

An Intelligent Technique for Fault Analysis in Smart Grids

Reem Alameri ^a, Muhammad Munir ^a, Sajid Hussain ^b, Ali Al Alili ^a *, Ehab El-Sadaany ^c

^a Department of Mechanical Engineering, Khalifa University of Science and Technology,
PO Box 12778, Abu Dhabi, United Arab Emirates

^b TELUS Telecommunication Inc., 3777 Kingsway, Burnaby, BC, CANADA V5H 3Z7

^c Department of Electrical Engineering and Computer Science, Khalifa University of Science and Technology, PO Box 127788, Abu Dhabi, United Arab Emirates

Abstract

Power systems are moving towards smart grids through the incorporation of new digital technologies and equipment that increases the system complexity. The power systems become more prone to many types of failures such as cyber-attacks and sensors failure. Therefore, efficient fault analysis is essential to maintain normal grid operations. In this research, an intelligent technique is proposed to detect, classify, and understand the propagation behaviors of the short circuit faults induced in an IEEE 39-bus system. The proposed method is also capable of identifying faults causes and consequences, and suggesting remedial actions. The IEEE 39 bus system is modeled under normal and faulty conditions. A number of datasets are created from the smart grid model, which are then processed separately by Discrete Wavelet Transform (DWT) for fault detection. After that, statistical features are extracted from the coefficients generated by the DWT. The most significant features are identified by the Random Forest algorithm, producing reduced features matrices, which are used to train and test four supervised machine learning techniques, namely: Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Bagged and Boosted Trees. The Frequency at the synchronous generators is found to be the best input to the classifiers based on the highest predictive accuracies attained by the classification techniques. The SVM achieved the highest average predictive accuracy of 98.4% and an F1 score of 0.98; therefore, it is selected as the best technique for fault classification. In addition, Bayesian Belief Networks (BBN) are built for fault propagation. The BBN can identify the fault location and the impacted buses through probability theory. Finally, a dynamic Fault Semantic Network (FSN) is constructed. The FSN utilizes the fault information and knowledge acquired by the Classifier and BBN for causes and consequences analysis and repair actions.

Keywords: *Fault detection; Classification; Bayesian Belief Networks; Supervised Machine Learning*

1. Introduction

Smart Grids (SGs) are very complex systems mainly intended for information and power transfer between generation units and consumers. The system complexity is increasing continuously through the incorporation of new digital and intelligent devices, equipment and technologies, which makes the grid more prone to failures such as equipment malfunction and cyber-attacks. The power system requires to be well monitored in order to avoid any disturbances to normal grid operations. Therefore, expanding our

fault knowledge is essential to mitigate the causes of the failures in the grid efficiently.

There are many different techniques used to detect, classify and locate faults in a smart grid. Various variables measured in the grid, such as the voltage difference and current amplitude, have been used for fault analysis in smart grids [1, 2]. Supervised machine learning techniques has also been implemented to classify and/or locate faults. Artificial Neural Network (ANN), Support Vector Machine (SVM), Decision Trees, Fuzzy Inference System (FIS), Markovian Models and many other techniques has been used to avoid any disturbances to grid normal operations. The supervised machine learning techniques are sometimes fused with a signal processing technique, which

* Corresponding author:

E-mail: ali.alalili@ku.ac.ae

© 2016 International Association for Sharing Knowledge and Sustainability

DOI: 10.5383/ijtee.18.01.006

increases the model ability to analyze the rapid changes occurring in the system. Aslan combined Discrete Fourier Transform (DFT), a signal processing technique that transforms the signals from the time domain into phasors in the frequency domain and an ANN for fault classification [3]. Another signal operator called Teager Energy Operator (TEO) has been integrated with FISs for fault detection and faulty phases identification [4]. “The TEO is a non-linear signal operator that could be used for extracting the energy of a signal” [4]. Moreover, a hybrid technique based on a moving average filter, a low pass filter that calculates the average of a signal, and a Random Forest (RF) has been employed for fault classification [5]. However, these signal processing techniques usually miss the rapid fluctuations occurring in power systems, which the Wavelet Analysis (WA) can detect at the exact time of occurrence; due to its capability of detecting even the rapid small changes (transients) usually occurring in power systems [6]. WA transforms a signal from the time domain to the time-frequency domain through the dilation and compression of a wavelet, called the mother wavelet, which decomposes the signal into approximate and detail coefficients. Wavelet Analysis, specifically the Discrete Wavelet Transform (DWT), has been widely used for fault detection due to its capability to reduce the volume of the input data better than the other types of WA [7-14].

However, there remains a number of research gaps in fault analysis in smart grids. The existing work focused mainly on fault detection, classification and location. The proposed techniques did not perform any root cause analysis and is incapable of suggesting remedial actions. This paper proposes a new intelligent technique for probabilistic fault propagation analysis, causes and consequences identification and repair actions. The technique starts with fault detection and classification through Discrete Wavelet Transform and four supervised machine learning techniques, namely: Support Vector Machine, K-Nearest Neighbors, Bagged and Boosted Trees. The fault information provided by Wavelet analysis and the Classifiers is utilized for fault propagation performed by Bayesian Belief Networks (BBNs). Both the Classifier and BBN provide the dynamic Fault Semantic Network (FSN) with the information and knowledge required to identify the fault causes and consequences, and finally, suggest remedial actions.

The paper is divided into three sections: Methodology, Results and Discussion, and Conclusion. The theory behind each selection and step performed in the suggested analysis is explained in the Methodology section. An analysis of the attained results is provided in Results and Discussion section followed by the main findings and recommendations in the Conclusion.

2. Methodology

2.1. Transmission Grid System Simulation

The study begins with modelling a smart grid on EMTP-RV software. EMTP-RV is a simulation and analysis software for power system transients. This computer program is used for the simulation of electromagnetic, electromechanical and control systems transients in electrical power systems. An IEEE 39-bus benchmark [15] is built for the fault analysis, which consists of 39 buses, 10 synchronous generators, 34 transmission lines, 12 transformers and 19 Three-Phase loads. The system is operating at four voltage levels: the generation is done at 20 kV, the transmission is done at 500 kV with the exception of Bus 12 and 20 which are rated at 25 kV and 300 kV, respectively. The total

generation capacity is 6250 MW, and the total active and reactive loads are 6150 MW and 1800 MVar. Static loads are assumed in the model. The system is regulated close to 60 Hz and the voltage levels in the grid are maintained between 0.95 and 1.05 of the nominal grid rating under normal conditions. **Error! Reference source not found.** 1 presents the 500 kV New-England Transmission Grid network considered in the study, i.e., the IEEE 39-bus system. The simulation is run for one second where all faults are induced at 0.5 seconds and lasted for 0.1 seconds.

Data were collected for normal and faulty conditions by inducing symmetric and asymmetric short circuit faults including Line-to-Ground (LG), Double-Line-to-ground (LLG), Triple-Line-To-Ground (LLLG), Line -to -Line (LL) and Triple Line to Line (LLL) faults. The faults are induced at the transmission buses and in the middle of the transmission lines. The data collected from the simulated model are the voltage measurements at all the buses and generators’ current and frequency measurements for the synchronous generators. The three phase voltage measurements at four buses under normal conditions are presented in Figure 2. When a LG fault in phase *a* is induced at bus 3, the voltage for phase *a* drops to zero and affects the other buses according to how close the buses are to the faulty bus, which could be clearly seen in Figure 3.

2.2. Fault Detection

To detect the induced faults in the grid, a signal processing technique is used to analyze the output of the grid simulation. The signal processing technique is Wavelet Analysis (WA), which is a multi-resolution analysis that transforms the signals in each dataset from the time into time-frequency domain.

