. The extracted features have been considered as the Root Mean Squared (RMS) value of first, second, third, fifth and ninth harmonic of voltage, the amount of Total Harmonic Distortion (THD), the length of each event and the symmetrical components and SVM algorithm has been deployed for the classification purpose. K-fold algorithm has been applied to divide data into train and test input. Applying different benchmark to the classifier algorithm, the veracity of the algorithm would be calculated for every input data.

Training of SVM is very slow which is not satisfactory when facing higher number of classes and the algorithms mentioned in [3-5] fail to respond perfectly. Hence it is advantageous to deploy other algorithm that can overcome the drawback of time-consuming for higher number of classes.

CVM algorithm is much faster than SVM and other algorithms and it shows accurate and enhanced results for a higher number of input data with less elapsed computation time. Therefore, in this paper for PQ events classification the CVM algorithm has been used to overcome the drawbacks of SVM algorithm for PQ events classification.

The rest of the paper has been organized as follows:

Section 2 describes the concept of SVM theory. In section 3 concept of CVM has been presented and compared with SVM. The proposed algorithm for PQ events classification has been introduced in section 4 and Feature Extraction (FE) and Feature Selection (FS) have also been explained in this section. Section 5 shows the simulation results and several indices are conducted to evaluate the effectiveness of the proposed method. Eventually section 6 summarizes a conclusion based on the simulation results and performance and preponderance of the proposed method.

2. Support Vector Machine Theory

Some journals use a new theory called Vapnik Chervonenkis (VC) which is based on statistical learning method and can be used for data without any particular density function [6]. Also SVM algorithm is fundamentally based on the VC theory. It is applicable to use SVM algorithm for regression, classification purposes, data mining and image processing [7, 8].

2.1. Support Vector Machine (SVM)

Support Vector Machines have been proposed to solve classification, date mining and other machine learning problems with generalization performance and ease of implementation in several application fields. SVM training is formulated as a quadratic programming (QP) problem. This QP problem can be optimized by numerical solvers. However, the complexity of a QP problem is about $O(m^3)$ in required time and $O(m^2)$ in required space, where m is the number of training samples [9]. Therefore, SVM is computationally inaccurate for large data sets.

2.2. Algorithm Formulation

Suppose a two class data set $S = (x_1, y_1), \dots, (x_n, y_n)$, where $x \; and \; y = \pm 1$ denotes the input vector and the label of the classes, respectively.

In binary linear classification, discrimination has been calculated by a line formulated as:

$$\omega^T x + b = 0 \qquad (1)$$

In fact ω and b determine the position of the line [10]. Figure. 1 shows a linear binary classification.

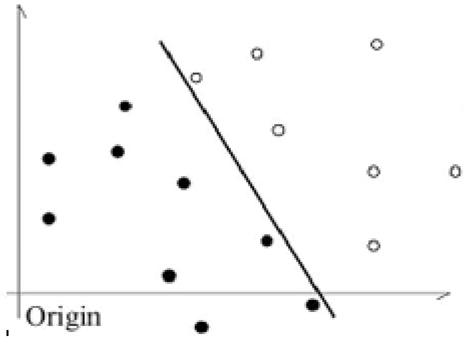

Figure. 1. Binary classification using SVM

The equations used for classification of each class can be formed as:

For first class:

$$\omega^T x + b \geq 0 \qquad for \quad y = 1 \qquad (2)$$

and

$$\omega^T x + b \leq 0 \qquad for \quad y = -1 \qquad (3)$$

Thus, for the correct classified data, it has been taken into account that:

$$y(\omega^T x + b \geq 0) \geq 0 \qquad (4)$$

On the other hand, the equation of the closest points of the first class can be shown:

$$\omega^T x + b = 1 \qquad (5)$$

For the other class is:

$$\omega^T x + b = -1 \qquad (6)$$

The margin of the two classes has been calculated by $x_1 - x_2$ where x_1 and x_2 are derived from (5) and (6). Figure 2 shows the margin between two classes of a binary classification.

