Efficient Genetic-Wrapper Algorithm Based Data Mining for Feature Subset Selection in a Power Quality Pattern Recognition Application

Brahmadesam Krishna[1] and Baskaran Kaliaperumal[2]
[1]Department of CSE, Saranathan College of Engineering, India
[2]Department of CSE, Government College of Technology, India

Abstract: *Power quality monitors handle and store several gigabytes of data within a week and hence automatic detection, recognition and analysis of power disturbances require robust data mining techniques. Literature reveals that much work has been done to evolve several feature extraction and subsequent classification techniques for accurate power disturbance pattern recognition .However the features extracted have been rarely evaluated for their usefulness. The objective of this work is to emphasize that feature selection is an important issue in power quality disturbance classification and that genetic algorithms can select good subsets of features. In this paper, a wrapper based approach that integrates multiobjective genetic algorithms and the target learning algorithm is presented in order to evolve optimal subsets of discriminatory features for robust pattern classification. The wavelet transform and the S-transform are utilized to produce representative feature vectors that can accurately capture the unique and salient characteristics of each disturbance. In the training phase the multiobjective genetic algorithms is used to find a subset of relevant attributes that minimizes both classification error rate and size of the classifier discovered by the classification algorithm, using the Pareto dominance approach. Two different classifiers were compared in this study using genetic feature subset selection: decision tree, a feed forward neural network. Moreover two different MOGAs namely elitism-based MOGA and Non-dominated sorting genetic algorithm have been employed separately in the training phase. Experimental results reveal that both of these proposed variants of MOGA combined with classifiers namely decision trees /FFNN yield improved classification performance and reduced classification time as compared to standard classifiers namely decision trees decision tree or standard feed -forward networks. Moreover NSGA performs better than the elitism based approach in terms of classification time.*

Keywords: *Data mining, feature selection, genetic algorithm, power quality, and disturbance recognition.*

Received June 2, 2009; accepted August 3, 2009

1. Introduction

Industrial and commercial customers have equipments that are sensitive to power disturbances, and, therefore, it is more important to monitor and understand the quality of power being provided. Power quality monitoring is the process of gathering and analyzing raw measurement data got by continuous monitoring of voltage and current waveforms and their instantaneous sampled values. The process of analysis and interpretation of this monitored data mostly consists of grouping the captured events into a number of disturbance classes: voltage dips, voltage swells, interruptions, dips and harmonics. This classification is usually called as Power Quality disturbance classification or Power Quality disturbance recognition. The major issues involved in dealing with Power Quality disturbance recognition are the first and foremost of them is to detect and classify disturbance waveforms automatically without any human intervention.

The extreme variability of the signals with regard to spectral content (25 Hz to 5MHz), duration of disturbance (50 ns to steady state) and the magnitude (0 - 4 p.u) makes disturbance recognition all the more difficult. For example, typical voltage fluctuation phenomena due to arc furnaces are characterized by frequencies of less than 25 Hz while impulsive transients due to lightning strikes could have a frequency spectrum of the order of MHz.

Also the recorded data at a customer site could run into gigabytes in a year and this poses a great challenge for data storage, analysis and identification of disturbance data though significant amount of literature exists on automatic classification of power quality disturbances, [1, 2, 3, 5, 6, 7, 8, 11, 12, 17, 18, 19] the quest for new characterization methods and tools for optimal power system disturbance classification still continues. The existing methods reveal that feature extraction is done directly from the original measurements (e.g., RMS values), from some transformed domain (e.g., Fourier [3] and wavelet transforms [2, 6, 7, 8], S-transform [5, 23]) or from the parameters of signal models and that these features are used by artificial neural networks [6, 7], fuzzy logic [3, 11] and other pattern recognition methods like C4.5

decision trees [1]. Though efforts have been taken to design robust feature extraction techniques the extracted features have been rarely evaluated for their usefulness as indicated in [9]. All the features extracted are used and the validity of the features for the classification model is known only at the end of the entire classification task. The main aim of this paper is to emphasize that feature selection prior to classification enhances the power quality pattern recognition task in the following ways

- Improves the accuracy of the classifier or predictor model.
- Provides faster and hence cost-effective model.
- To provide better comprehension of the induced model because a smaller one makes it easier for the user to infer the underlying process.