In wavelet transform, a wavelet, derived from the mother wavelet (MW), is dilated and translated over the input signal to generate approximate and detail coefficients. Discrete Wavelet Transform (DWT), is utilized for fault detection and feature extraction. Figure 4 illustrates the decomposition process, where low and high pass filters produce approximate and detail coefficients. The approximate coefficients continue to break down into other sets of approximate and detail coefficients as the level of decomposition increases.

The wavelets used in this analysis ($\psi_{a,b}(t)$) are derived from a mother wavelet ($\psi(t)$) by Equation (1):

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad (1)$$

where *a* and *b* are the wavelet scale and translation over the signal. In DWT, the scaling and shifting parameters are discretized and set to powers of 2: $a = 2^j$, $b = 2^j \times k$ [7]. The wavelet equation changes to Equation (2).

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t}{2^j} - k\right) \quad (2)$$

The parameters *j* and *k* represent the level of decomposition and location. The detail coefficients that will be utilized in the fault classifier are derived by [7]:

$$d_{j,k} = \int_{-\infty}^{\infty} x(t) \psi_{j,k}^*(t) dt = \langle x(t), \psi_{j,k}(t) \rangle \quad (3)$$

where $x(t)$ is the input signal in the time domain and $\psi_{j,k}^*(t)$ is the complex conjugate of $\psi_{j,k}(t)$.

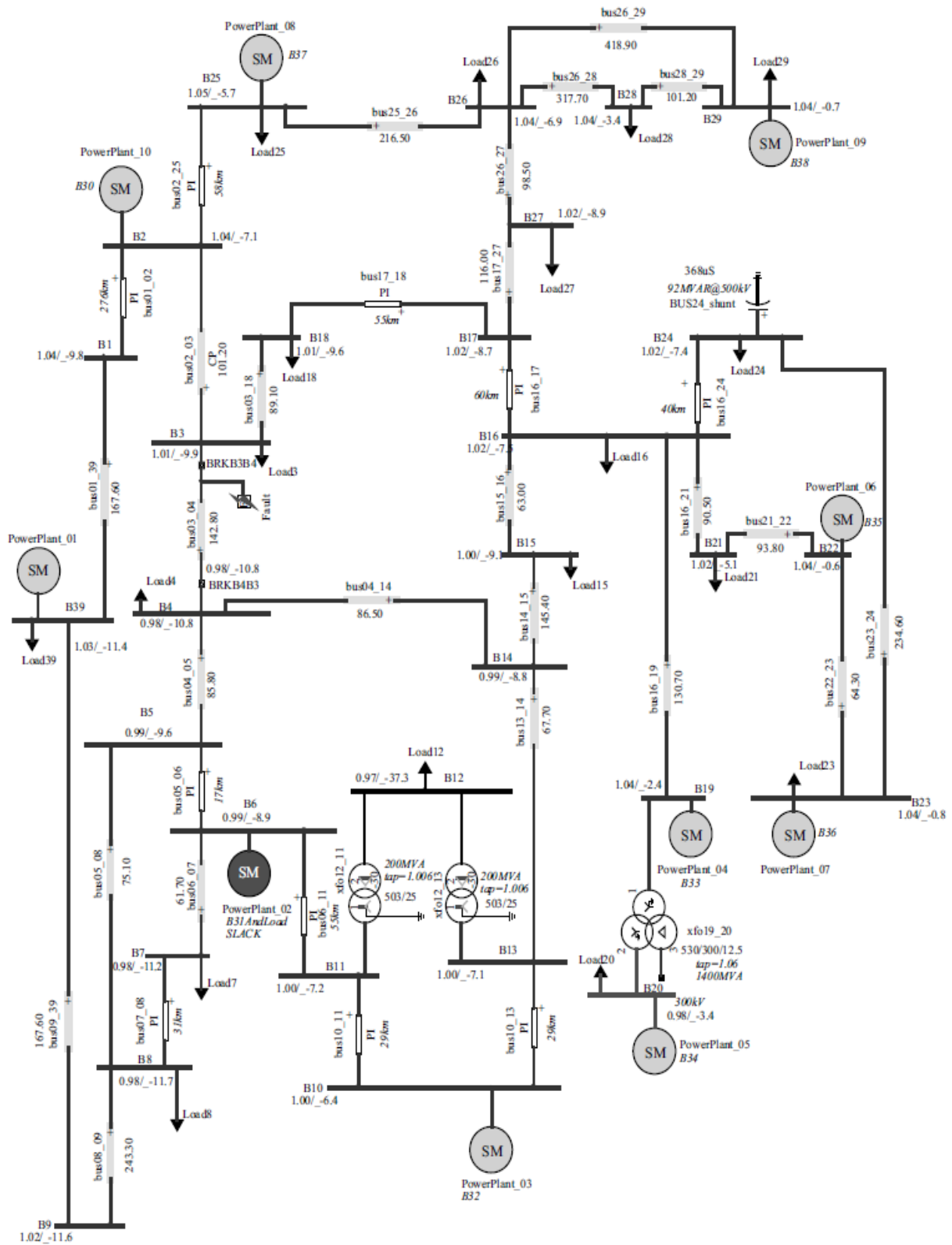


Fig. 1. IEEE-39 Bus Transmission Grid System [15]

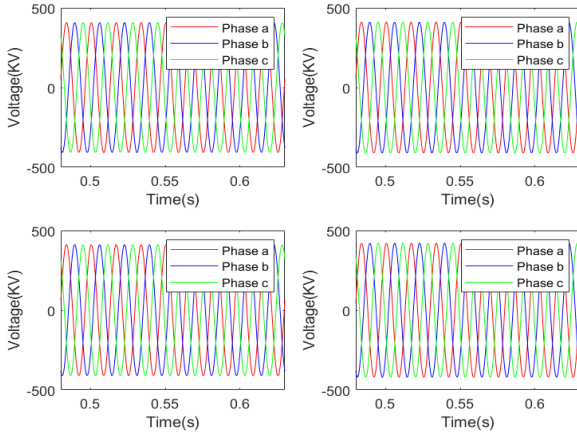


Fig. 2. Three-Phase Bus Voltage at Bus 3, 17, 27 and 39 under normal conditions

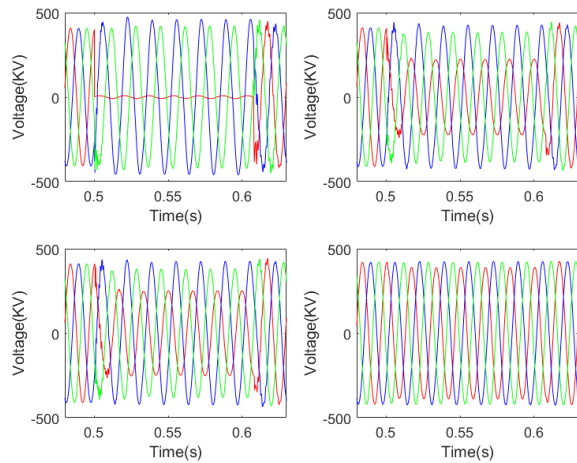


Fig. 3. Three-Phase Bus Voltage at Bus 3, 17, 27 and 39 when a LG fault in Phase (a) is induced at Bus 3