It can be shown that the margin would be $\frac{1}{||\omega||}$.

To maximize the margin, the following equation can be solved:

$$\min(\omega) = \frac{1}{2}(||\omega||)^2 \qquad (7)$$

In order to solve an optimization problem of second order, the Lagrange equation in terms of (4) has been defined as:

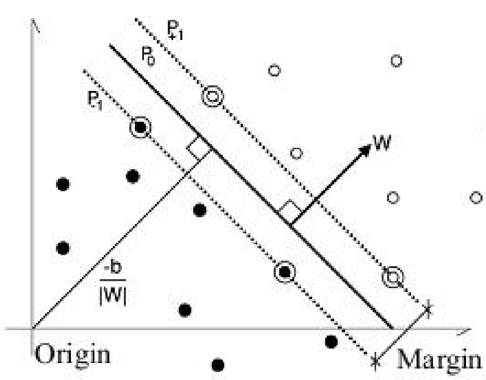

Figure. 2. Margin between the two classes

$$L = \frac{1}{2}(\omega.\omega) - \sum_{i}^{m} \alpha_i(y_i((\omega.x + b) - 1)) \quad (8)$$

Which$\alpha_i = 0$. It can be shown that the dual Lagrangian of (8) is [11-13]:

$$W(\alpha) = -\frac{1}{2}\sum_{i}^{m} \alpha_i \alpha_j y_i y_j (x_i, x_j) \qquad (9)$$

The above equation must be maximized where $\sum_i^n \alpha_i y_i = 0$ and $\alpha_i > 0 \ \forall i$. Using mapped input data accurately in feature space, $x_i.x_j$ can be replaced with$\phi(x_i)\phi(x_j)$. The mapping function can be replaced by a Kernel function$\kappa(x_i, x_j)$. Several kernel functions such as Polynomial, Radial basis function and sigmoidal are applicable for mapping function. Hence (9) can be reformulated as:

$$W(\alpha) = \sum_{i}^{m} \alpha_i - \frac{1}{2}\sum_{i,j}^{m} \alpha_i \alpha_j y_i y_j \kappa(x_i, x_j) \qquad (10)$$

where $\sum_i^n \alpha_i y_i = 0$ and $\alpha_i > 0 \ \ \forall i$. Determining the value of α_i, the decision function is calculated by:

$$f(x) = sign(\sum_{i}^{m} \alpha_i y_i \, \kappa(x_i, x_j)) \qquad (11)$$

and the value of the bias is:

$$b = -\frac{1}{2}(\left\{min_{\{y|x_i\}=-1}\middle| \sum\{sv\} \, \alpha_i y_i\right\} + \left\{max_{\{y|x_i\}=-1}\middle| \sum\{sv\} \, \alpha_i y_i\right\}) \qquad (12)$$

3. Core Vector Machine Algorithm

3.1. Core Vector Machine (CVM)

This algorithm is a technique providing equivalence between Minimum Enclosing Ball (MEB) and QP problem of SVM [14]. SVM quadratic programming can be reformulated by CVM as a MEB problem. An efficient $(1 + \xi)$ approximation algorithm is applied to MEB problem to obtain approximate optimal solution for SVM. This algorithm expands the current core-set size by considering the furthest point from the current center.

3.2. Minimum Enclosing Balls (MEB)

Given a finite set of data points $S = \{x_i \in R^d\}$, the minimum enclosing ball of S is defined as the smallest ball that contains all the points in S. Let κ be a Kernel function with features of map ϕ, where $\kappa(x_1, x_2) = <\phi(x_i)\phi(x_j)>$, then the primal MBE problem in Kernel feature space to find $B(C,R)$ .This problem can be expressed as follows:

$$B(C,R) = minR^2 \ \ s.t. \ ||-\phi(x_i)||^2 \leq R^2 \ \forall i \quad (13)$$

The Lagrange dual of this problem is as follows [15]:

Such that:

$$max_{\alpha_i} \sum_{i=1}^{m} \alpha_i \kappa(x_i, x_j) - \sum_{i,j=1}^{m} \alpha_i \alpha_j \kappa(x_i, x_j) \quad (14)$$

$$\sum_{i=1}^{m} \alpha_i = 1 \qquad \alpha_i \geq 0 \quad i = 1, \dots, m \quad (15)$$

When the involved Kernel satisfies $\kappa(x_i, x_j) = k$ as constant value, any SVM quadratic programming can be identified as a MBE. Then the Lagrange dual problem of the MBE can be rewritten as the following QP formulation:

$$\sum_{i=1}^{m} \alpha_i = 1 \qquad \alpha_i \geq 0 \quad i = 1, \dots, m \quad (17)$$

$$min_{\alpha_i} \sum_{i=1}^{m} \alpha_i \kappa(x_i, x_j) - \sum_{i,j=1}^{m} \alpha_i \alpha_j \kappa(x_i, x_j) \quad (16)$$

3.2.1. Approximate MBE

Having support vector machines quadratic problem equation reformulated as a MBE problem, CVM can present a solution for SVM problem with the $(1+\xi)$ approximate MBE problem. The MBE algorithm can terminate within $O(\frac{1}{\xi})$ trials. Therefore, the time and space complexities of CVM are $O(\frac{1}{\xi^2} + \frac{1}{\xi^4})$ and $O(\frac{1}{\xi^2})$ respectively.

Based on these points, a flowchart of CVM algorithm is presented in Figure 3.

4. Proposed Algorithm

The proposed procedure has four steps presented in the flowchart illustrated in Figure 4. The sub-sections of this part explain the steps and structure of this flowchart.

4.1. Data Acquisition (Step I)

The off-line results of different PQ type events are used for knowledge base creation. Voltage signal is used for classification procedure.

4.2. Data Processing (Step II)

Data processing systems typically manipulate raw data into information and likewise information systems typically take raw data as input to produce information as output. Data is basically facts (either organized or unorganized) which can be converted into other forms to make it useful. Figure 5 shows the data processing step.

4.2.1. Feature Extraction (FE)

In the pattern classification problems, it is essential to determine the dimensions of the pattern representation at network input to keep them small to obtain higher classification accuracy and lower computational load and time. Features in this paper are first extracted from data originated from voltage signal and then some of the extracted features are discarded. The detail and approximation coefficients are not directly used as the classifier inputs. Reducing the feature dimension, FE methods are generally implemented to these coefficients at each decomposition level. The algorithm of FE is shown in figure 6.

4.2.1.1. Wavelet Transformation Theory

Wavelet Transform (WT) extracts both time and frequency information of signal simultaneously. Using this aspect of WT, a tremendous technique for analyzing the system waveform can be implemented. WT is particularly useful for studying disturbance or transient waveform, where it is necessary to analyze different frequency components separately. WT can be continuous or discrete. Decomposition Wavelet Transform (DWT) uses low-pass and high-pass filters (LPF/HPF) to divide the frequency-band of the input signal with respect to low and high-frequency components into octave bands.