To substantiate this, a wrapper based approach that integrates MultiObjective Genetic Algorithms (MOGA) and the target learning algorithm is presented in order to evolve optimal subsets of discriminatory features for robust power quality disturbance pattern classification. The subset of features selected by this approach yield a more accurate, comprehensible, compact and hence faster classifier model as compared to the one built using all the extracted features. The rest of the paper is organized as follows: section 1 emphasizes the need for feature selection and introduces our proposed method followed by the system architecture in section 2, feature selection using genetic algorithms in section 3 and experimental results in section 4 respectively. Conclusions and future work are presented in section 5.

1.1. Need for Feature Selection

Most power quality disturbance classification methods reported in the literature [2, 7, 8, 11] use all the features extracted for classification purposes. As a result, irrelevant information might be fed to the classifier which might not allow the classifier to generalize nicely. For instance in decision trees, since the induction C4.5 algorithm uses information gain as the impurity function ,the features used by the index are not always the optimal ones. The presence of irrelevant attributes misleads the impurity functions towards producing bigger, less comprehensible, more error-prone decision tree classifiers. Similarly avoiding irrelevant and redundant features is desirable in artificial neural networks because not only they increase the size of the network and the training time, but also because they may reduce the accuracy of the network .Therefore, the selection of relevant features, prior to using induction, can be expected to lead to a better index. Feature selection reduces the number of features, removes irrelevant, redundant, or noisy data, and brings the immediate effects for applications: speeding up a data mining algorithm, improving mining performance such as predictive accuracy and result comprehensibility [10]. Feature selection algorithms fall under two categories namely filter and wrapper approaches. The wrapper method searches for an optimal feature subset tailored to a particular learning algorithm and a domain whereas the filter approach attempts to assess the merits of features from the data alone. Advantages of wrapper approaches include the interaction between feature subset search and model selection, and the ability to take into account feature dependencies.

1.2. Overview of Proposed Method

Automatic feature subset selection distinguishes the proposed Power Quality disturbance classification method from all other reported approaches. In particular, GAs are employed to select features that encode important power quality information and improve classification performance. GAs belong to the class of randomized heuristic search techniques, offering an attractive approach to feature subset selection. Although GAs have been employed in [15, 14] for effectively training neural networks and hence for recognizing Power Quality disturbance types, they have been rarely used for optimisation of feature selection in a Power Quality pattern recognition.

In this paper, a wrapper based approach using MOGAs is presented for feature subset selection and robust classification of Power Quality disturbances. Two Classifiers Namely Decision Trees (CART) and FeedForward Neural Networks (FFNN) have been used as the target algorithm for the wrapper approach and compared in this study. The goal of the proposed multiobjective GA is to find a subset of attributes that leads to a reduction in both the classification error rate and the complexity of the rule set CART or size of the network FFNN discovered by the classification algorithm. Two different MOGAs namely elitism-based MOGA and Non-dominated Sorting Genetic Algorithms (NSGA) have been employed separately in the training phase. Experimental results reveal that both of these proposed MOGAs combined with CART or FFNNs yield improved classification performance and reduced classification time as compared to CART decision trees or FFNN employing the whole set of features.

2. System Architecture

The power quality classification system which employs hybrid learning using MOGA and decision trees/FFNN is divided into several stages as given in Figure 1. The raw power system data is split into two data sets namely training and test data.