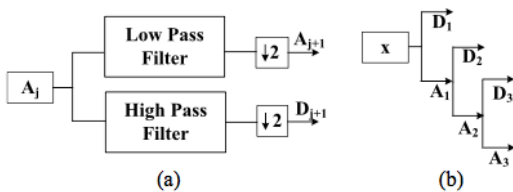


Fig. 4. (a) Decomposition into detail and approximate coefficients (b) Decomposition of a signal 'x' up to level 3 [16]

The DWT multi-resolution signal decomposition is mathematically expressed by Equation (4):

$$s(t) = \sum_k a_{M,k} \frac{1}{\sqrt{2^M}} \varphi\left(\frac{t}{2^M} - k\right) + \sum_j \sum_k d_{j,k} \frac{1}{\sqrt{2^j}} \psi\left(\frac{t}{2^j} - k\right) \triangleq A_M(t) + \sum_j D_j(t) \quad (4)$$

where $a_{M,k}$ are the approximate coefficients at level M and $\varphi(t)$ is the scaling function [10]. The signal in the time domain, $s(t)$, is decomposed through DWT into approximate coefficients, $A_M(t)$, and a series of detail coefficients, $\sum_j D_j(t)$, at level M [10]. There are many types of mother wavelets that are utilized for various purposes. The Daubechies (Db) wavelet family is known for its high effectiveness in detecting faults in power systems [17]. Many studies have used Db4 as a mother wavelet for their analysis, which has shown good results in terms of accuracy and time [7-10, 12-14]. Therefore, Db4 is employed in this research for fault detection. Another important parameter

that needs to be selected carefully when performing wavelet analysis, is the level of decomposition. To make sure all the fault transients are captured and included in the analysis, the maximum level of decomposition is computed by Equation (5).

$$L = \left\lceil \log_2 \frac{N}{F-1} \right\rceil \quad (5)$$

where L , N and F represent the maximum level of decomposition, length of the input signal and the filter size of the MW [10]. The maximum level of decomposition is found to be 7.

2.3. Feature Extraction

The performance of the classifiers depends on the quality of the input features extracted from the coefficients generated by DWT. Therefore, the input features should be capable of representing the system condition accurately. According to Yu et al., the following features are effective in fault detection and classification [10]:

- The maximum detail coefficient
- The minimum detail coefficient
- The mean of detail coefficients
- The standard deviation of detail coefficients
- The skewness of the detail coefficients
- The energy of the detail coefficients

The entropy of the DWT coefficients has also been used by Adewole et al. [7]. For this work, entropy will be extracted from the gathered data besides the list of features suggested by Yu et al. [10]. In addition, the following statistical features are considered:

- The kurtosis of the detail coefficients
- The median of the detail coefficients
- The Root Mean Square (RMS) of the detail coefficients

These ten features are extracted from the detail coefficients from level one to seven and the approximate coefficients at level seven. Table 1 summarizes all the important parameters selected for feature extraction.

Table 1: Feature Extraction Parameters

Input	1. Three-Phase Voltage at all Transmission Buses 2. Three-Phase Currents at Generator Terminals 3. Three-Phase Transmission Line Currents 4. Frequency of Synchronous Generators and at Grid Interconnection
Mother Wavelet	Db4
Level of Decomposition	1-7 (Detail Coefficients), 7 (Approximate Coefficients)
Extracted Features	1. The maximum coefficients 2. The minimum coefficients 3. The mean of Coefficients 4. The standard deviation of coefficients 5. The skewness of the coefficients 6. The energy of the coefficients 7. The log energy entropy of the coefficients 8. The Kurtosis of the coefficients 9. The median of the coefficients 10. The Root Mean Square (RMS) of the coefficients

2.4. Sensitivity Analysis

After extracting the features from the input signals, a supervised machine learning technique is used to select the most important features for fault classification. One of the most popular techniques employed for feature ranking is the Random Forest (RF). The RF is an ensemble of Decision Trees (DTs) where multiple de-correlated trees are used to find the most probable output class. At the root node of each tree, a bootstrap (sampling with replacement) of the extracted features is input to each DT to find out the fault type through a series of yes/no questions. The features at each node that achieves the greatest reduction in Gini impurity split the data into internal or leaf nodes. Gini impurity is the probability of a feature causing output misclassification, which is calculated by Equation (6).

$$I_G(n) = 1 - \sum_{i=1}^J (p_i)^2 \quad (6)$$

where n , J and p_i represent the node, number of classes at the node and the fraction of the input data present in each class. The features producing low reductions in Gini Impurity receive low rankings and are removed from the input datasets.

2.5. Fault Classification

The Reduced Features Matrices generated by the RF from each input dataset is used to train and test four different classifiers. The utilized supervised machine learning techniques are the Support Vector Machine (SVM), Bagged and Boosted Trees and K-Nearest Neighbors (KNN).

2.5.1. Support Vector Machine

Support Vector Machines can classify faults by first mapping the data in non-linear patterns into high dimensional space, where the data becomes linearly separable, through a kernel function [18]. A hyperplane is then designed to divide the data into different classes. This hyperplane is identified by finding the maximum margin from all the data in the different classes [19]. The margin is the sum of the minimum distances between the data in all the classes and the hyperplane. Quadratic and Cubic SVMs are used for fault classification in this study.

2.5.2. Bagged Trees

Two types of ensembles of DTs are used to classify the faults in the system, which are the Bagged and Boosted Trees. Bagged Trees are similar to the RF, but differs in the way each model select the random sample data at each tree. In RF, each DT receives a random sample of features, while in Bagged Trees, random samples of observations including all the extracted features are utilized as inputs. Table 2 shows the Bagged Trees parameters selected for each input dataset in the analysis.

Table 2: Bagged Trees Selected Parameters

Input	Parameters
Frequency of synchronous generators	Max Number of Splits: 5,463 Number of learners: 30 Learning Rate: 0.1
Three-Phase Voltage at all Transmission Buses	Max Number of Splits: 57,371 Number of learners: 30 Learning Rate: 0.1
Three-Phase Voltage at all Buses	Max Number of Splits: 74,591 Number of learners: 60 Learning Rate: 0.1
Three-Phase Currents at Generator Terminals	Max Number of Splits: 3000 Number of learners: 60

	Learning Rate: 0.1
Three-Phase Transmission Line Currents	Max Number of Splits: 69,665 Number of learners: 40 Learning Rate: 0.1

2.5.3. Boosted Trees

Boosted trees are used for fault classification, which differs from the Bagged trees in DTs arrangement and the model mechanism for classifying faults. DTs in a Boosted Tree are arranged in series, where weights of input samples and outputs/predictors are adjusted at each tree by learning from the errors produced by the previous tree. The parameters of Boosted Trees employed for fault classification of each input are presented in Table 3.

Table 3: Boosted Trees Selected Parameters

Input	Parameters
Frequency of synchronous generators	Max. Number of splits: 100 Number of Learners: 70 Learning rate: 1
Three-Phase Voltage at all Transmission Buses	Max Number of Splits: 100 Number of learners: 70 Learning Rate: 0.1
Three-Phase Voltage at all Buses	Max Number of Splits: 100 Number of learners: 70 Learning Rate: 0.1
Three-Phase Currents at Generator Terminals	Max Number of Splits: 100 Number of learners: 60 Learning Rate: 0.1
Three-Phase Transmission Line Currents	Max Number of Splits: 100 Number of learners: 70 Learning Rate: 0.1

2.5.4. K-Nearest Neighbors

KNNs are built to train and test the data for fault classification. KNN classify the input data into different classes by the majority vote of its neighbors. The input is mapped into an output by the most common output among its k nearest neighbors measured by a distance function [20]. The KNN used in this work utilizes correlation distance function and a squared inverse distance weight. Different numbers of neighbors are selected for each input depending on the nature of each input dataset, **Error! Reference source not found.**

Table 4: Number of Neighbors in KNNs

Input	Parameters
Frequency of synchronous generators	20
Three-Phase Voltage at all Transmission Buses	10
Three-Phase Voltage at all Buses	10
Three-Phase Currents at Generator Terminals	10
Three-Phase Transmission Line Currents	10

All of the parameters are selected through trial and error. The outputs of each classifier are represented by numerical values that indicate normal conditions and fault types, Table 5.

