International Journal of Engineering & Technology, 7 (4.19) (2018) 550-555

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Data Mining Based Power Quality Disturbance Detection Using Wavelet Transform

P. Hariramakrishnan[1*], S. Sendilkumar[2]

[1]*Research scholar, Sathyabama Institute of Science and Technology, Chennai-119*
[2]*Professor, Panimalar Engineering College, Chennai-123*
**Corresponding Author Email: harirmk@gmail.com*

Abstract

This paper presents the identification and classification of power quality disturbances using wavelet transform and data mining models. The various power quality events like oscillatory transient, voltage swell and voltage sag are simulated for a small distribution network using MATLAB/SIMULINK. The simulated events are then processed by the wavelet transform with Daubechies Db5 mother wavelet and wavelet coefficients are generated. The energy, standard deviation, mean, skewness, kurtosis and variance are obtained for these wavelet coefficients for the purpose of distinguishing the events. A data mining based Random Forest model and Decision Tree model are then generated using these features of voltage signal and these models are used for final classification into sag, swell and transients. The test result shows that the wavelet algorithm along with decision tree is more sharp in identifying events.

Keywords: *Power quality, oscillatory transients, voltage swell, voltage sag, Parseval's theorem, wavelet transform and Decision Tree(DT).*

1. Introduction

The power companies and their customers are giving more importance for Power Quality (PQ) issues. Both the power companies and their customers at distributed centers equipped with sensitive devices may have power outages for a short period of time that will lead to a large amount of loss. In order to reduce the occurrence of outage, and to improve the quality of electrical power the study of power quality disturbance detection is must [1]. Examining of PQ events such as transients, sag and swell using various signal transforming methods like Wavelet Transform (WT), Wavelet Packet Transform (WPT), S-Transform, Fourier Transform (FT), Discrete Fourier Transform (DFT) and Short-Time Fourier Transform (STFT) is discussed in literature. FT and STFT employ only a constant window size which is non-preferable for the study of the transient signals [2]. In reference [3], various electric system disturbances are analysed in time scale domain using wavelet transform. A complex wavelet network may also be used for analysing power system disturbances. In this paper authors employed various mother wavelet functions to identify the voltage sag [4]. In another method WPT is used for examining power quality disturbances [5].
Discrete Fourier Transform (DFT) is used for assessing signals of constant amplitude with predefined window portion only. DFT is not efficient in following the changes in fast changing transient signals. However for regular and constant amplitude signals, the quantity and angle of frequency components are obtained using this DFT method. The current or voltage waveform changes are obtained by combining the output groups of rectangular DFT sample-windows. In any case, DFT gives details in frequency-domain only and its resolution is based on time-window width [6].

Wavelets are suitable for the analysis of fast fluctuating transient signals. The wavelets are characterised by their fluctuating waveform, having zero average value and its amplitude is decaying to zero instantly. In contrast to DFT analysis, concurrent assessment of a waveform with distinct resolutions in frequency and time-domain is possible in wavelet analysis[7]. When the time details are essential rather than the details of frequency of the disturbance in a transient investigation, then wavelets are more preferred[8]. The digital depiction of Continuous Wavelet Transform (CWT) is called Discrete Wavelet Transform (DWT). DWT can be realized with a multistage filter bank having wavelet as low pass filter (LP) and its duple as high pass filter (HP). Using the multistage filter bank, the given signal is decomposed or down sampled by two in each stage and is given as input to subsequent stages [9].
The proposed method presents a power quality disturbances (PQD) detection technique for a small distribution network using wavelet transform and data mining based decision tree model. The power system disturbances are obtained using a SIMULINK model of an actual electrical energy distribution system developed in MATLAB. The PQD features in the voltage signal are obtained using wavelet transform and decision tree(DT) model and random Forest(RF) model of data mining package are developed using these features. The classification of PQD is based on the DT model(10). The rest of this paper is organized as follows: Section 2 describes the introduction and application of wavelet transform for obtaining PQD features. Section 3 presents the methodology of the proposed work. Section 4 and section 5 presents the simulation model of the system studied and different PQD features to be extracted respectively. The simulation results and the DT model and RF model for feature classification are illustrated in section 6 and section 7 respectively. Finally the results and discussions and conclusion are illustrated in section 8 and 9 respectively.