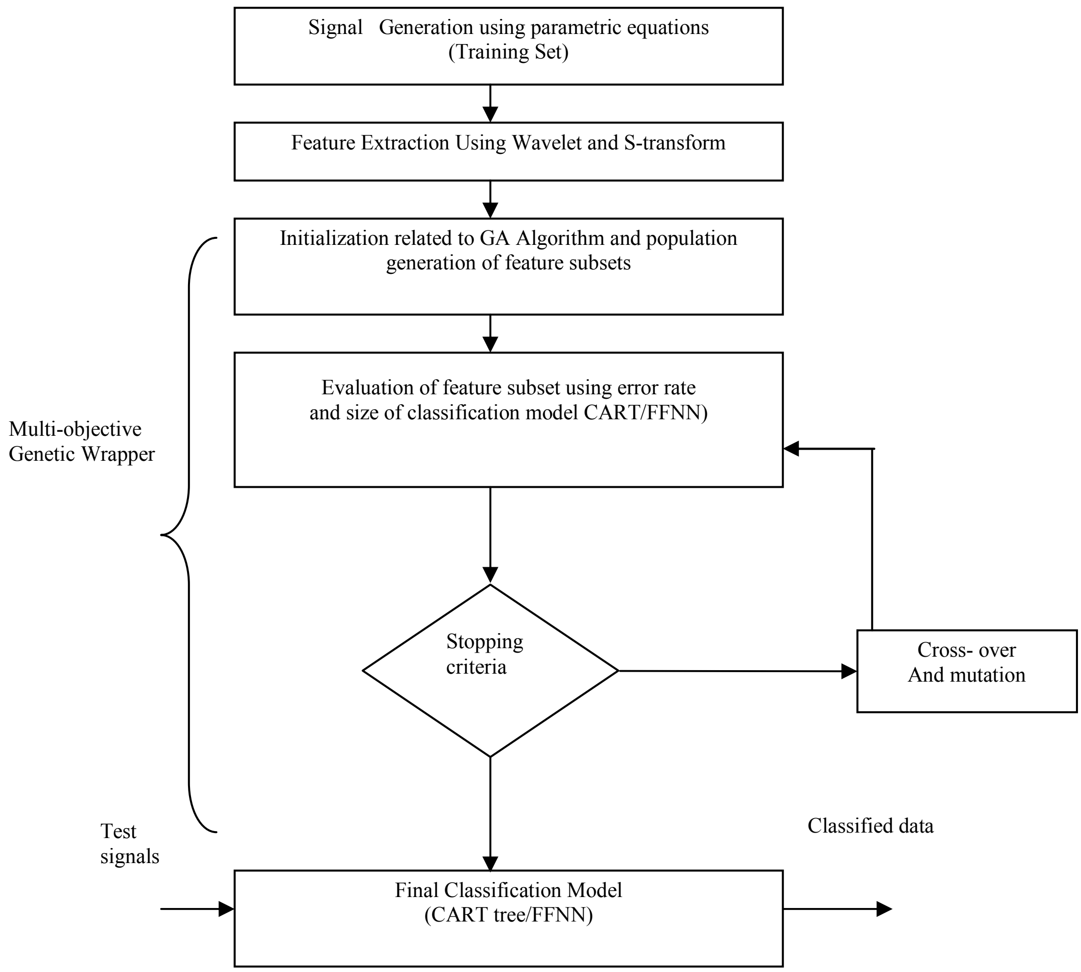

Figure 1. System architecture block diagram.

Initially pertinent features are extracted from the training data using wavelet and S-transform. Then the optimal feature subset is extracted by applying MOGA with the two objectives of minimizing classification error and size of the classifier .This optimal feature subset is used for building the final CART decision tree or FFNN which is used in classification of the test data. It is to be noted that only the optimal features are extracted in the feature extraction of the test data. In this paper power quality disturbances were simulated using parametric equations in Matlab 7.0.

Ten cycles of each disturbance were considered (f=50Hz, amp= 1p.u).The signal sampling rate fs was set at 1.6 KHz. Six classes of disturbances were simulated (sag, swell, sagh, swellh, outage and harmonics).

2.1. Feature Extraction

In power quality applications FFT [3], the discrete wavelet transforms MRA [1, 6, 7, 8] and the S-transform [5, 23] have been used extensively to extract features from power quality data. In our work the discrete wavelet transform MRA and the S-transform were used to extract four and six features respectively resulting in a total of ten features.

2.1.1. Wavelet Based Multiresolution Analysis

The Wavelet transform is a well-suited tool for analyzing the high frequency transients in the presence of low frequency components such as non-stationary and non-periodic wideband signals. The first main characteristic in DWT is the MRA technique that can decompose the original signal into several other signals with different levels (scales) of resolution. From these decomposed signals, the original time-domain signal can be recovered without losing any information. The signal was decomposed into three levels using wavelet based Multiresolution Analysis (MRA). The mother wavelet used for decomposition was ‘db4’. Then by employing Parseval’s theorem the energy distribution at each level was obtained yielding four features (3 detailed versions at each level and one approximate version) as given in Table 3, A(refer appendix).

2.1.2. S-Transform

The S-transform of a disturbance is a complex matrix whose rows are frequencies and columns are time. The S-transform yields contours which are very much viable to visual inspection.Six features were extracted using S-transform as given in Table 3, B (refer appendix).