Table 5: Outputs/Predictors of all Datasets

Class Output	Representation
0	No Fault Exists (Normal)
1	LG(a)
2	LG(b)

3	LG(c)
4	LL (ab)
5	LL (ac)
6	LL (bc)
7	LLG (ab)
8	LLG (ac)
9	LLG (bc)
10	LLL
11	LLG

2.5.5. Training and Testing

The classification training has been performed with 10-folds cross validation. The total observations in each dataset, Table 6, were divided into 10 equal folds, and only one-fold was used for assessing each classifier at a time. This process is repeated for 10 times until each unique group has been used as a test set, which gives a better indication of the model performance on unseen data compared to the hold-out validation. In hold-out validation, the data is split into training and testing sets; therefore, the model is tested by a fixed sample of data.

Table 6: Total Number of Observations

Reduced Feature Matrices Extracted From	Number of Observations
Frequency of synchronous generators	5,465
Three-Phase Voltage at all Transmission Buses	57,373
Three-Phase Voltage at all Buses	74,593
Three-Phase Currents at Generator Terminals	20,491
Three-Phase Transmission Line Currents	69,667

2.6. Fault Propagation

In order to eliminate the occurring faults in the grid efficiently, the propagation behaviors of the faults should be studied first. Therefore, a probabilistic graph, Bayesian Belief Network (BBN), is designed specifically for a fault to show how a fault is spread throughout the system. The BBN is a probabilistic graph that represents conditional dependencies between the process variables using a Directed Acyclic Graph (DAG). The model consists of nodes representing different process variables such as buses' voltages and currents. These nodes are connected through arcs that represent the probabilistic dependencies. The probabilities are calculated in the BBN Inference Engine through Equation (7). This Equation is used by the BBN to discover the posterior probabilities between the variables for fault propagation.

$$P(X_1, \dots, X_N) = \prod_{i=1}^N P(X_i | Parents(X_i)) \tag{7}$$

The BBN is selected for determining the impact of a single fault on the whole system due to its capability of modelling and reasoning uncertainty in the system. Furthermore, this probabilistic graph can consider both subjective probabilities, usually obtained from experts in the field, and probabilities based on observed data. The BBN can also revise the probabilities through new data.

To construct the BBN and identify the prior and conditional probabilities, a learning algorithm is required [21]. The K2

algorithm is chosen to build the BBN. The K2 algorithm is a greedy search algorithm that build the BBN structure and learn the parameters through learning input data. This learning algorithm selects the BBN structure that maximizes prior probabilities at the parent's nodes using observed data [21].

After designing the BBN structure, the Junction Tree Algorithm transforms the BBN into an undirected Acyclic Graph to allow efficient inferences for fault information and knowledge acquisition performed by the BBN inference engine. The Junction Tree Algorithm transforms the BBN into an appropriate data structure while keeping the joint probabilities fixed, and ensuring that the marginal probabilities can be computed. The BBN learning framework is demonstrated in Figure 5. The Inferences done by the BBN can provide the most probable fault location and the probabilities of observing disturbances in the surrounding buses.

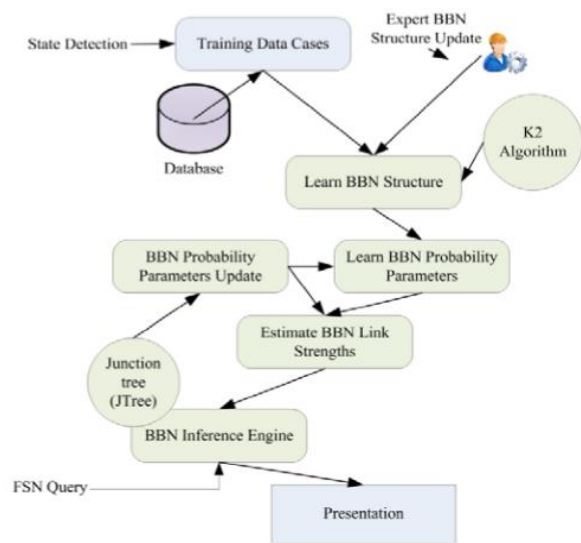


Fig. 5. BBN learning framework [22]

2.7. Normal Grid Operations Restoration

The fault information and knowledge attained by the Classifiers and BBNs are utilized by the dynamic Fault Semantic Network to identify the causes and consequences of the faults and suggest repair actions. Semantic Networks are networks that reveal the interactions between objects [22]. The structure of Semantic Networks is illustrated in Figure 6. The FSN is a way of representing our fault knowledge by the relationships between the objects, which are the faults/causes/consequences. The FSN models and describes faults by symptoms, enablers, variables, causes, consequences and remedial actions [22]. The knowledge represented by the FSN could be updated either manually or through observed data, which makes the technique dynamic and adaptable to any changes in the system.

Normally, when a fault occurs in the electric grid, either the protection system consisting of electrical devices such circuit breakers are initialized or the grid operator would take the necessary remedial actions based on the Standard Operating Procedures (SOPs). The dynamic FSN proposed in this study would provide the operator with any fault causes and consequences, and suggest repair actions that are set by experts and updated through observed data. Therefore, the Grid Operator do not need to get back to the fixed SOPs for mitigating actions required to restore normal grid operations. Furthermore, the FSN

is adaptable to any changes in the system, which makes it a better reference for fault elimination. The proposed method can accelerate the remedial processes with limited resources.

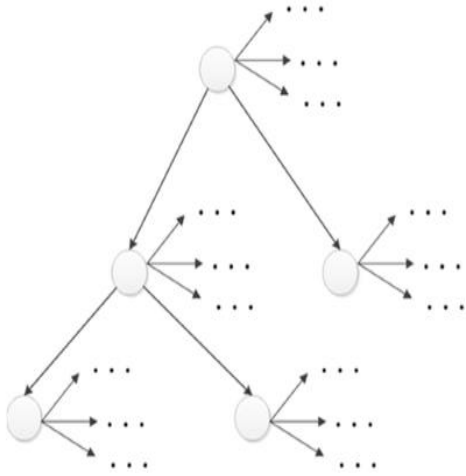


Fig. 6. Semantic Network Structure [22]

3. Results and Discussion

3.1. Fault Detection

The first step was to decompose the input signals into approximate and detail coefficients through DWT on MATLAB. The voltage signal decomposition into detail and approximate coefficients from level one to seven is shown in Fig. 7. Detail Coefficients at levels one to five were capable of detecting the fault at the exact time of occurrence through large deviations.