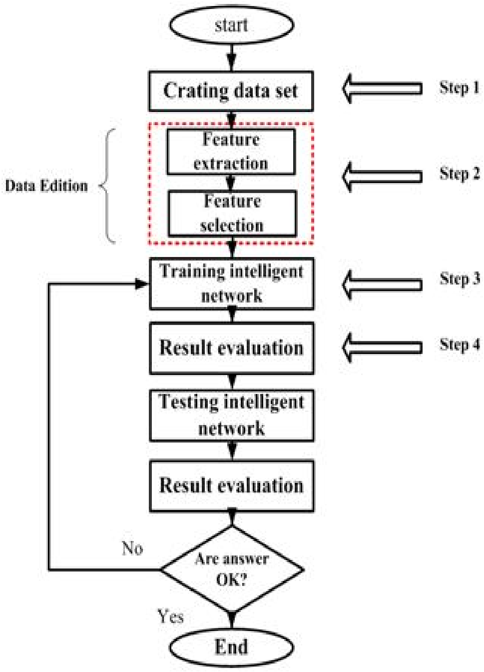

Figure. 3. Flowchart of CVM method

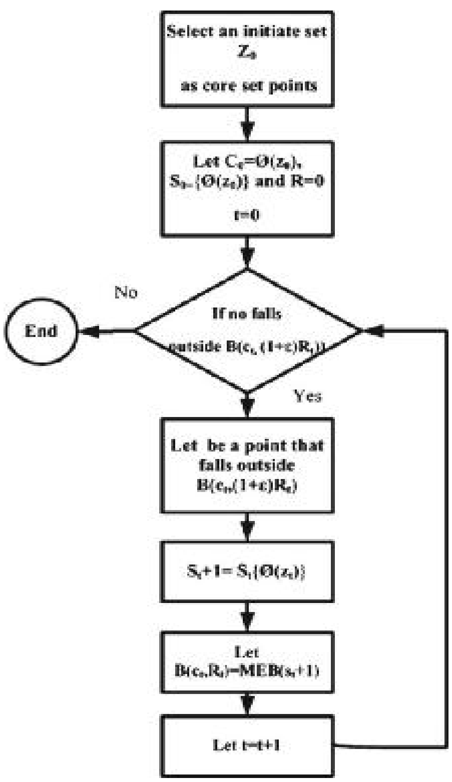

Figure. 4. Flowchart of power quality classification

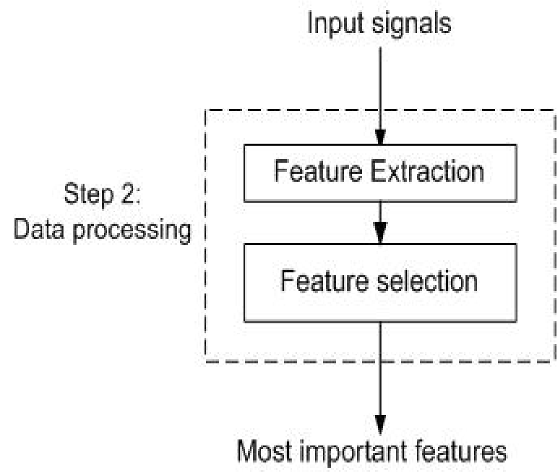

Figure. 5. Data processing

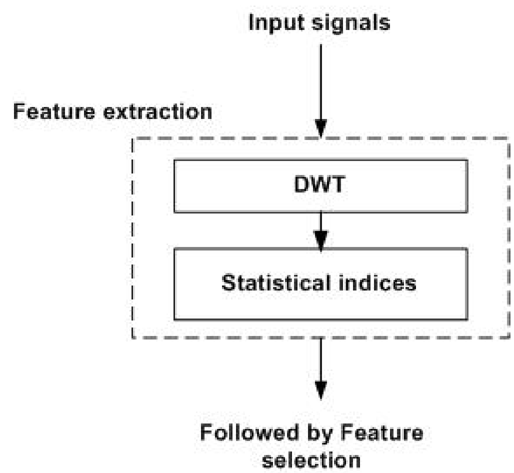

Figure. 6. Feature extraction algorithm

Table.1 Formulations of statistical feature extraction technique.