- std1: the standard deviation of the S-transform contour representing the fundamental frequency denoted by f(f=50Hz).
- std2: max (std(3*f- contour), std(5*f- contour), std(7*f- contour), std(9*f- contour), std(11*f- contour)).
- ampmax1: the max value of amplitude Vs time curve of the fundamental frequency.
- ampmin1: the min value of amplitude Vs time curve of the fundamental frequency.
- ampmax2: max (max (3*f amp Vs t curve), max (5*f- amp Vs t curve), max (7*f- amp Vs t curve), max (9*f- amp Vs t curve), max(11*f- amp Vs t curve)).
- ampmin2: min(min(3*f amp Vs t curve), min (5*f- amp Vs t curve), min (7*f- amp Vs t curve), min (9*f- amp Vs t curve),min(11*f- amp Vs t curve)

3. Learning Using a Multiobjective Genetic Wrapper Approach

The basic idea of our wrapper approach is to use MOGAs to efficiently explore the space of all possible efficiently explore the space of all possible subsets of a given feature set in order to find feature subsets which are of low order and high discriminatory power. In order to achieve this goal, we felt that fitness evaluation had to involve direct measures of classification performance. The goal of the proposed MOGA is to find a subset of relevant attributes that leads to a reduction in both the classification error rate and the complexity (size) of the rule set or network discovered by the data mining algorithm. In this paper the data mining algorithm is CART or FFNN.

3.1. Individual Encoding and Fitness Function

In the proposed MOGA, each individual consists of M genes, where M is the number of original attributes in the data. Each gene is either 0 or 1 indicating attribute occurs or not (respectively) in the candidate subset of selected attributes. The fitness function measures the quality of a candidate attribute subset represented by an individual. The fitness of an individual consists of two quality measures:

a. The error rate of data-mining algorithm.
b. The size of the decision tree built by CART or size of the network built by FFNN.

Both (a) and (b) are computed by running the data-mining algorithm with the individual's attribute subset only, and by using a hold-out method to estimate data-mining algorithm's error rate, as follows. First, the training data is partitioned into two mutually-exclusive data subsets, the building subset and the validation subset. Then we run the data-mining algorithm using as its training set only the examples (records) in the building subset. Once the decision tree or neural network has been built, it is used to classify examples in the validation set.

3.2. Selection Method and Genetic Operators

Two different MOGAs namely elitism-based MOGA and Non-dominated Sorting Genetic Algorithm (NSGA) have been used separately in the training phase and compared in terms of classification accuracy, tree size, number of features and classification.

3.2.1. Elitism-Based MOGA

At each generation, selection is performed as follows. First the MOGA selects all the non-dominated individuals (the Pareto front) [4] of the current population. These non-dominated individuals are passed unaltered to the next generation by elitism. Let *N* be the population size, and let *Nelit* be the number of individuals reproduced by elitism, where *Nelit*□(*N*/2). Then the other *N-Nelit* individuals to reproduce are chosen by performing *N-Nelit* times a tournament selection procedure, as follows. First the MOGA randomly picks k individuals from the current population, where k is the tournament size, a user-defined parameter which is set to 2 in our work. Then the MOGA compares the fitness values of the two individuals on the basis of Pareto dominance and selects the winner. In case none of them dominate the other a tie-breaking procedure is adopted as follows. For each remaining individual *Ii*, the MOGA computes *Xi* as the number of individuals in the current population that are dominated by *Ii* and *Yi* the number of individuals that dominate *Ii*. Depending on *Xi-Yi*, each individual is assigned a fitness value. Finally, if *Xi-Yi* is same for both, the winner is selected randomly. This tie-breaking criterion is also used to reduce the elitist set when *Nelit*>(*N*/2).Then the individuals selected by tournament selection undergo crossover and mutation.