3.1.1. Feature Extraction and Sensitivity Analysis

Features were extracted from the detail and approximate coefficients for fault classification. Sensitivity analysis is then performed by the RF on Salford Predictive Modeler (SPM). Fig. 7. DWT Level 7 decomposition results of the Phase (a) Voltage signal at Bus 2 when a LG (a) fault is induced at Bus 2

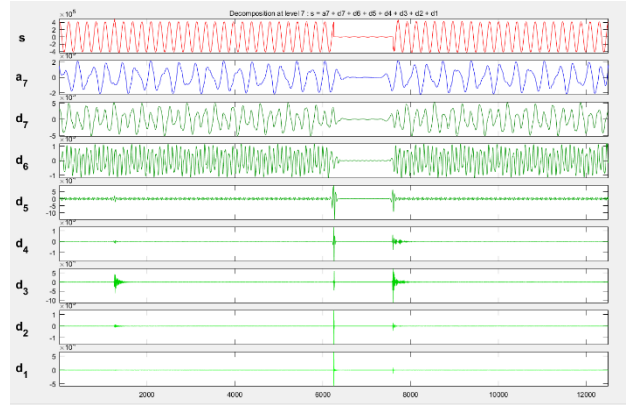


Fig. 7. DWT Level 7 decomposition results of the Phase (a) Voltage signal at Bus 2 when a LG (a) fault is induced at Bus 2

Table 7: Cut-off relevancies for the input datasets

Input	Cut-off Relevancy (%)
Frequency of synchronous generators	19
Three-Phase Voltage at all Transmission Buses	5
Three-Phase Voltage at all Buses	16
Three-Phase Currents at Generator Terminals	6
Three-Phase Transmission Line Currents	11

3.1.2. Fault Classification

The reduced features matrices generated by the Random Forest are used to train and test the four classifiers. Figures 8-12 show the predictive accuracies, which are calculated by Equation (8), achieved by each classifier using the original and reduced features matrices for each input. The predictive accuracies attained by the reduced matrices are greater than the accuracies obtained by the original matrices for all inputs.

$$Predictive\ Accuracy = \frac{Correct\ Predictions\ for\ the\ testing\ sample}{Total\ Number\ of\ Testing\ Observations} * 100 \quad (8)$$

Table 7: Cut-off relevancies for the input datasets

Input	Cut-off Relevancy (%)
Frequency of synchronous generators	19
Three-Phase Voltage at all Transmission Buses	5
Three-Phase Voltage at all Buses	16
Three-Phase Currents at Generator Terminals	6
Three-Phase Transmission Line Currents	11

lists the cut-off relevancies values for each input. Features with relevancies lower than the minimum percentage were eliminated to improve classification accuracy.

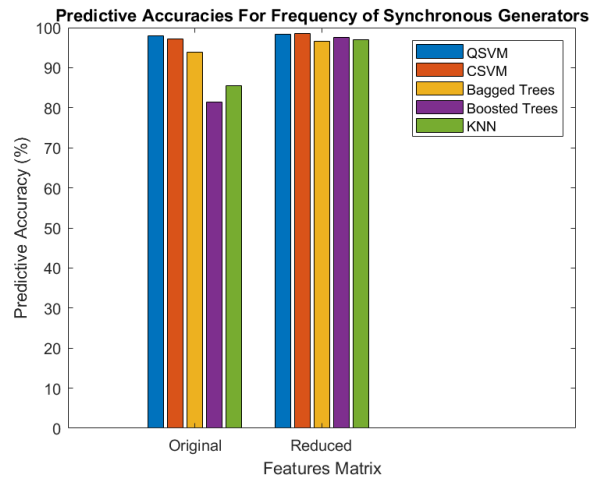


Fig. 8. Predictive accuracies for frequency of synchronous generators

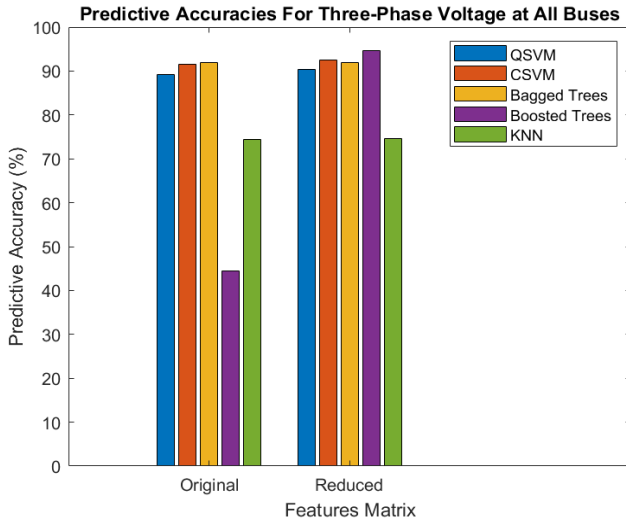


Fig. 9. Predictive accuracies for three-phase voltage at all buses

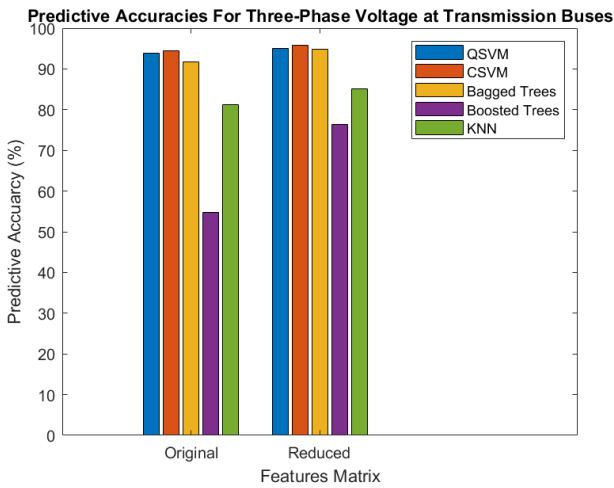


Fig. 10. Predictive accuracies for three-phase voltage at transmission buses

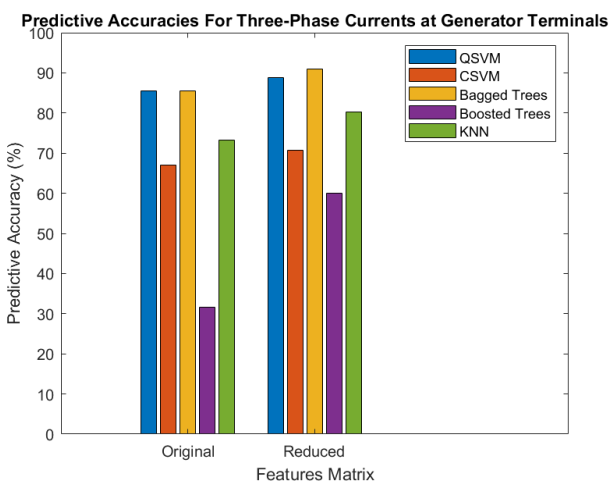


Fig. 11. Predictive accuracies for three-phase current at generators terminals

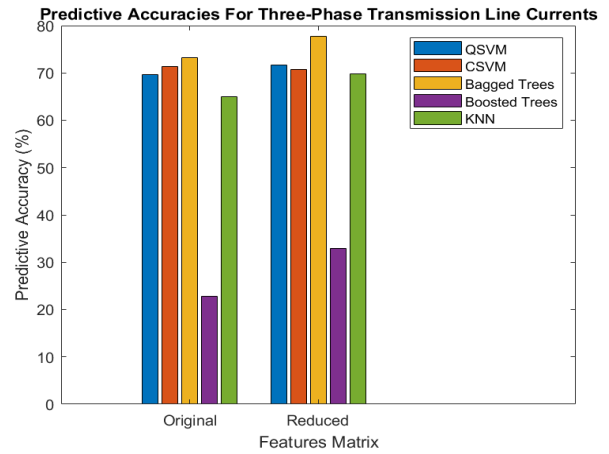


Fig. 12. Predictive accuracies for three-phase line currents

3.1.3. Best Input Selection

The predictive accuracies achieved by the classifiers using the reduced features matrices generated from each input are compared for best input selection. Figure 13 exhibits the predictive accuracies obtained from the inputs shown in Table 8. All classifiers achieved higher predictive accuracies using the frequency of synchronous generators dataset. Zhang et al., who modeled a 39-bus grid system, found that the frequency of synchronous generators data leads to high effectiveness in fault detection and classification [23]. Therefore, the frequency of synchronous generators input is used for further analysis.