2. Introduction to Wavelet Transform

The STFT and FT utilize windows of constant width for the analysis of whole signal considered. During 1975-1981, a French geophysicist J.Morlet proposed wavelet theory in which the high frequency elements are evaluated with small time period and the low frequency elements are evaluated with long time period in distinct frequencies. The wavelets are small duration fluctuating waveforms, having zero average value and zero decaying amplitude. Since the wavelet window function is having finite length, they are said to be compactly supported. Since the function is oscillatory, it is referred as wave. The mother wavelet function $\psi(\tau,s)$ in the following CWT equation 1 performs the transforming function where τ is the translation factor [11].

$$CWT_x^{\Psi}(\tau,s) = \Psi_x^{\Psi}(\tau,s) = \frac{1}{\sqrt{|s|}}\int x(t)\Psi^*\left(\frac{t-\tau}{s}\right)dt \qquad (1)$$

The term mother is used because the function that provides different regions of support for the transformation process is derived from the main function, called mother wavelet. The other window functions are produced on the basis of the mother wavelet. Wavelet coefficients are coincidence indicators computed between the mother wavelet and the signal considered [12]. The voltage signals are decomposed into different levels using low pass and high pass filters of a multistage filter bank as shown in Figure 1. The LP filter will result in approximation coefficients of the original signal, and the HP filter results in detailed coefficients of the signal [13].

2.1 Reasons for using the Wavelet Transform

The wavelet transform has some advantages when compared to the FT. Wavelets have compact support, which means that wherever the function is not defined it will have a value of zero. It helps in speeding up the computations and also tells us about the locality of the wavelet in the time domain.

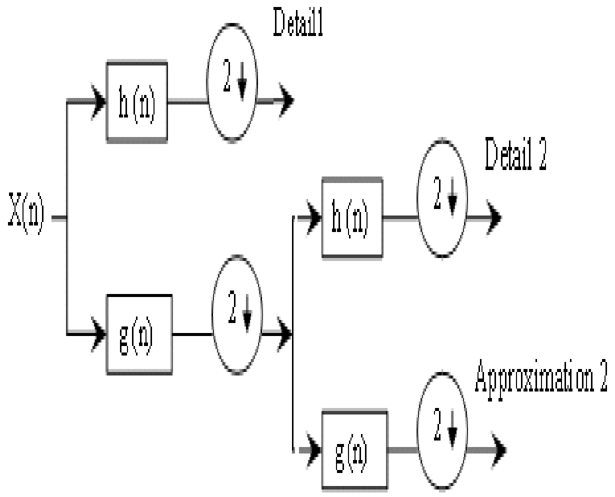

Fig. 1: Implementation of the discrete wavelet transform

The FFT is a tool for high speed implementations of the Fourier transform that provides sines and cosines as basis functions, which are non-local, and stretch out to infinity. They do a very poor job in detecting power quality disturbances and cannot provide the time localization of the frequency components in a signal.

2.2 Selection of mother wavelets

The selection of the mother wavelet plays an important role in the time frequency analysis. These have to be carefully selected for better resemblance and for acquiring the spikes of the signal considered. Proper selection of mother wavelet results in better estimation of the original signal, and it also affects the fault signal frequency spectrum. One of the simplest wavelet in the wavelet family is the Haar wavelet. Harr wavelet is the first proposed wavelet and is seldom used in practice, because it is not continuous and inadequate energy compaction. The Daubechies wavelet is one of the most progressive wavelets that was demostrated by several researchers. While sharp transitions and fast attenuation are not allowed in Haar, Daubechies allows continuous derivatives that behaves well for discontinuous signals. The frequency property of Daubechies wavelets is better than Haar wavelets. Also decomposition of a signal into low frequency and high frequency bands is not efficient in Haar wavelet The Daubechies wavelets permits the user to set the amount of acceptable fluctuations in the high frequency band.

3. Procedure of the proposed work

Figure 2 shows the steps organized for the proposed system.

- Simulation of PQ events such as swell, sag and transients are developed using a power distribution system model and built employing MATLAB/SIMULINK software.
- Data collected from events using MATLAB/SIMULINK are treated with wavelet transform and DB5 mother wavelet is used.
- Data of energy, standard deviation and other features are calculated for data collection.
- Decision tree model is developed for the extracted features.
- Classification of power quality events are performed using extracted features.

