

Classification of Power Quality Disturbances Using GA Based Optimal Feature Selection

K.R. Krishnanand¹, Santanu Kumar Nayak¹, B.K. Panigrahi², V. Ravikumar Pandi²,
and Priyadarshini Dash³

¹Department of Electrical Engineering, Silicon Institute of Technology, Bhubaneswar, India

²Department of Electrical Engineering, Indian Institute of Technology, Delhi, India

³Department of Electrical Engineering, NIT Durgapur, India

krishkr09@gmail.com, santanu.nayak87@gmail.com,
bkpanigrahi@ee.iitd.ac.in, ravikumarpandi@gmail.com,
priyadarshini.nitdgp@gmail.com

Abstract. This paper presents a novel technique for power quality disturbance classification. Wavelet Transform (WT) has been used to extract some useful features of the power system disturbance signal and Gray-coded Genetic Algorithm (GGA) have been used for feature dimension reduction in order to achieve high classification accuracy. Next, a Probabilistic Neural Network (PNN) has been trained using the optimal feature set selected by GGA for automatic Power Quality (PQ) disturbance classification. Considering ten types of PQ disturbances, simulations have been carried out which show that the combination of feature extraction by WT followed by feature reduction using GGA increases the testing accuracy of PNN while classifying PQ signals.

Keywords: Gray-coded Genetic Algorithm, Power quality disturbances, Wavelet transform, Probabilistic Neural Network.

1 Introduction

In recent years, power quality has become a significant issue for both utilities and customers. Power quality issues [1] and the resulting problems are the consequences of the increasing use of solid state switching devices, non-linear and power electronically switched loads, unbalanced power systems, lighting controls, computer and data processing equipments as well as industrial plant rectifiers and inverters. A power quality (PQ) problem usually involves a variation in the electric service voltage or current, such as voltage dips and fluctuations, momentary interruptions, harmonics and oscillatory transients causing failure or mal-operation of the power service equipment. Hence to improve power quality, fast and reliable detection of the disturbances and the sources and causes of such disturbances must be known before any appropriate mitigating action can be taken. However, in order to determine the causes and sources of disturbances, one must have the ability to detect and localize these disturbances. In the current research trends in power quality studies, Wavelet transform (WT) [2-4] is widely used in analyzing non-stationary signals for power quality assessment [5,6]. In order to identify the type of disturbance present in the power

signal more effectively, several authors have presented different methodologies based on combination of wavelet transform (WT) and artificial neural network (ANN) [7]. Using the multi-resolution properties of WT [6], the features of the disturbance signal are extracted at different resolution levels and are used for the classification purpose. Gaing [8] demonstrated the classification of 7 types of PQ events by using wavelets and probabilistic neural network (PNN) [9].

In this paper, we have tried to decompose the power system disturbance signal up to 13 level of decomposition as mentioned in [8]. Instead of taking only one statistical feature of the decomposition levels, seven different statistical measures like energy, entropy, standard deviation, mean, kurtosis, ITD and skewness are applied to each of the decomposition level. This leads to increase the feature set length up to 91. It is generally well known that as the length of the feature set increases, the classification accuracy increases but the computational complexity becomes very high. Hence, an attempt has been made to reduce the dimensionality of the feature space by finding the optimal feature set using gray coded Genetic Algorithm [12].

2 Wavelet Transform

The Discrete Wavelet Transform (DWT) is a special case of the WT that provides a compact representation of a signal in time and frequency that can be computed efficiently. The DWT is calculated based on two fundamental equations: the scaling function $\phi(t)$, and the wavelet function $\psi(t)$, where

$$\phi(t) = \sqrt{2} \sum_k h_k \phi(2t - k) \tag{1}$$

$$\psi(t) = \sqrt{2} \sum_k g_k \phi(2t - k) \tag{2}$$

The scaling and wavelet functions are the prototype of a class of orthonormal basis functions of the form

$$\phi_{j,k}(t) = 2^{\frac{j}{2}} \phi(2^j t - k); \quad j, k \in Z \tag{3}$$

$$\psi_{j,k}(t) = 2^{\frac{j}{2}} \psi(2^j t - k); \quad j, k \in Z \tag{4}$$

The parameter j controls the dilation or compression of the function in time scale and amplitude. The parameter k controls the translation of the function in time. Z is the set of integers. Once a wavelet system is created, it can be used to expand a function $f(t)$ in terms of the basis functions

$$f(t) = \sum_{l \in Z} c(l) \phi_l(t) + \sum_{j=0}^{J-1} \sum_{k=0}^{\infty} d(j,k) \psi_{j,k}(t) \tag{5}$$