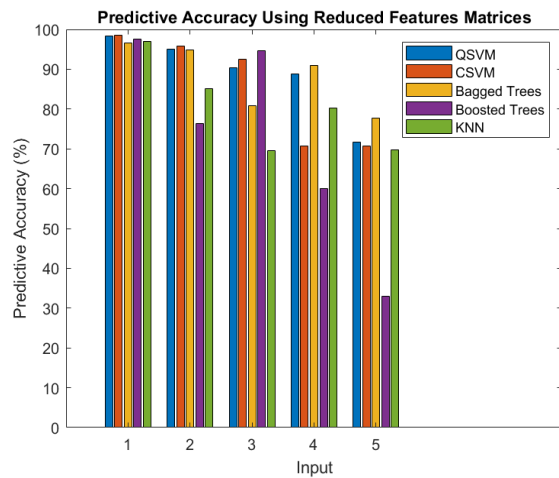


Fig. 13. Predictive accuracies using reduced features matrices for all inputs

Table 8: Input datasets from the simulated model

Input
1. Frequency of Synchronous Generators
2. Three-Phase Voltage at Transmission Buses
3. Three-Phase Voltage at all Buses
4. Three-Phase Currents at Generator Terminals
5. Three-Phase Transmission Line Currents

3.1.4. Best Classifier Selection

These predictive accuracies alone are not good indicators of the model performance in classifying faults due to the presence of imbalanced data and multiple classes/outputs. Therefore, the confusion matrix is utilized for an accurate model comparison. Due to classes imbalance present in the data, which occurs when

an unequal number of observations are used to classify each output class, and having more than two outputs, the *F1* score is calculated for best model selection. The *F1* score uses the information provided by the confusion matrix to assess the performance of the classifiers. The confusion matrix shows the model performance in classifying all the outputs through observed data. The matrix identifies the number of correctly classified and misclassified observations. The matrix can also show the rate at which the model performed the classification correctly (Positive Predicted Value) and the rate of misclassification (False Discovery Rate). The Figures below show the confusion matrices for the four supervised machine learning techniques utilized for fault classification using the frequency input.

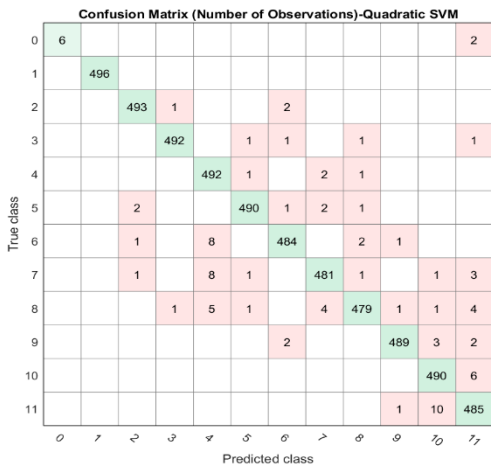


Fig. 14. Quadratic SVM confusion matrix

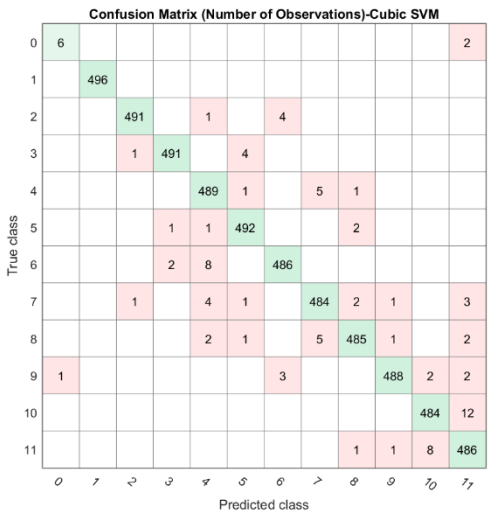


Fig. 15. Cubic SVM confusion matrix

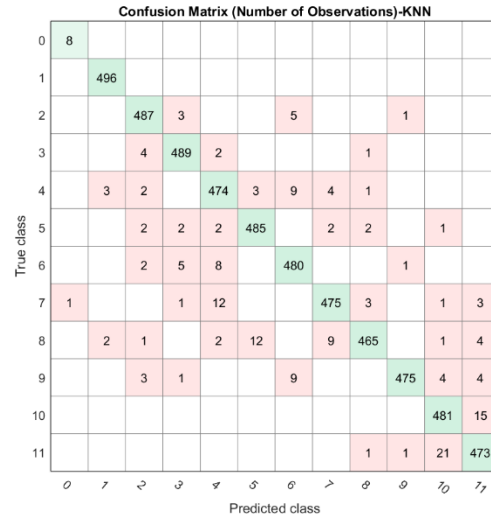


Fig. 16. KNN confusion matrix

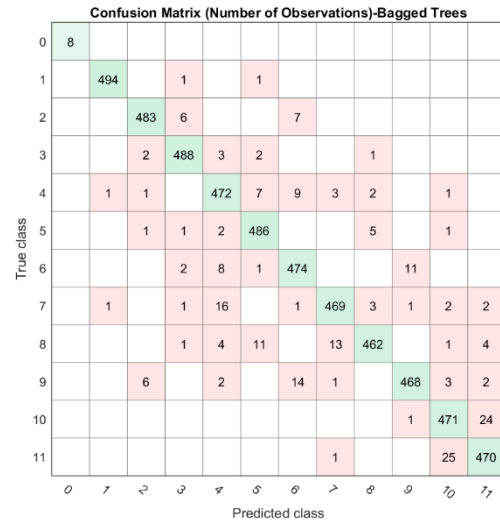


Fig. 17. Bagged Trees confusion matrix

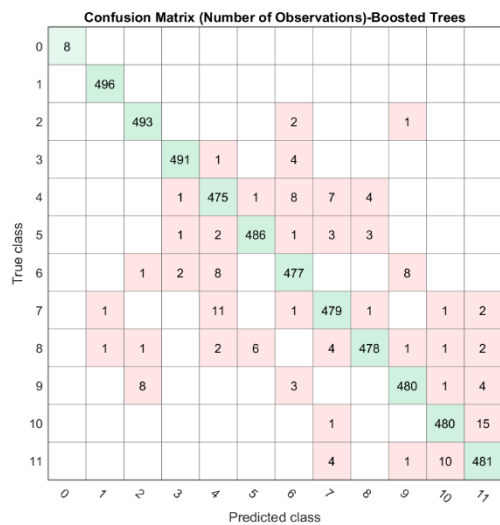


Fig. 18. Boosted Trees confusion matrix

The information contained in the confusion matrix are used to calculate the *F1* score, which is a metric used to quantify the performance of each classifier in predicting each class. The *F1* score is the harmonic mean of precision (Ratio of correctly

predicted positive observations to the total observations for a class) and recall (Ratio of correctly predicted positive observations to all the correct and incorrect predictions for each output), which is calculated as follows:

$$Precision = \frac{True\ Positive}{Total\ Observations} \quad (8)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (9)$$

$$F1\ Score = 2 \left(\frac{Precision \cdot Recall}{Precision + Recall} \right) \quad (10)$$

The Total positive and False Negative of each class/output are attained from the confusion matrix of each classifier. The True positive represents correctly identified predictions for each output, and False Negative is the incorrect predictions for each class. The True positives are the diagonal observations highlighted in green by the Confusion Matrix. Moreover, the False Negatives are the red observations presented in each column for all classes in the Confusion Matrix. Table 9 presents the calculated F1 score for the five classifiers.