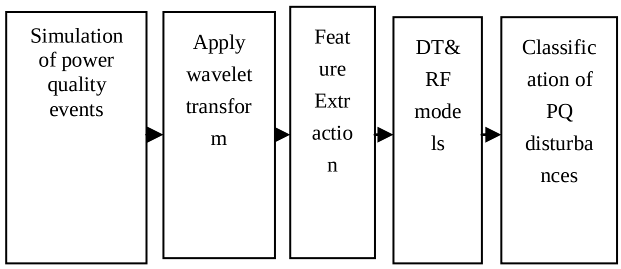

Fig. 2: Methodology of the proposed work

4. System Studied

For the present study, the power quality disturbances are simulated for a real electrical energy distribution system associated with actual data of the network using MATLAB/SIMULINK. The simulation model shown in Figure 3 is constructed with a 11kV, 30MVA, 50Hz three phase source block which supplies a 100kW resistive and 100VAR inductive load through a 1MVA, 11kV/400V delta/wye transformer. The voltage swell is simulated by disconnecting a considerable amount of load connected in the same network. Disconnection of smaller loads does not have much impact on the rest of network. In order to have an observable swell in the voltage waveform, a load of 500kW is disconnected from the 400V bus through a three phase breaker as shown in Figure 3.

A line fault model is constructed to simulate voltage sag caused by line fault. A three phase fault block is located at 11kV bus to simulate the line fault. With the line fault model, various models like single line to ground, double line to ground, line to line and three phase fault can be simulated. In the simulation model, fault block is set to simulate single line to ground fault with a fault impedance of 1Ω. The oscillatory transient voltage wave can be simulated by introducing a capacitor bank energizing model for power factor correction in the power system as shown in Figure 3. The capacitor bank is connected to the 400V bus through a three phase capacitor contactor. The voltage transient frequency is less for a large size capacitor bank. A 0.2 second simulation time is set for all the simulation models.

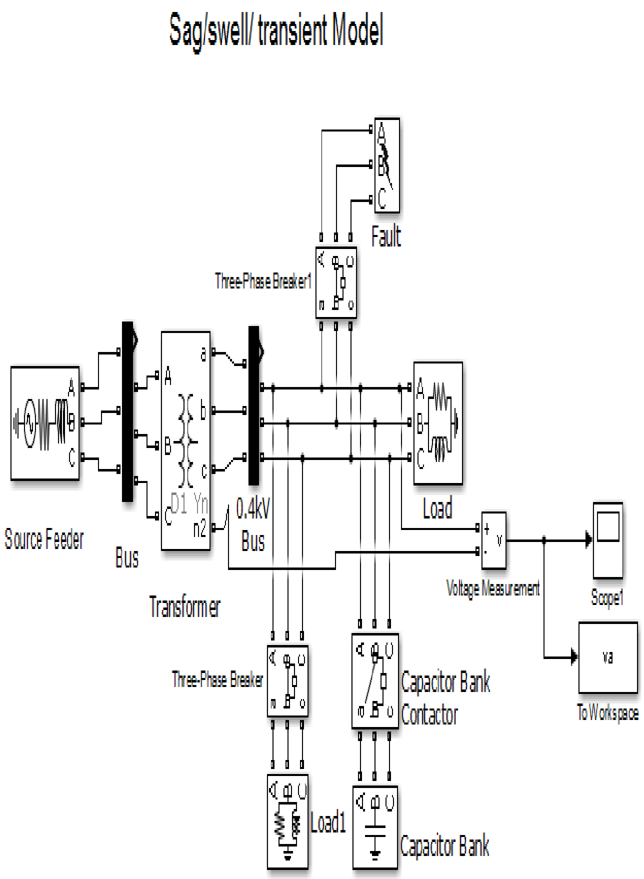

Fig. 3: Simulink model of the system studied

5. Feature extraction

From the statistical analysis of voltage signal disturbances, different primary and secondary statistical parameters can be obtained. A set of these parameters such as energy, standard deviation, mean, kurtosis, skewness and variance are selected for our study. The statistical features are described as below.