Feature extraction techniques	Formulation of detail coefficients	Formulation of approximation coefficients
Mean	$\mu_d = \frac{1}{N}\sum_{j=1}^{N} d_{ij}$	$\mu_a = \frac{1}{N}\sum_{j=1}^{N} a_{ij}$
Standard Deviation	$s_d^2 = \frac{1}{N}\sum_{j=1}^{N}(d_{ij}-\mu_d)^2$	$s_a^2 = \frac{1}{N}\sum_{j=1}^{N}(a_{ij}-\mu_a)^2$
Skewness	$S = \sqrt{\frac{1}{s\ N}}\sum_{j=1}^{N}(\frac{d_{ij}-\mu_d}{s_d})^3$	$S = \sqrt{\frac{1}{s\ N}}\sum_{j=1}^{N}(\frac{a_{ij}-\mu_a}{s_a})^3$
Kurtosis	$K = \sqrt{\frac{N}{24}}\{\frac{1}{N}\sum_{j=1}^{N}\left(\frac{d_{ij}-\mu_d}{s_d}\right)^4 - 3\}$	$K = \sqrt{\frac{N}{24}}\{\frac{1}{N}\sum_{j=1}^{N}\left(\frac{a_{ij}-\mu_a}{s_a}\right)^4 - 3\}$
Root Mean Square (RMS)	$RMS_d = \sqrt{\frac{1}{N}\sum_{j=1}^{N} d_{ij}^2}$	$RMS_a = \sqrt{\frac{1}{N}\sum_{j=1}^{N} a_{ij}^2}$
Form Factor	$FF = \frac{\mu_d}{RMS_d}$	$FF = \frac{\mu_a}{RMS_a}$
Crest Factor	$CF = \frac{peak-of-signal-value}{RMS_d}$	$CF = \frac{peak-of-signal-value}{RMS_a}$
Energy	$E = \sum_{j=1}^{N} \lvert d_{ij} \rvert^2$	$E = \sum_{j=1}^{N} \lvert a_{ij} \rvert^2$
Shanon Energy	$SE = -\sum_{j=1}^{N} d_{ij}^2 \log(d_{ij}^2)$	$SE = -\sum_{j=1}^{N} a_{ij}^2 \log(a_{ij}^2)$
Log Energy Entropy	$LEE = \sum_{j=1}^{N} \log(d_{ij}^2)$	$LEE = \sum_{j=1}^{N} \log(a_{ij}^2)$
Inter quartile energy	$IE = d_i\left(\%75th\right) - d_i(\%25th)$	$IE = a_i\left(\%75th\right) - a_i(\%25th)$

Table 2 : Formulation of simulated classes

PQ ev cl	Formulation	Parameters' value
Pure C sinus	$y = A\sin(wt)$	$0.9<A<1.1$
Voltag C	$y = (1-a_{sag}(u(t_2)-u(t_1)))\sin(wt)$	$0.1<\alpha_{sag}<0.9$ and $T<t_2-t_1<9T$
Voltag C swell	$y = (1+a_{swell}(u(t_2)-u(t_1)))\sin(wt)$	$0.1<\alpha_{swell}<0.9$ and $T<t_2-t_1<9T$
Harm C	$y = a_1\sin(wt)+a_3\sin(3wt)+a_5\sin(5wt)+a_7\sin(7wt)$	$0.05<a_3<0.15$, $0.05<a_5<0.15$, $0.05<a_7<0.15$, $\sum a_i^2 = 1$
Sag C harm	$y = (1-a_{sa\&h}(u(t_2)-u(t_1)))a_1\sin(wt)+a_3\sin(3wt)+a_5\sin(5wt)$	$0.1<\alpha_{sa\&h}<0.9$ and $T<t_2-t_1<9T$ $0.05<a_3<0.15$, $0.05<a_5<0.15$, , $\sum a_i^2 = 1$
Swell C harm	$y = (1+a_{sw\&h}(u(t_2)-u(t_1)))a_1\sin(wt)+a_3\sin(3wt)+a_5\sin(5wt)$	$0.1<\alpha_{sw\&h}<0.8$ and $T<t_2-t_1<9T$ $0.05<a_3<0.15$, $0.05<a_5<0.15$, , $\sum a_i^2 = 1$
Interr C	$y = (1-a_{out}(u(t_2)-u(t_1)))\sin(wt)$	$0.9<\alpha_{out}<1$ and $T<t_2-t_1<9T$
Transi C	$y = \sin(wt)+a_{osc}e^{\frac{-(t-t_b)}{t}}\sin(2pf_{osc}(t-t_b))$	$0.1<\alpha_{osc}<0.9, 0.008<$ <0.04