3.2.2. NSGA Selection

In this work, we have used the Nondominated Sorting Genetic Algorithm NSGA proposed by Srinivas and Deb in [20]. The idea behind NSGA is that a ranking selection method is used to emphasize good points and a niche method is used to maintain stable subpopulations of good points. It varies from simple genetic algorithm only in the way the selection operator works. The crossover and mutation remain as usual. Before the selection is performed, the population is ranked on the basis of an individual's nondomination. The nondominated individuals present in the population are first identified from the current population. Then, all these individuals are assumed to constitute the first nondominated front in the population and assigned a large dummy fitness value. The same fitness value is assigned to give an equal reproductive potential to all these nondominated individuals. In order to maintain the diversity in the population, these classified individuals are then shared with their dummy fitness values. Sharing is achieved by performing selection operation using degraded fitness values obtained by dividing the original fitness value of an individual by a quantity proportional to the number of individuals around it. Thereafter, the population is reproduced according to the dummy fitness values. Since individuals in the first front have the maximum fitness value, they get more copies than the rest of the population. The efficiency of NSGA lies in the way multiple objectives are reduced to a dummy fitness function using nondominated sorting procedures. In our work, for the MOGA modules we have used constant population size of 50, 10 generations, □ share of 0.397 (sharing for NSGA algorithm), a crossover rate of 0.7 and a mutation rate of 0.005.

4. Experimental Results

Six classes of disturbances were simulated using parametric mathematical equations in Matlab 7.0 as given in Table 2.

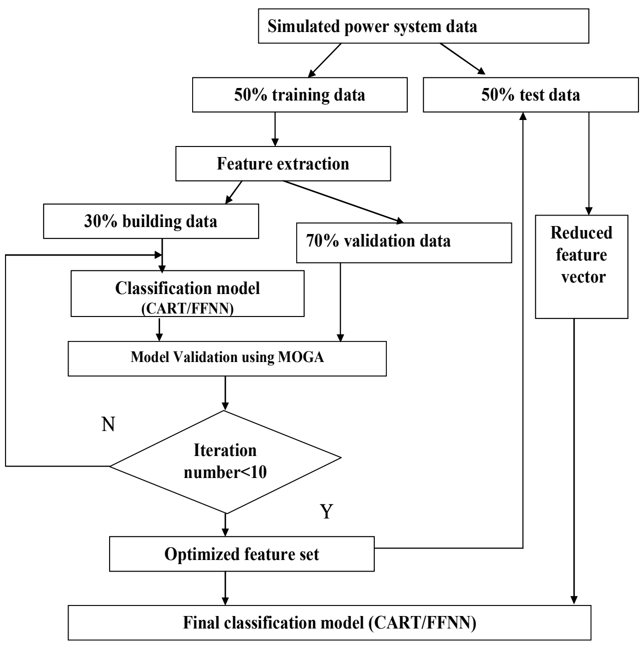

Figure 2. Set-up for GA-CART/FFNN experiments.

The power system data was split into two sets namely training and testing sets. The characteristics of power system data are given in Table 1. The training data consisted of 2160 examples (360 examples per class of disturbance) and the test data consisted of 2160 examples (360 examples per class of disturbance). The experimental setup for the wrapper based learning is given in Figure 2. In the training set 30% of examples were used for inducing decision trees or constructing feed-forward networks and the other 70% for validation of the learned description .Ten runs of the experiment were performed. Again Pareto's dominance concept was applied to the set of dominant solutions obtained from the ten runs and the resulting non-dominated solutions are given in Tables 5, 6, 8 and 9 [see Appendix]. Tables 4 and 7 show classification accuracy and size of model for the CART tree and feed-forward networks respectively (built with all features). It is evident that solution 1 has lower error rate when compared to solution 2 in each case whereas solution 2 has reduced classification time in both elitism based MOGA and NSGA. It is left to the user or domain expert to choose between these solutions. Also FFNN classifier outperforms the decision trees in classification accuracy. However the classification time in FFNN (also training time) is about 1 second higher on an average which is significant when classification has to be done in real time. 5, 6, 8 and 9 show that the NSGA algorithm yields reduced feature set and hence lesser classification time. Experimental results show that high classification accuracy up to 98.98% is obtained from the proposed nsga -based ga+ffnn learning technique and reduced classification time of .2161 sec is obtained from the proposed nsga -based ga+CART learning.

5. Conclusions

The identification and classification of power quality disturbances are critical tasks for precise monitoring of distribution systems. This paper introduces a wrapper based approach that integrates MOGA and the target learning algorithm is presented in order to evolve optimal subsets of discriminatory features for robust pattern classification. By reducing irrelevant information and using only optimised feature subsets, the two classifiers CART decision trees, FFNN showed significant performance improvements in terms of classification accuracy and classification time. Experimental results reveal that our approach could provide valuable insights into other pattern classification problems- how to extract and use only the relevant features for a particular pattern classification task.