5.1. Energy

The energy of DWT decomposed voltage signal disturbances can be given by

$$\frac{1}{N}\sum_{T}\left|X(t)^{2}\right| = \frac{1}{N_j}\sum_{h}\left|cA_{j,h}\right|^{2} + \sum_{J=1}^{J}\left(\frac{1}{N_j}\sum_{h}\left|cD_{j,h}\right|^{2}\right) \tag{2}$$

Where N is the period of sampling. In equation 2, the total energy on right of equal sign is the sum of average energy of the detailed and approximated versions. In this paper, five level of detailed version energy distribution and one level of approximate version energy distribution are considered. The various feature of signals are extracted from different PQ events by employing the coefficients cA and cD of approximated and detailed versions.

5.2 Standard deviation (σ)

It is a measure of the effective energy or power content of the given signal. It gives a measure of how much a signal is deviated from the normal waveform.

$$\sigma = \sqrt{\frac{n\sum x^{2} - (\sum x)^{2}}{n(n-1)}} \tag{3}$$

5.3 Mean ($\bar{x}$)

It gives the average value of the given signal over a period of time.

5.4 Skewness

Skewness is a the degree of asymmetry of a distribution around its mean value. It is computed using the formula given in equation 4.

$$Skewness = \frac{n}{(n-1)(n-2)}\sum\left(\frac{x_i = \bar{x}}{\sigma}\right)^{3} \tag{4}$$

where n - number of samples taken over a period of time.

5.5 Kurtosis

It indicates whether the waveform is normal or spiky. For a spiky waveform, its value is high.

$$Kurtosis = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum\left(\frac{x_i = \bar{x}}{\sigma}\right)^{4}\right] - \frac{3(n-1)^{2}}{(n-2)(n-3)} \tag{5}$$

5.6 Variance

It is the average of the squared differences from the mean. It is also given by the square of standard deviation.

$$Variance = \frac{n\sum x^{2} - (\sum x)^{2}}{n(n-1)} \tag{6}$$

6. Simulation Results

The simulation is carried out for the distribution system shown in Figure 3. The necessary PQ disturbances are generated by either disconnecting a considerable part of load or by introducing a line to ground fault or by energizing a capacitor bank. The simulation is carried out for the following 675 test cases for the three type of PQ disturbances considered. For each event, the simulation is done on all the 3 phases with 5 source voltage phase angles (0^o, 45^o, 90^o, 210^o, 225^o) and 5 disturbance inception angles (0^o, 45^o, 90^o, 210^o, 240^o). The voltage swell is simulated for the disconnection of three different load values. Sag is also simulated for three cases by introducing line to ground fault in all the three phases individually. Similarly three different capacitor banks are considered for the simulation of voltage oscillatory transients. A set of statistical parameters such as energy, standard deviation, mean, kurtosis, skewness and variance are computed for these 675 cases by considering the entire 10 cycles of the output waveform. And an another set of 675 cases are computed by considering only the cycles for which the disturbance prevails i.e. for two cycles from 0.10sec. Therefore a total of 1350 cases given in Table 1are taken for our study.

Table 1: Test cases

Events	Simulation conditions	Total
Swell	Disconnection of three different values of load (300kW, 500kW, 700kW) Parameters computation for (i) entire cycle and (ii) cycles during which PQD presents in the waveform, here two cycles considered Source voltage phase angle (0^o, 45^o, 90^o, 210^o, 225^o) Disturbance inception angles (0^o, 45^o, 90^o, 210^o, 240^o) Parameters computation for all the phases	450 (3x2x5x5x3)
Sag	LG fault introduced at all three phases individually Parameters computation for (i) entire cycle	450 (3x2x5x5x3)

	and (ii) cycles during which PQD presents in the waveform, here two cycles considered Source voltage phase angle (0^o, 45^o, 90^o, 210^o, 225^o) Disturbance inception angles (0^o, 45^o, 90^o, 210^o, 240^o) Parameters computation for all the phases	
Transients	Energization of three different values of capacitor bank (10kVAR, 20kVAR, 30kVAR) Parameters computation for (i) entire cycle and (ii) cycles during which PQD presents in the waveform, here two cycles considered Source voltage phase angle (0^o, 45^o, 90^o, 210^o, 225^o) Disturbance inception angles (0^o, 45^o, 90^o, 210^o, 240^o) Parameters computation for all the phases	450 (3x2x5x5x3)