Where, the coefficients $c(l)$ and $d(j,k)$ are calculated by inner product as

$$c(l) = \langle \phi_l | f \rangle = \int f(t) \phi_l(t) dt \tag{6}$$

$$d(j,k) = \langle \psi_{j,k} | f \rangle = \int f(t) \psi_{j,k}(t) dt \tag{7}$$

The expansion coefficients $c(l)$ represent the approximation of the original signal $f(t)$ with a resolution of one point per every 2^j points of the original signal. The expansion coefficients $d(j,k)$ represent details of the original signal at different levels of resolution. $c(l)$ and $d(j,k)$ terms can be calculated by direct convolution of $f(t)$ samples with the coefficients h_k and g_k , which are unique to the specific mother wavelet chosen. Daubechies wavelet family is one of the most suitable wavelet families in analyzing power system transients. In the present work, the db4 wavelet has been used as the wavelet basis function for power quality disturbance detection and classification.

3 Feature Extraction and Reduction

3.1 Feature Extraction

Data preparation is same as the one detailed in [10]. The detailed co-efficient D_{ij} at each decomposition level is used to extract the features. Statistical features like energy, standard deviation, mean, kurtosis, skewness, Shannon entropy and, Instantaneous Transient Disturbance (ITD) of the decomposition coefficients D_{ij} are calculated.

Thus in the present case, with a '13' level of decomposition of the original signal, the feature vector adopted is of length '7*13' i.e., 91. The feature vector is denoted as Feature= [ED₁ED₂ ...ED₁₃σ₁σ₂ ...σ₁₃μ₁μ₂ ...μ₁₃KRT₁ KRT₂ ... KRT₁₃SK₁ SK₁ SK₁₃ ENT₁ENT₂ ...ENT₁₃ITD₁ITD₂ITD₁₃]. (8)

3.2 Gray Coded Genetic Algorithm and Optimal Feature Selection:

Genetic Algorithm (GA) [11] is a well known meta-heuristic algorithm which mimics the evolution process employed by nature for finding the best chromosome. Here, each decision variable provides a value for finding the best vector all together. A variant of GA termed as Gray Coded Genetic Algorithm [12], developed by replacing the binary representation scheme of classical GA with the gray coded binary representation scheme was proposed to remove the problem of premature convergence encountered in binary GA. In this paper a discrete version of GGA is used, wherein variables are allowed to pass through discrete values only; for better performance in problems where the search range consists of discrete values. Chaos Gray coded genetic algorithm (CGGA) is also developed in [13]. Pseudo-code for GGA is given below:

1. Initialization: Initialize the chromosome population with k random vectors. $\vec{x}_1, \dots, \vec{x}_k$ which are binary coded vectors
2. Evaluation: Each vector is evaluated to obtain the respective fitness value
3. Crossover: The exchange of binary values between chromosomes is done by roulette wheel selection and subsequent crossover.
4. Mutation: The mutation is done by random change of 1 to 0 and 0 to 1.
5. Reinsertion: The new chromosomes thus evolved are compared against the current population and the chromosomes with lesser fitness value in the population are replaced by them.
6. Repeat from step 2 until a certain number of iterations has been performed.

The parameters of the search are

- k , the size of the memory. A typical value is in the order of 10 to 50.
- $P_{\text{crossover}}$ the probability of crossover. A typical value is 0.6
- P_{mutation} the probability of mutation. A typical value is 0.01

Here in our problem we have to select some of the features among the whole set in order to get better classification accuracy. In the present work, dimension of each vector of the GGA was taken to be 13 and each component of the vectors are initialized with 13 values out of random permutation of integers between 1 and 91. Thus, the proposed technique allows the algorithm to select a number of features which is equal to 13. Also, during each mutation step the target vectors are allowed to pass through integral values in this region only. As such, the proposed technique searches a small range and hence is expected to show better convergence rate. During the search process the fitness of each vector is evaluated using PNN classification accuracy so that our objective becomes maximization of classification accuracy. Finally, when a stopping criterion is reached (either by end of iterations or through stagnation) components of the vector with the best fitness in the population, gives the optimally selected features.

4 Classification of Disturbances Using Probabilistic Neural Network (PNN)

The PNN model is one among the supervised learning networks. The learning speed of the PNN model is very fast, making it suitable for fault diagnosis and signal classification problems in real time. Fig. 1 shows the architecture of a PNN model that is composed of the radial basis layer and the competitive layer. The learning and recalling processes of the PNN for classification problems can be found in [14].