References

[1] Abdel K., Kamel M., Youssef M., Youssef F., and Saadany F., "Power Quality Disturbance Classification Using the Inductive Inference Approach," *Computer Journal of IEEE Transactions on Power Delivery*, vol. 19, no. 4, pp. 159-163, 2004.

[2] Angrisani L., Daponte P., and Apuzzo D., "Wavelet Network-Based Detection and Classification of Transients," *Computer Journal*

of IEEE Transactions on Instrumentation and Measurement, vol. 50, no. 5, pp. 1425-1435, 2001.

[3] Dash K., Mishra M., Salama A., and Liew C., "Classification of Power System Disturbances Using a Fuzzy Expert System and a Fourier Linear Combiner," *Computer Journal of IEEE Transactions on Power Delivery*, vol. 15, no. 2, pp. 472-477, 2000.

[4] Deb K., *Multi-Objective Optimization Using Evolutionary Algorithms*, John Wiley and Sons, 2001.

[5] Fengzhan Z. and Rengang Y., "Power Quality Disturbance Recognition Using S-Transform," *Computer Journal of IEEE Transactions on Power Delivery*, vol. 22, No 2, pp. 422-427, 2007.

[6] Gaing L., "Implementation of Power Disturbance Classifier Using Wavelet-Based Neural Networks," *in Proceedings of IEEE Bologna Power Tech Conference*, Italy, pp. 7-7, 2003.

[7] Gaing L, "Wavelet-Based Neural Network for Power Disturbance Recognition and Classification," *Computer Journal of IEEE Transactions on Power Delivery*, vol. 19, no. 4, pp. 1560-1568, 2004.

[8] Gaouda M., Salama A., Sultan R., and Chikhani Y., "Power Quality Detection and Classification Using Wavelet-Multiresolution Signal Decomposition," *Computer Journal of IEEE Transactions on Power Delivery*, vol. 14, no. 4, pp. 1469-1476, 1999.

[9] Gunal E., Rifat G., Omer E., and Dogan G., "Feature Selection for Power Quality Event Classification," *in Proceedings of Signal Processing, Communications and Applications Conference*, pp. 1-4, 2008.

[10] Huan L. and Lei Y., "Toward Integrating Feature Selection Algorithms for Classification and Clustering," *Computer Journal of IEEE Transactions on Data and Knowledge Engineering*, vol. 17, no. 4, pp. 491-502, 2005.

[11] Huang J., Negnevitsky M., and Nguyen T., "A Neural-Fuzzy Classifier for Recognition of Power Quality Disturbances," *Computer Journal of IEEE Transactions on Power Delivery*, vol. 17, no. 2, pp. 609-616, 2002.

[12] Heydt T. and Galli W., "Transient Power Quality Problems Analyzed Using Wavelets," *Computer Journal of IEEE Transaction Power Delivery*, vol. 12, no. 2, pp. 908-915, 1997,

[13] Lee H. and Nam W., "Efficient Feature Vector Extraction for Automatic Classification of Power Quality Disturbances," *Computer Journal of Electronics Letters*, vol. 34, no. 11, pp. 1059-1061, 1998.

[14] Li G., Zhou M., and Zhang Z., "Power Quality Disturbance Automatic Recognition Based on Wavelet and Genetic Network," *in Proceedings of Tencon*, China, pp. 1923-1926, 2002.

[15] Ouyang S., Song Z., Wang J., and Chen D., "Application of LVQ Neural Networks Combined with Genetic Algorithm in Power Quality Signals Classification," *in Proceedings PowerCon*, China, pp. 491-495, 2002.

[16] Santoso S. and Lamoree D., "Power Quality Data Analysis: From Raw Data To Knowledge Using Knowledge Discovery Approach," *in Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference*, USA, pp. 172-177, 2000.

[17] Styvaktakis E., Bollen J., and Gu H., "Expert System for Classification and Analysis of Power System Events," *Computer Journal of IEEE Transactions on Power Delivery*, vol. 17, no. 2, pp. 423-428, 2002.

[18] Styvaktakis E., *Automating Power Quality Analysis*, PhD Thesis, Chalmers University of Technology, Sweden, 2002.