6.1 Voltage swell

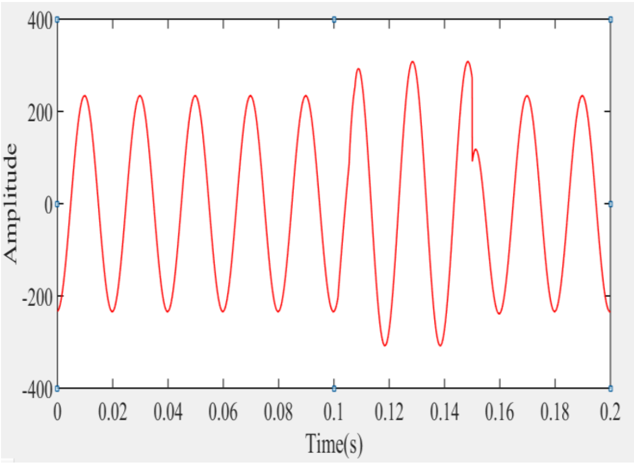

Fig. 4a: Generation of voltage swell

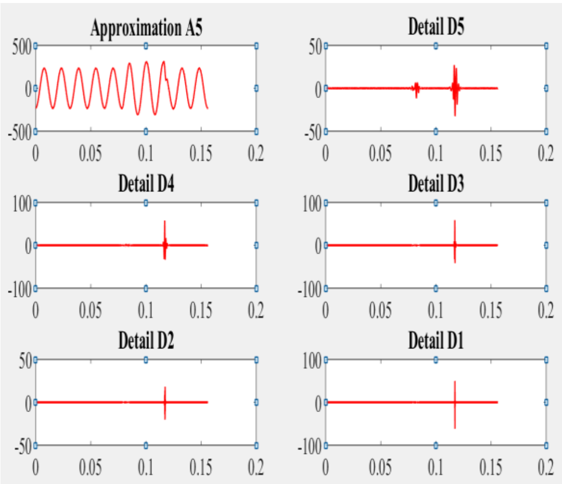

Fig. 4b: Amplitudes of wavelet co-efficient for voltage swell
Fig. 4: Simulation results of voltage swell

Figure.4 depicts the characteristics of voltage swell. Figure. 4a shows the voltage swell generation and 4b shows wavelet transform amplitudes. Voltage swell is generated as a result of de-energization of large amount of load connected to the secondary of transformer and Figure 3c shows the amplitudes of wavelet

coefficients of voltage swell for which energy, standard deviation and other features are extracted for the duration of entire voltage swell waveform as well as for the duration for which a swell presents in the waveform. The energy is computed using Parseval's theorem and a set of calculated parameters are given in the Table 2.

6.2 Voltage sag

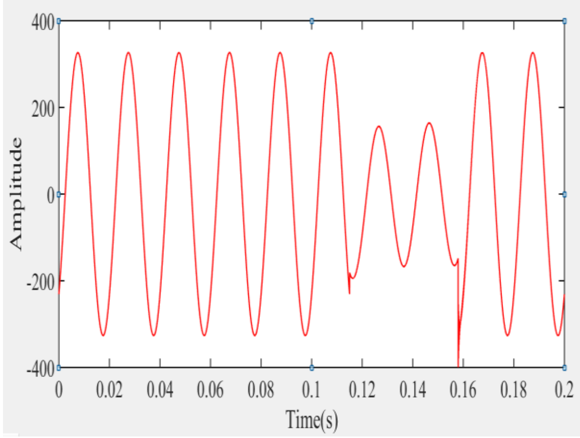

Fig. 5a: Generation of sag

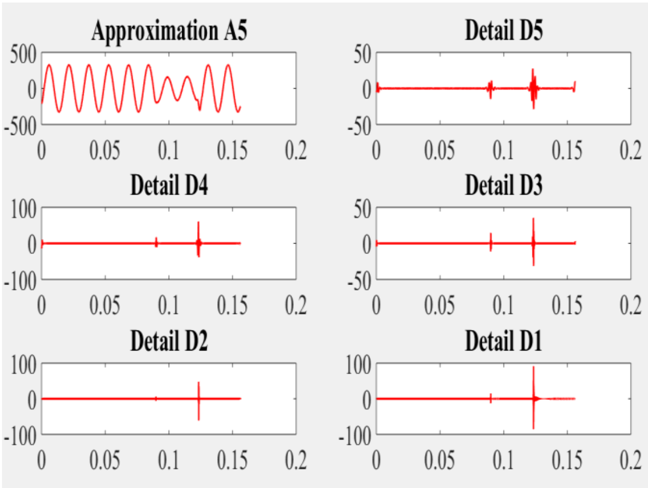

Fig. 5b: Voltage sag wavelet co-efficient amplitudes
Fig. 5: Simulation results of voltage sag