5 Results and Discussion

Ten classes (C1-C10) of different PQ disturbances are taken for classification and they are as follows:

C1→	Voltage Sag	C6→	Sag with Harmonic
C2→	Voltage Swell	C7→	Swell with Harmonic
C3→	Harmonic Distortion	C8→	Interruption
C4→	Flicker	C9→	Spike
C5→	Oscillatory transients	C10→	Notch

The power quality signals corresponding to these ten classes are generated in Matlab [15] using parameterized models with different parameter values. Wavelet transform of these data samples are then performed to decompose the signals up to 13th level. The statistical features of the decomposed levels constitute the feature vector. Based on the extracted feature, the feature data sets for training and testing are constructed separately. The classification accuracy of the data set is computed using PNN

Table 1. Optimally selected features by GGA

Serial no.	Selected Index	Selected Feature Type	Serial no.	Selected Index	Selected Feature Type
1	24	Std. deviation of 11 th level	8	69	Entropy of 4 th level
2	40	Kurtosis of 1 st level	9	75	Entropy of 10 th level
3	41	Kurtosis of 2 nd level	10	81	ITD of 3 rd level
4	43	Kurtosis of 4 th level	11	82	ITD of 4 th level
5	56	Skewness of 4 th level	12	88	ITD of 10 th level
6	57	Skewness of 5 th level	13	90	ITD of 12 th level
7	60	Skewness of 8 th level			

Table 2. Comparison of classification accuracy results

Features	Classification accuracy	Features	Classification accuracy
Energy (13)	98.86	Entropy (13)	96.84
Std. deviation (13)	98.77	ITD (13)	98.59
Mean (13)	85.76	Energy and Entropy (26)	98.60
Kurtosis (13)	98.86	All (91)	93.89
Skewness (13)	97.21	GGA selected features(13)	99.30

classifier for automatic classification of PQ events using all 91 features, only energy features, only entropy features and, entropy-energy features combined together, successively. Next, GGA is used to select the optimal feature set amongst available 91 features and overall classification accuracy is next compared with that of the previous cases. 13 features are selected by GGA as shown in table 1 and overall classification accuracy is compared in table 2.

6 Conclusion

In this paper an attempt has been made to improve the overall performance of a PNN classifier using optimal feature selection. GGA algorithm has been used for that purpose and results corroborate the feature extraction mechanism showing high classification accuracy of PQ disturbance at reduced complexity.

Acknowledgement

The authors acknowledge the financial grant by Department of Science and Technology (DST), Govt of India, for the research project on Power Quality Assessment in Distribution Network, to be carried out in EE Department, IIT Delhi (SR/S3/EECE/46/2007).

References

1. Bollen, M.H.J.: *Understanding Power Quality: Voltage sags and Interruptions*. IEEE Press, NewYork (2000)
2. Daubechies, I.: The wavelet transform, time/frequency location and signal analysis. *IEEE Transactions on Information Theory* 36, 961–1005 (1990)
3. Mallat, S.G.: A theory of multi resolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11, 674–693 (1989)
4. Meyer, Y.: *Wavelets and Operators*. Cambridge University Press, London (1992)
5. Santoso, S., Powers, E.J., Grady, W.M., Hofmann, P.: Power quality assessment via wavelet transform analysis. *IEEE Transactions on Power Delivery* 11, 924–930 (1996)
6. Gaouda, A.M., Salama, M.M.A., Sultan, M.K., Chikhani, A.Y.: Power Quality Detection and Classification Using Wavelet-Multi resolution Signal Decomposition. *IEEE Transactions on Power Delivery* 14, 1469–1476 (1999)
7. Santoso, S., Powers, E.J., Grady, W.M., Parsons, A.: Power quality disturbance waveform recognition using wavelet-based neural classifier, Part 1: theoretical foundation. In: *The 1997 IEEE/PES Winter Meeting*, New York,U.S.A (1997)
8. Gaing, Z.L.: Wavelet-Based Neural Network for Power Disturbance Recognition and Classification. *IEEE Trans. on Power Delivery* 19, 1560–1568 (2004)
9. Specht, D.F.: Probabilistic neural networks. *Neural Networks* 3, 109–118 (1990)
10. Panigrahi, B.K., Ravikumar Pandi, V.: Optimal feature selection for classification of power quality disturbances using wavelet packet-based fuzzy k-nearest neighbour algorithm. *IET Generation Trans. Distr.* 3, 296–306 (2009)
11. Michalewicz, Z.: *Genetic Algorithms + Data Structures = Evolution Programs*. Springer, Berlin (1999)
12. Chakraborti, N., Mishra, P., Erkoc, S.: A Study of the Cu Clusters Using Gray-Coded Genetic Algorithms and Differential Evolution. *Journal of Phase Equilibria and diffusion* 25, 16–21 (2004)
13. Yang, X., Yang, Z., Yin, X., Li, J.: Chaos gray-coded genetic algorithm and its application for pollution source identifications in convection–diffusion equation. *Comm. In non linear science and numerical simulation* 13, 1676–1688 (2008)
14. MATLAB, Math Works, Inc., Natick, MA, USA (2000)