[19] Santoso S., "A Scalable PQ Event Identification System," *Computer Journal of IEEE Transactions on Power Delivery*, vol. 15, no. 2, pp. 738-743, 2000.

[20] Srinivas N. and Deb K., "Multiobjective Optimization Using Non-Dominated Sorting in Genetic Algorithms," *Computer Journal of MIT Press Journal* of *Evolutionary Computation*, vol. 2, no. 3, pp. 221-248, 1994.

[21] Wang M. and Mamishev V., "Classification of Power Quality Events Using Optimal Time-Frequency Representations-Part 1: Theory," *Computer Journal of IEEE Transactions on Power Delivery*, vol. 19, no. 3, pp. 1488-1495, 2004.

[22] Wang M., Rowe I., and Mamishev V., "Classification of Power Quality Events Using Optimal Time-Frequency Representations-Part 2: Application," *Computer Journal of IEEE Transactions on Power Delivery*, vol. 19, no. 3, pp. 1496-1503, 2004.

[23] Yong Z., Hao Z., Yi-Feng D., Gan-Yun L., and Yi-Bin S., "S-Transform-Based Classification of Power Quality Disturbance Signals by Support Vector Machines," *in Proceedings of the Chinese Society for Electrical Engineering*, China, vol. 25, no. 4, pp. 51-56, 2005.

Brahmadesam Krishna received her Bachelor of engineering degree in electrical and electronics engineering from Government College of Technology, Coimbatore, India in 1994 and Masters in computer science and engineering from Government College of Engineering, Tirunelveli India in 2002. She is now pursuing PhD in data-mining from Anna University-Coimbatore, India. She is a Chennai, India in 2006. He is a member of IEEE and member of ISTE. He is now assistant professor in computer science and engineering, Government member of Institution of Engineers, India and member of ISTE. She is working as assistant professor in computer science and engineering, Saranathan College of Engineering, Trichy, India. Her area of interest includes signal processing, artificial intelligence, and machine learning algorithms.

Baskaran Kaliaperumal received his Bachelor of engineering degree in electrical and electronics engineering from the Annamalai University, India in 1989, Master of engineering degree in computer science engineering from Bharathiar University, India in 2002 and PhD degree from Anna University-College of Technology, Coimbatore, India. His area of interest includes adhoc networks, network security.

Appendix

Table 1. Characteristic of power system data.

Event	Controlling Parameter	Equation
Pure sine	NA	$v(t) = \mathrm{A}\sin(\omega t)$
Sudden Sag	$.1 \le \alpha \le .9, T \le t_2 - t_1 \le 9T$	$v(t) = \mathrm{A}(1 - \alpha(u(t_2) - u(t_1)))\sin(\omega t)$
Sudden Swell	$.1 \le \alpha \le .8, T \le t_2 - t_1 \le 9T$	$v(t) = \mathrm{A}(1 + \alpha(u(t_2) - u(t_1)))\sin(\omega t)$
Harmonics	$.05 \le \alpha_i \le .15$ $\Sigma \alpha_i^2 = 1$ i=3,5,7	$v(t) = A\begin{pmatrix} \sin(wt) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5wt) \\ + \alpha_7 \sin(7wt) \end{pmatrix}$
Outage	$.9 \le \alpha \le .1, T \le t_2 - t_1 \le 9T$	$v(t) = \mathrm{A}(1 - \alpha(u(t_2) - u(t_1)))\sin(\omega t)$
Sag with harmonics (sagh)	$.9 \le \alpha \le .1, T \le t_2 - t_1 \le 9T$ $.05 \le \alpha_i \le .15$ $\Sigma \alpha_i^2 = 1$ i=3,5,7	$v(t) = \mathrm{A}(1 - \alpha(u(t_2) - u(t_1)))*$ $\begin{pmatrix} \sin(wt) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5wt) \\ + \alpha_7 \sin(7wt) \end{pmatrix}$
Swell with harmonics (swellh)	$.8 \le \alpha \le .1, T \le t_2 - t_1 \le 9T$ $.05 \le \alpha_i \le .15$ $\Sigma \alpha_i^2 = 1$ i=3,5,7	$v(t) = \mathrm{A}(1 + \alpha(u(t_2) - u(t_1)))*$ $\begin{pmatrix} \sin(wt) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5wt) \\ + \alpha_7 \sin(7wt) \end{pmatrix}$
A= 1 , ω=2 □ ft , f=50Hz		

Table 2. Parametric equations for disturbance modelling.