Figure.5 shows the characteristic of voltage sag. The voltage sag generated due to fault on the transmission line is shown in Figure. 5a and its amplitudes shown in Figure. 5b are produced by processing them with wavelet transform. Duration of the voltage sag is occurring from 0.1 to 0.4 sec in Figure. 5a. The energy, standard deviation and other features are calculated for the entire duration of cycles and for the cycles during which the sag presents in the waveform. A sample of extracted features are given in the Table 2.

6.3 Transients

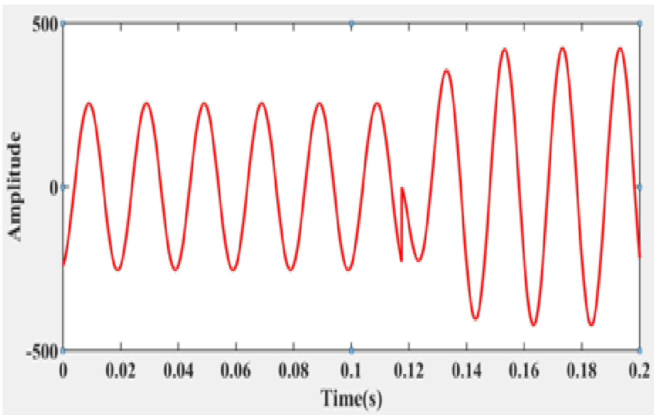

Fig. 6a: Generation of transients due to capacitor switching

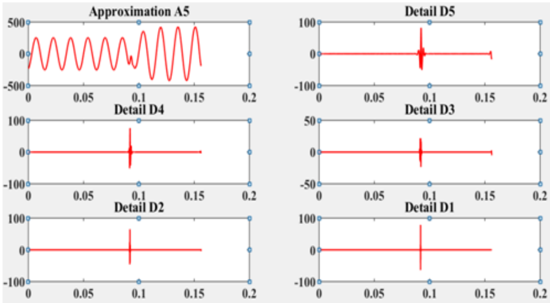

Fig. 6b: Wavelet coefficient amplitudes of oscillatory transients
Fig. 6: Simulation results of oscillatory transients

Power system model for simulation of transients is given in Figure. 3. Oscillatory transients due to capacitor switching in power system is given in Figure. 6a and Figure.6b shows the results of wavelet transform of transients due to capacitance

switching. Here the features like energy, variance, standard deviation and other features are computed for the duration of transients for 0.10 to 0.14sec and for the complete cycle.. This calculated features can be discriminated with the features of other events like sag and swell.

Table 2: Sample of PQ events with statistical parameters

Power quality events		**Standard Deviation (STD)**	**Mean**	**Energy**	**Kurtosis**	**Skewness**	**Variance**
Voltage sag	Phase A	233.04	2.42	1.76E+08	1.82E+03	79.44	4.46E+04
	Phase B	233.26	0.30	2.13E+08	1.80E+03	98.28	5.33E+04
	Phase C	231.18	1.13	1.94E+08	1.85E+03	80.79	4.81E+04
Voltage swell	Phase A	168.63	0.52	1.09E+08	2.58E+03	92.58	2.72E+04
	Phase B	167.87	0.63	1.08E+08	2.50E+03	88.32	2.69E+04
	Phase C	170.58	1.23	1.06E+08	2.57E+03	91.75	2.67E+04
Transients	Phase A	237.35	8.79	1.98E+08	2.66E+03	95.99	4.98E+04
	Phase B	236.75	-2.19	2.01E+08	2.64E+03	95.33	5.00E+04
	Phase C	230.70	-2.85	2.09E+08	2.35E+03	85.62	5.22E+04

Table 2 shows a sample set of energy, standard deviation, mean, kurtosis, skewness and variance values obtained from the wavelet coefficients of power quality events. These values are computed for different source voltage angle and for different instants of disturbance inception angle. Figure 4a, Figure 5a and Figure 6a shows the generation of swell, sag and transients respectively and their corresponding amplitudes of wavelet coefficients are shown in Figure 4b, Figure 5b and Figure 6b respectively. By feeding these values to a decision tree model, the type of power quality disturbance occurring in the network can be easily identified.