Figure 7 shows different decomposition levels using low and high pass filters. While low-pass filtering produces the approximation coefficients A_j, high-pass filtering produces detail coefficients D_j of the decomposition. The relationship of the approximation coefficients and detail coefficients between two adjacent levels are given as follows:

The DWT is based on decomposition of the original signal into different signals at various levels of resolution. First, the original signal is passed through the two filters producing the detail coefficient D_1 and approximate coefficient A_1 for $j = 1$. After down-sampling by a factor of 2, the approximate coefficients A_1 are passed through the same filters again to obtain the coefficients for $j = 2$. After another down-sampling, the approximate coefficients A_2 are then filtered again to obtain the next level of coefficients. This process continues in an iteration until it reaches the last point n. There are many wavelet functions named as mother wavelets. The choice of mother wavelet is important because different types of mother wavelets have different properties. Popular wavelet functions are Haar, Morlet, Coiflet, Symlet and Daubechies wavelets [15-17].

4.2.1.2. Statistical Indices for FE

Taking into account the frequency components of signals obtained by wavelet transform, statistical indices are implemented to extract

$$A_{j+1}(k) = \sum_n h(n-2k)A_j(n) \qquad (18)$$

$$D_{j+1}(k) = \sum_n g(n-2k)A_j(n) \qquad (19)$$

several indices, such as mean, standard deviated (STD), skewness, kurtosis, RMS, form factor, crest-factor, energy, Shannon-entropy, log-energy entropy and inter quartile energy. Table 1 shows statistical indices that are used as feature extractor [15-17].

4.2.2. Feature Selection

It is applicable to use feature selection methods to deviate the dimensionality of the feature space. Feature selection methods realize the selection of the best subset of the input feature set. In this paper, FISHER and Minimum- Redundancy Maximum-Relevance (MRMR) are used as two feature selection method. Fisher Score introduced selects features that assign similar values to the samples from the same class and assign different values to samples from different classes [18] . Fisher Score is an effective supervised feature selection algorithm. MRMR selects dual features that are mutually far away from each other, while they still have "high" correlation with the classification variable. It is an approximation method to maximize the dependency between the joint distribution of the selected features and the classification variable.

4.3. CVM Training (Step III)

In this step, all feature vectors and the correspondent classes are used to build the SVM and CVM. When full CVM is built, K-fold cross validation is used for the best result from CVM.

4.4. Performance Evaluation (Step IV)

Three different indices are used to evaluate trained CVM. Indices are Total Accuracy (TA), Sensitivity Index (SI) of i^{th} class and Precision Rate Index (PRI) of i^{th} class.
Indices are introduced as follows:

$$TA = \frac{\sum_{i=1}^{K} T_{K_i}}{\sum_{i=1}^{K} T_{K_i} + \sum_{i=1}^{K} F_{K_i}} \tag{20}$$

$$SN_{K_i} = \frac{T_{K_i}}{T_{K_i} + F_{K_i}} \tag{21}$$

$$PR_{K_i} = \frac{T_{K_i}}{T_{K_i} + F_{K_{oi}}} \tag{22}$$

where K_i is the classes index; T_{K_i} is the number of cases which are in class K_i and classified correctly by the trained CVM; F_{K_i} is the number of cases of other classes classified in class K_i; $F_{K_{oi}}$ is the number of cases of other classes classified in class K_i by the trained CVM and K is the number of the classes [14].

5. Simulation Results

The proposed algorithm is used for PQ classification purpose. Eight different classes of power quality events are assessed and simulated. These classes are pure sinusoid, voltage sag, voltage swell, harmonic, sag and harmonic, swell and harmonic, interruption and transients.