Characteristic	Description
Number of examples	Training data of 2160 examples and test data of 2160examples
Number of features	10
Feature values	The feature value is numerical and continuous in the range 0-1
Number of classes	6
Distribution of classes in training data	360 numbers in each class
Distribution of classes in testing data	360 numbers in each class

Table 3. Features for each class of power disturbance A. using DWT.

S.No	App	One	Two	Three	Type	Class
1	0.44	2.44E-05	4.10E-04	5.39E-03	sag	1
2	0.50	1.13E-05	2.75E-04	5.96E-03	swell	2
3	0.42	1.44E-02	6.54E-03	2.98E-02	sagh	3
4	1.47	4.60E-02	2.17E-02	8.85E-02	swellh	4
5	8.E-02	1.70E-03	7.05E-04	1.00E-02	out	5
6	0.47	1.20E-02	5.87E-03	2.17E-02	harmonic	6

B. using S-transform.

S.No	Std1	Std2	Ampmax1	Ampmin1	Ampmax2	Ampmin2	Type	Class
1	9.35E-02	2.18E-02	0.49	0.16	0.108	0.0012	sag	1
2	1.09E-02	5.29E-03	0.53	0.49	0.097	1.14E-02	swell	2
3	9.11E-02	5.82E-02	0.49	0.1757	0.204	5.38E-04	sagh*	3
4	0.15	0.105	0.94	0.509	0.388	6.46E-03	swellh*	4
5	0.199991	0.041129	0.497075	9.45E-16	0.1118	0	out	5
6	7.16E-03	5.15E-02	0.49	0.472	0.185	2.70E-03	harmonic	6

Sagh ind** Sagh indicates sag with harmonics * Swellh indicates swell with harmonics

Table 4. Solution obtained by feedforward neural network (all features applied).

FFNN		
Classifier Error Rate	**Network Size [i/p-hidd-o/p]**	**Classifier Time**
6.01%	36 [10-20-6]Nodes 346 weights	2.5059sec

Table 5. Set of Dominant Solutions by Elitism GA+FFNN.

ELITISM-BASED GA+ FFNN						
Classifier Error Rate	**Network Size [I/P-Hidd-O/P]**	**No of Features out of Ten**	**Classifier Time**	**%Reduction in Weight Matrix**	**%Reduction in Features**	**%Reduction in Time**
2.97%	32[6-20-6]nodes 266 weights	6	1.4 sec	23.12	30	44.13
3.65%	31 [5-20-6]Nodes 246 weights	5	1.365 sec	28.90	40	45.53

Table 6. Set of dominant solutions by NSGA GA+FFNN.

NSGA BASED GA+ FFNN						
Classifier Error Rate	**Network Size [I/P-Hidd-O/P]**	**No of Features out of Ten**	**Classifier Time**	**%Reduction in Weight Matrix**	**%Reduction in Features**	**%Reduction in Time**
1.02%	30[4-20-6]Nodes 226 weights	4	1.35 sec	34.68	60	46.13
3.65%	15[3-6-6]Nodes 66 weights	3	1.31 sec	80.92	70	47.72

Table 7. Solution obtained by CART Decision Tree.

CART DECISION TREE		
Classification Error Rate	**Tree Size**	**Classification Time**
7.73%	29 nodes	.3878sec

Table 8. Set of dominant solutions by elitism GA+CART.

ELITISM-BASED GA+CART						
Classifier Error Rate	**Tree Size**	**No of Features out of Ten**	**Classifier Time**	**%Reduction in Weight Matrix**	**%Reduction in Features**	**%Reduction in Time**
3.45%	23	7	0.2769 sec	20.69	30	28.59
3.72%	23	6	0.2674 sec	20.69	40	31.04

Table 9. Set of dominant solutions by NSGA +CART.

NSGA -BASED GA+CART						
Classifier Error Rate	**Tree Size**	**No of Features out of Ten**	**Classifier Time**	**%Reduction In Weight Matrix**	**%Reduction in Features**	**%Reduction in Time**
3.45%	23	5	0.2459 sec	20.69	50	36.59
3.72%	***23***	***4***	*0.2161 sec*	***20.69***	***60***	***44.28***