7. Decision Tree Model for Feature Classification

A decision tree(DT) is a tree based knowledge representation methodology for predicting classification rules. It consists of a root and a number of branches and leaves. The decision tree model is generated for predicting the performance of given system based on the input data given[14]. It is a part of data mining tool with which a model can be created for classification. For most of the power system problems it is commonly used for decision making[15,16]. It is very simple and user friendly. Along with DT, another data mining model namely Random Forest(RF) is also used and compared with DT. RF is more accurate in classifying the given input. In this paper, the DT and RF models are created using the data mining package 'R'.
DT and RF models are fed with 945 training data(70%) and 405 testing data(30%) out of the total 1350 cases. A section of DT model for the given input data is shown in Figure 7. Even though six parameters are given for classification to DT model, only four parameters are utilized for decision making. The following five rules can be obtained from the DT model shown in Figure 7.

Rule 1: If (STD<=230.48625) & (Energy<=136650000), then status is set to sag.
Rule 2: If (STD<=230.48625) & (Energy>136650000), then status is swell.
Rule 3: If (STD>230.48625) & (skewness>84.5923), then status is swell.
Rule 4: If (STD>230.48625) & (skewness<=84.5923) & (kurtosis<=2090), then status is sag.
Rule 5: If (STD>230.48625) & (skewness<=84.5923) & (kurtosis>2090), then status is transient.

The model was tested with a mixture of all the three type of PQ disturbances. The confusion matrix of DT and RF models is given in Table 3. The RF model classified 403 out of 405 test cases correctly whereas DT model classified 399 out of 405 test cases correctly.

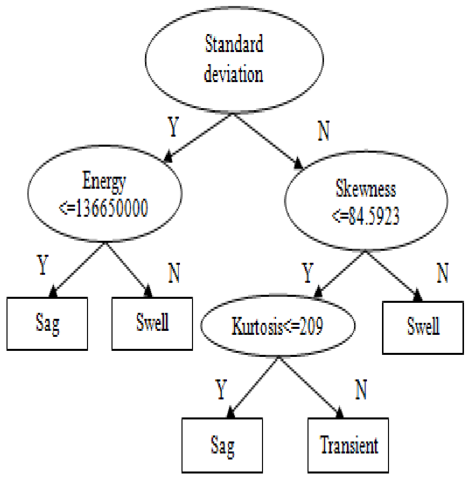

Fig. 7: A section of DT model

A rule set is formed from the extracted features and it can be fed as an input to the next stage of automatic classification tools like a fuzzy classifier. A fuzzy rule set is obtained mostly by intuition and domain knowledge[13]. This paper presents the use of decision tree to generate the rules automatically from the feature set. The statistical features of the PQD are extracted and good features that discriminate the different PQD conditions are selected using decision tree.

8. Discussions

The performance of wavelet transform method combined with data mining models was evaluated based on the confusion matrix given in Table 3. From the results, it can be stated that both DT and RF models were good in detecting the power quality disturbances. Even though the RF model is accurate in prediction, its complexity cause it difficult for real time implementation of data signal processing.

Table 3: The confusion matrix

Predicted	DT model			RF model		
	1(Swell)	2(Sag)	3(Transient)	1(Swell)	2(Sag)	3(Transient)
1(Swell)	133	3	1	134	1	0
2(Sag)	1	132	0	1	134	0
3(Transient)	1	0	134	0	0	135

While classifying swell, the DT model misclassified 1 test case as sag and 1 test case as transient and remaining all were classified correctly. Similarly, 3 cases in sag and 1 case in transient were

misclassified. However, in RF model there is only one misclassification is sag and swell and all cases are classified correctly in transients.