5.1. Date Acquisition

For 3000 random values of the parameters in the formulations of each class, the total 24000 data samples are achieved. Table 2 shows the formulations of the above events.

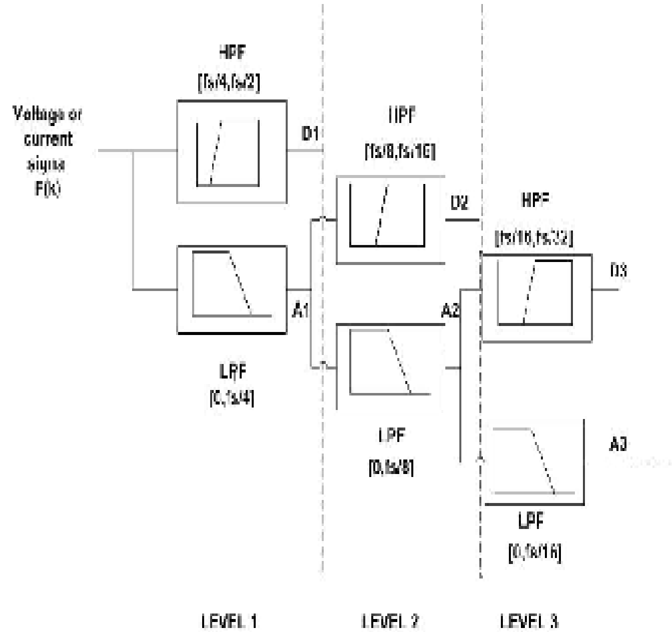

Figure. 7. Algorithm for DWT coefficient decomposition
***HPF: High Pass Filter **LPF: Low Pass Filter**

5.2. CVM Training

Taking into account the K-fold method, the data set is divided into training data and test data. The training data is applied to the algorithm to train CVM. Testing data is used to evaluate the quality of the training algorithm. In this paper, 21000 samples of the total 24000 samples are considered as training data and the rest are used to test the trained algorithm.

Table 3:Output results of the CVM algorithm (Percent)

Cla	C_1	C	C_3	C_4	C_5	C	C_7	C_8
C_1	10	0	0	0	0	0	0	0
C_2	0.6	9	0	0	1.9	0	0	0
C_3	0	0	10	0	0	0	0	0
C_4	0	0	0	10	0	0	0	0
C_5	0	0	0	0	10	0	0	0
C_6	0	0	0	2.9	0	9	0.3	0
C_7	0	0	0	1.6	0	0.	98	0.52
C_8	0	0	0	0	0	0	0	100

5.3. Performance Evaluation

For 26 features, the algorithm is trained for 21000 training data samples. Then three indices are applied to the algorithm for 3000 outputs of the testing data. The output answers are depicted in table 3. This table shows that percentage of outputs of the relevant class are inaccurately 1; for instance, 0.26 percent of outputs of this class are inaccurately 1. The amount of indices for each class are calculated, and the results are shown in tables 4, 5, and 6.

5.3.1. Analysis of Sensitivity Index

The sensitivity analysis is applied to the algorithm output and the results are obtained and shown in table 4. This index shows the main diagonal of table 3. A sensitivity of 100%, as

classes	Sensitivity index Amount
C_1	100%
C_2	97.6%
C_3	100%
C_4	100%
C_5	100%
C_6	96.88%
C_7	98.15%
C_8	100%

depicted in table 4, meaning that all the outputs of that class are correct.

classes	Sensitivity index Amount
C_1	100%
C_2	97.6%
C_3	100%
C_4	100%
C_5	100%
C_6	96.88%
C_7	98.15%
C_8	100%

Table 4: Sensitivity analysis of the results

Table 5: Purity degree analysis of the results

5.3.2. Analysis of Degree of Purity Index

Table 5, shows the degree of purity index of the algorithm results. This index is equal to 100%, for C_3. For instance, meaning that the results of

classes	Purity Degree Index Amount
C_1	99.45%
C_2	100%
C_3	100%
C_4	95.45%
C_5	98.27%
C_6	99.73%
C_7	99.73%
C_8	100%

none of the classes are not 3 except for C_3.