Table 9: Average predictive accuracy and F1 score of each classifier

Classifier	Average Predictive Accuracy (%)	Average F1 Score
Quadratic SVM	98.4	0.98
Cubic SVM	98.4	0.97
KNN	96.8	0.97
Bagged Trees	96.2	0.97
Boosted Trees	97.4	0.97

The Quadratic SVM achieved the highest average *F1* score. Therefore, the Quadratic SVM is selected for fault classification based on the highest *F1* score and predictive accuracy attained by the model. The SVM is expected to show good performance in terms of correct classes classifications due to the SVM’s ability to map non-linear input data into a high dimensional space where the data becomes linearly separable.

3.1.5. Fault Propagation

Many Bayesian Networks are built for different short circuit faults that could discover the impact of a fault on the whole system. Moreover, any fault occurring at a bus in the network can affect the other buses at different severities depending on how close the buses are to the faulty bus. A BBN is built for a LG fault to show the way a fault can spread in the smart grid. The constructed DAG for a LG fault in phase *a* is presented in Figure 19.

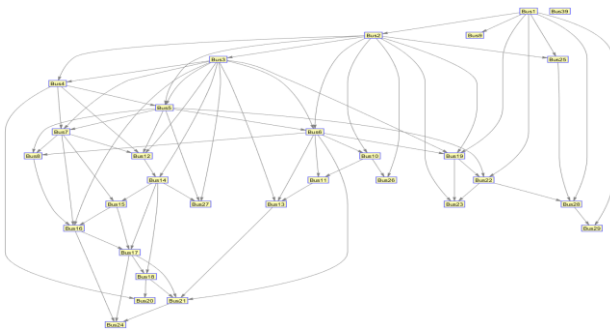


Figure 19: Constructed BBN for a LG fault in Phase (a)

After building the BBN structure and learning the necessary parameters through observed data using the K2 algorithm, the BBN is transformed into an undirected acyclic graph by the Junction Tree Algorithm for inferences held by the BBN inference Engine. The following inferences were done with different fault scenarios for fault propagation:

Fault Propagation BUS 16

- What is the probability of observing disturbance in Bus 24, given a fault on Bus 16 has occurred?
P (FB24 | FB16) = 99.6%
- What is the probability of observing disturbance in Bus 1, given a fault on Bus 16 has occurred?
P (FB1 | FB16) = 89.33%
- What is the probability of observing disturbance in Bus 29, given a fault on Bus 16 has occurred?
P (FB29 | FB16) = 93.63%
- What is the probability of observing disturbance in Bus 28, given a fault on Bus 16 has occurred?
P (FB17 | FB16) = 94.18%

Fault Propagation BUS 29

- What is the probability of observing disturbance on Bus 24, given a fault on Bus 29 has occurred?
P (FB24 | FB29) = 88.89%
- What is the probability of observing disturbance on Bus 1, given a fault on Bus 29 has occurred?
P (FB1 | FB29) = 0%
- What is the probability of observing disturbance in Bus 16, given a fault on Bus 29 has occurred?
P (FB16 | FB29) = 89.59%
- What is the probability of observing disturbance in Bus 17, given a fault on Bus 29 has occurred?
P (FB17 | FB29) = 94.13%

The inferences performed by the BBN can be verified by referring to the buses locations in the IEEE 39-bus system. The following inferences are tested by the system topology in Figure:

- What is the probability of observing disturbance in Bus 24, given a fault on Bus 16 has occurred?
P (FB24 | FB16) = 99.6%
- What is the probability of observing disturbance in Bus 1, given a fault on Bus 29 has occurred?
P (FB1 | FB29) = 0%

The involved buses in the first inference are presented in red circles, Figure 20. The two buses are close to each other, which proves that the high probability revealed by the BBN is logical. Bus 1 and Bus 29 are far away from each other as illustrated in Figure 20; therefore, Bus 1 is expected to operate normally when a fault is detected in Bus 29, which is proved by the extremely low posterior probability calculated by the BBN inference Engine. The posterior probabilities provided by the BBN will help the FSN in the causes and consequence analysis and repair actions.

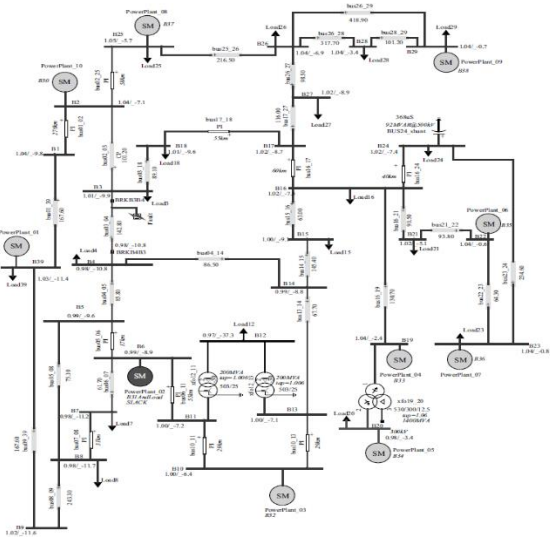


Figure 20: System Topology [15]

3.1.6. Normal Grid Operations Restoration

The dynamic FSN utilizes the fault information and knowledge obtained by the previous steps for causes and consequences analysis and repair actions. The FSN comes in the form of dynamic tables linked together as shown in Tables 10-12. The FSN presents the symptoms, hypotheses, diagnosis and repairs. Each has a weight value that determines the frequency of contribution in the analysis. The weight values presented in **Error! Reference source not found.** are initial values that need be updated either by experts or historical data. Moreover, the symptoms are the measurements sensors used to detect any abnormal behavior in the smart grid that initializes the analysis performed by the FSN. Hypotheses are the possible causes for the detected fault. The diagnosis are the sensors and features (found in separate tables linked to the main table) utilized for detecting the causes of a fault, which determine the corrective actions required to restore normal grid operations. An example is shown below, where an abnormal behavior is detected by the frequency sensor. The FSN asks the Classifier to reveal the fault type, which is found to be LG fault in phase *a*, and initializes the BBN trained for this fault type. This fault is mapped to a number of hypotheses, which are heavy wing and falling tree (H1 and H2). H1 and H2 are diagnosed by D1 and D2. On board weather stations (S1) are used to obtain the wind speed and direction (F2) to detect heavy wind. UAV drones (S5) use Light Detection and Ranging (LiDAR) feature (F4) to detect tree contact. If heavy wind and falling tree are detected in the smart grid, the FSN suggests Tripping Circuit Breakers in the Zones of Protection (R2) and Tree trimming and vegetation removal around power lines (R7) as repair actions.