9. Conclusion

In this paper, a simple and efficient method for identification and classification of power quality events using wavelet transform and DT model and RF model is presented. The wavelet transform processes the voltage waveforms and produces the wavelet coefficients. Data of energy, standard derivation, mean, kurtosis, skewness and variance values of these wavelet coefficients are found using Parseval's theorem and other statistical formulae. Using these data, the power quality events are easily identified and can be classified by using the data mining models. The effectiveness of the proposed method has been verified on a real time energy distribution system for different operating conditions. The obtained results are found to be good in terms of accuracy and speed. The work proposed can be extended to automatic detection and prevention of power quality disturbances.

References

[1] P.Kanirajan, V.Suresh Kumar, "Power quality disturbance detection and classification using wavelet and RBFNN", Applied Soft Computing, Volume 35, October 2015, Pages 470-481.

[2] SafdarRaza, HazlieMokhlis, HamzahArof, J.A.Laghari, LiWang, "Application of signal processing techniques for islanding detection of distributed generation in distribution network: A review" Energy Conversion and Management", Volume 96, 15 May 2015, Pages 613-624.

[3] Raj Kumar, Bhim Singh, D.T. Shahani, Ambrish Chandra, Kamal Al-Haddad, " Recognition of Power- Quality Disturbances Using S-Transform-Based ANN Classifier and Rule-Based Decision Tree", IEEE Transactions on Industry Applications, Volume 51(2), 2015 , pp.1249-1258.

[4] Mohammad BarghiLatran, AhmetTeke, "A novel wavelet transform based voltage sag/swell detection algorithm", International Journal of Electrical Power & Energy Systems, Volume 71, October 2015, Pages 131-139

[5] Naik C, Hafiz Faizal ; Swain Akshya, Kar AK, "Classification of power quality events using wavelet packet transform and extreme learning machine", IEEE 2nd Annual Southern Power Electronics Conference (SPEC). IEEE, New York, USA, 2016. DOI:10.1109/SPEC.2016.7846169.

[6] ZhengyouHe, ShibinGao, XiaoqinChen, JunZhang, ZhiqianBo, QingquanQian, "Study of a new method for power system transients classification based on wavelet entropy and neural network", International Journal of Electrical Power & Energy Systems, Volume 33, Issue 3, March 2011, pp. 402-410.

[7] Samir Avdakovi, Nejra Čišija, "Wavelets as a tool for power system dynamic events analysis – State-of-the-art and future applications", Journal of Electrical Systems and Information Technology, Volume 2(1), May 2015, pp. 47-57.

[8] Abdelazeem A. Abdelsalam, Azza A. Eldesouky, Abdelhay A. Sallam, "Characterization of power quality disturbances using hybrid technique of linear Kalman filter and fuzzy-expert system", Electric Power Systems Research, vol.83,No.1,Feb.2012,pp.41-50.

[9] Prakash K.Ray, Soumya R.Mohanty, NandKishor, " Disturbance detection in grid-connected distributed generation system using wavelet and S-transform", Electric Power Systems Research, Volume 81, Issue 3, March 2011, Pages 805-819.

[10] M.Saimurugan, K.I.Ramachandran, V.Sugumaran, N.R.Sakthivel, "Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine", Expert Systems with Applications, Volume 38, Issue 4, April 2011, Pages 3819-3826.

[11] ZhangYang, LiCe, LiLian, " Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods", Applied Energy, Volume 190, March 2017, pp. 291-305.

[12] RuqiangYan, Robert X.Gao, XuefengChen, "Wavelets for fault diagnosis of rotary machines: A review with applications", Signal Processing, Volume 96, Part A, March 2014, Pages 1-15.

[13] S.Sendilkumar, B.L.Mathur, Joseph Henry, "Differential protection of power transformer using wavelet transform and PNN", International Journal of Electrical and Electronics Engineering, vol.4, No.7, 2010, pp.468-474.

[14] M. Amarnath, V. Sugumaran, Hemantha Kumar, "Exploiting sound signals for fault diagnosis of bearings using decision tree", Measurement, 2013, Volume 46, 1250–1256.

[15] Mishra DP, Samantaray SR, Zadeh MD, "A combined wavelet and data-mining based intelligent protection scheme for microgrid", IEEE Transaction on Smart Grid, 2016, Volume 7(5), 2295-304.

[16] KavaskarSekar, Nalin KantMohanty, "Data mining-based high impedance fault detection using mathematical morphology", Computers & Electrical Engineering, July 2018, Volume 69, Pages 129-141.