Table 6 : Trust ability analysis of the results

5.3.3. Analysis of Trust Ability Index

Determining how trustworthy the output of the algorithm can be, either the answer actually belongs to the correct class or to the wrong answers of other classes. In fact, this index indicates the probability of accuracy of answers. In table 6, this index is applied to results of the algorithm. This index gives a view of total accuracy of classification.

6. Conclusion

In this paper, the application of CVM in intelligent classification of PQ events has been investigated and successfully validated. In order to evaluate the effectiveness of the proposed method, a four-step algorithm has been conducted. In the first step, a proper data set has been obtained. In the second step, data processing has been evaluated in which a combination of discrete wavelet transform and a statistical coefficient has been used as feature extraction method. Moreover, the feature selection algorithm has also been implemented. In the third step, CVM training has been done, and eventually in the last step, performance evaluation has been investigated using various indices. Also the proposed algorithm has been applied to a test case and the simulation results showed the effectiveness and rapid response of the CVM to classify PQ events.

References

[1] 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 with Applications, vol. 36, pp. 5962-5975, 2009.
[2] H. Erişti and Y. Demir, "A new algorithm for automatic classification of power quality events based on wavelet transform and SVM," Expert systems with applications, vol. 37, pp. 4094-4102, 2010.
[3] K. Silva, B. Souza, and N. Brito, "Fault detection and classification in transmission lines based on wavelet transform and ANN," Power Delivery, IEEE Transactions on, vol. 21, pp. 2058-2063, 2006.
[4] P. G. Axelberg, I.-H. Gu, and M. H. Bollen, "Support vector machine for classification of voltage disturbances," Power Delivery, IEEE Transactions on, vol. 22, pp. 1297-1303, 2007.
[5] P. G. Axelberg, I. Y. Gu, and M. H. Bollen, "Automatic Classification of Voltage Events Using the Support Vector Machine Method," in 19th International Conference on Electricity Distribution, 2007, p. 6.
[6] V. Vapnik, The nature of statistical learning theory: springer, 2000.
[7] L. Jack and A. Nandi, "Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms," Mechanical systems and signal processing, vol. 16, pp. 373-390, 2002.
[8] Z. Moravej, S. Banihashemi, and M. Velayati, "Power quality events classification and recognition using a novel support vector algorithm," Energy Conversion and Management, vol. 50, pp. 3071-3077, 2009.
[9] C. J. Burges, "A tutorial on support vector machines for pattern recognition," Data mining and knowledge discovery, vol. 2, pp. 121-167, 1998.
[10] S. K. Shevade, S. S. Keerthi, C. Bhattacharyya, and K. R. K. Murthy, "Improvements to the SMO algorithm for SVM regression," Neural Networks, IEEE Transactions on, vol. 11, pp. 1188-1193, 2000.
[11] C. Campbell, "Kernel methods: a survey of current techniques," Neurocomputing, vol. 48, pp. 63-84, 2002.
[12] G. Fung and O. L. Mangasarian, "Finite Newton method for Lagrangian support vector machine classification," Neurocomputing, vol. 55, pp. 39-55, 2003.
[13] S.-X. Lu and X.-Z. Wang, "A comparison among four SVM classification methods: LSVM, NLSVM, SSVM and NSVM," in Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on, 2004, pp. 4277-4282.
[14] M. Mohammadi and G. Gharehpetian, "Application of core vector machines for on-line voltage security assessment using a decision-tree-based feature selection algorithm," IET generation, transmission & distribution, vol. 3, pp. 701-712, 2009.
[15] I. Daubechies, Ten lectures on wavelets vol. 61: SIAM, 1992.
[16] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 22, pp. 4-37, 2000.
[17] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 11, pp. 674-693, 1989.
[18] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern classification: John Wiley & Sons, 1999.