Table 10: Dynamic Fault Semantic Network

Symptom	S_ID	Symptom	Weight	Semantic Network (BBN)	Fault Type	Hypotheses
	1	Frequency	0.50	BBN_LG_B	LG (a)	H1, H2
	2	Voltage	0.40			
	3	Current	0.10			
Hypothesis	H_ID	Hypothesis	Weight			Diagnosis
	1	Heavy Wind	0.31			D1
	2	Falling Tree	0.04			D2
Diagnosis	D_ID	(Sensor, Features)	Weight			Repair

		Pairs			
	1	(S1, F2)	0.17		R2
	2	(S5, F4)	0.17		R7
Repair	R_ID	Repair	Weight		
	1	Detect and identify Fault Location	0.091		
	2	Tripping Circuit Breakers in the Zones of Protection	0.091		
	3	Transmission Line Switching	0.091		
	4	Excess Load Shedding	0.091		
	5	Back-up Distributed Generation	0.091		
	6	Replacing Damaged Equipment	0.091		
	7	Tree trimming and vegetation removal around power lines	0.091		
	8	Dispatching Security Staff	0.091		
	9	Removing Unauthorized Personnel	0.091		
	10	Firewalls Activation, firmware update	0.091		
	11	Minimize the network traffic through whitelisting for approved IPs only	0.091		
Version	Last Updated	Case Quality	Success	Failure	Conditions
	27 Dec, 2019				Dust storms

Table 11: Sensors used in the diagnosis process.

Sensor ID	Name
S1	Weather Radar
S2	Current Transformer
S3	Potential Transformer
S4	Digital Frequency Relay
S5	UAV Drone Inspection

Table 12: Features used in the diagnosis process.

Feature ID	Feature Name
F1	Precipitation
F2	Wind Speed and Direction
F3	Atmospheric pressure
F4	Light Detection and Ranging (LiDAR)

4. Conclusion

An IEEE 39 bus system is simulated using EMTP-RV software package, where short circuit faults were induced at different locations. DWT successfully detected the induced faults in the smart grid model. Features matrices, extracted from the detail and approximate coefficients of the data collected from the simulated grid model, were reduced by the Random Forest

algorithms. The reduced feature matrices were used to train and test various supervised machine learning techniques (Quadratic and Cubic SVM, KNN, Bagged and Boosted Trees) for fault classification. The predictive accuracies obtained by the classifiers, using each input datasets separately, have shown that the Frequency of synchronous generators is the best input for fault classification, with a predictive accuracy of 98.4%. The Support Vector Machine is selected as the best classifier based on the high average predictive accuracy and *F1* score attained. Furthermore, the built BBN for a LG fault was capable of providing the most probable fault location and identifying the disturbed buses using probability theory. The dynamic FSN uses the fault information and knowledge utilized the classifier and BBN to identify the faults causes and consequences, and suggest repair actions.

Acknowledgement

This publication is based upon work supported by ADEK Award for Research Excellence (AARE-2017) under Award No. AARE17-176.

References

- [1] S. Brahma and A. Girgis, "Fault Location on a Transmission Line Using Synchronized Voltage Measurements," *IEEE Transactions on Power Delivery*, vol. 19, no. 4, pp. 1619–1622, 2004.
- [2] M. M. Mahfouz and M. A. El-Sayed, "Smart grid fault detection and classification with multi-distributed generation based on current signals approach," *IET Generation, Transmission & Distribution*, vol. 10, no. 16, pp. 4040–4047, 2016.
- [3] Y. Aslan, "An alternative approach to fault location on power distribution feeders with embedded remote-end power generation using artificial neural networks," *Electrical Engineering*, vol. 94, no. 3, pp. 125–134, 2011.
- [4] B. Chaitanya and A. Yadav, "An intelligent fault detection and classification scheme for distribution lines integrated with distributed generators," *Computers & Electrical Engineering*, vol. 69, pp. 28–40, 2018.
- [5] D. Patil, O. Naidu, P. Yalla, and S. Hida, "An Ensemble Machine Learning Based Fault Classification Method for Faults During Power Swing," *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, 2019.
- [6] R. Shariatinasab, B. Rahmani, and M. Akbari, *Application of Wavelet Analysis in Power Systems*. INTECH Open Access Publisher, 2012.
- [7] A. C. Adewole, R. Tzoneva, and S. Behardien, "Distribution network fault section identification and fault location using wavelet entropy and neural networks," *Applied Soft Computing*, vol. 46, pp. 296–306, 2016.
- [8] E. Koley, K. Verma, and S. Ghosh, "An improved fault detection classification and location scheme based on wavelet transform and artificial neural network for six phase transmission line using single end data only," *SpringerPlus*, vol. 4, no. 1, 2015.
- [9] N. Perera and A. D. Rajapakse, "Recognition of Fault Transients Using a Probabilistic Neural-Network Classifier," *IEEE Transactions on Power Delivery*, vol. 26, no. 1, pp. 410–419, 2011.
- [10] J. J. Q. Yu, Y. Hou, A. Y. S. Lam, and V. O. K. Li, "Intelligent Fault Detection Scheme for Microgrids With Wavelet-Based Deep Neural Networks," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1694–1703, 2019.
- [11] A. Pradhan, A. Routray, S. Pati, and D. Pradhan, "Wavelet Fuzzy Combined Approach for Fault Classification of a Series-Compensated Transmission Line," *IEEE Transactions on Power Delivery*, vol. 19, no. 4, pp. 1612–1618, 2004.
- [12] A. Khaleghi, M. O. Sadegh, M. Ghazizadeh-Ahsaei, and A. M. Rabori, "Transient Fault Area Location and Fault Classification for Distribution Systems Based on Wavelet Transform and Adaptive Neuro-Fuzzy Inference System (ANFIS)," *Advances in Electrical and Electronic Engineering*, vol. 16, no. 2, 2018.
- [13] M. Reddy and D. Mohanta, "Adaptive-neuro-fuzzy inference system approach for transmission line fault classification and location incorporating effects of power swings," *IET Generation, Transmission & Distribution*, vol. 2, no. 2, p. 235, 2008.
- [14] H. Dehghani, B. Vahidi, R. Naghizadeh, and S. Hosseini, "Power quality disturbance classification using a statistical and wavelet-based Hidden Markov Model with Dempster-Shafer algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 47, pp. 368–377, 2013.
- [15] Power system test cases for EMT-type simulation studies. 2018. Paris: CIGRÉ, pp.19 - 38.
- [16] R. Shariatinasab, B. Rahmani, and M. Akbari, *Application of Wavelet Analysis in Power Systems*. INTECH Open Access Publisher, 2012.
- [17] N. Perera and A. D. Rajapakse, "Recognition of Fault Transients Using a Probabilistic Neural-Network Classifier," *IEEE Transactions on Power Delivery*, vol. 26, no. 1, pp. 410–419, 2011.
- [18] S. S. Gururajapathy, H. Mokhlis, and H. A. Illias, "Fault location and detection techniques in power distribution systems with distributed generation: A review," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 949–958, 2017.
- [19] K. Chen, C. Huang, and J. He, "Fault detection, classification and location for transmission lines and distribution systems: a review on the methods," *High Voltage*, vol. 1, no. 1, pp. 25–33, 2016.
- [20] Cigdem, O. and Demirel, H. "Performance analysis of different classification algorithms using different feature selection methods on Parkinson's disease detection". *Journal of Neuroscience Methods*, 309, pp.81-90, 2018
- [21] G. F. Cooper and E. Herskovits, "A Bayesian method for the induction of probabilistic networks from data," *Machine Learning*, vol. 9, pp. 309- 347, 1992.
- [22] S. Hussain, A. Hossein, and H. A. Gabbar, "Tuning of fault semantic network using Bayesian theory for probabilistic fault diagnosis in process industry," *2013 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE)*, 2013.

[23] H. Jiang, J. J. Zhang, W. Gao, and Z. Wu, "Fault Detection, Identification, and Location in Smart Grid Based on Data-Driven Computational Methods," *IEEE Transactions on Smart Grid*, vol. 5, no. 6, pp. 2947–2956, 2014.