4
Hibert-Huang Transform Based Algorithm

In this chapter, Hilbert-Huang Transform is utilized for fault diagnosis in the presence of DG. It comprises of the principle of HHT, threshold based algorithm without DG and machine learning approach with DG and noise. As EMD is already discussed in the previous chapter, only Hilbert Transform is presented here along with mathematical analysis of Support Vector Machine.

4.1 HILBERT TRANSFORM

The Hilbert transform of a continuous signal $S(t)$ is the convolution of the signal with $\frac{1}{t}$, which is given as $H(t) = S(t) * (1/t)$ where, $*$ indicates convolution. It is computed conventionally as

$$H(t) = \int_{-\infty}^{\infty} \frac{S(\tau)}{t-\tau} d\tau \tag{4.1}$$

The continuous Hilbert transform involves a $\frac{\pi}{2}$ radian phase shift (for positive frequencies only) in the frequency domain. The transfer function $H(\omega)$ of discrete Hilbert transform is defined in equation (4.2).

$$H(\omega) = \begin{cases} +j, & 0 < \omega < \pi. \\ 0, & \omega = 0 \text{ and } \omega = \pi. \\ -j, & -\pi < \omega < 0. \end{cases} \tag{4.2}$$

The discrete Hilbert transform can be computed using a combination of DFT and its transfer function. Firstly, the DFT of the given signal is computed. Secondly, the obtained DFT coefficients are multiplied by the masking coefficients M_{even} and M_{odd} depending on the odd or even sequence. M_{even}=$\{0, j, j, \cdots, 0, -j, -j, \cdots, -j\}$ and M_{odd}=$\{0, j, j, \cdots, j, -j, -j, \cdots, -j\}$. Subsequently, inverse DFT is applied on the computed signal after multiplication to obtain the time sequence [Shukla *et al.*, 2009].

A real data sequence when transformed into an analytic signal with the imaginary part equal to the Hilbert transform of the real data returns instantaneous amplitude, phase and frequency. Thus, by applying HT on first level residue R_1(t) we obtain,

$$\begin{aligned} Z(t) &= R_1(t) + j(H(t)) \\ &= Amp(t)e^{j\theta(t)} \end{aligned} \tag{4.3}$$

where,
$Z(t)$ is the analytic function.
$H(t)$ is the Hilbert transform of the residue R_1(t).
$Amp(t)$ is the instantaneous amplitude.

$$Amp(t) = [R_1(t)^2 + H(t)^2]^{1/2} \tag{4.4}$$

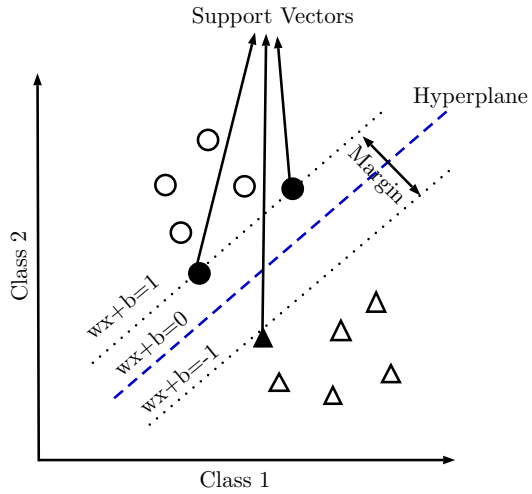


Figure 4.1: Support Vector Machine.

$\theta(t)$ is the instantaneous phase.

$$\theta(t) = tan^{-1}(H(t)/R_1(t)) \tag{4.5}$$

$\omega(t)$ is the instantaneous frequency derived as,

$$\omega(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt} \tag{4.6}$$

4.2 SUPPORT VECTOR MACHINE

Support vector machines gained prominence among the supervised learning algorithms due to their capability to draw non-linear boundaries to separate the classes in variable space. It solves the quadratic optimization problem even in high-dimensional feature space. The underlying concept of binary SVM lies in identifying the hyperplane that separates two classes. There may exist an infinite number of possible hyperplanes that separate the two classes but the optimal solution is obtained by using only a portion or subset of the training samples known as support vectors. The superiority of SVM algorithm over ANN lies in determining the support vectors from the training samples contrary to the number of hidden layers assigned by trial and error in the case of ANN. Also SVM performs better than ANN with less number of training samples. The distance between the support vectors also known as the margin (M_s) is increased to obtain the optimal hyperplane or decision boundary. SVM maximizes M_s by introducing a cost parameter which determines the penalty for support vector misclassification. In such cases where the classes are not linearly separable, SVM transforms the input variables using a kernel function to a high dimensional feature space to obtain a linear hyperplane.

It can be observed from Fig. 4.1, that two classes class 1 and class 0 are separated by an optimum hyperplane. Let the total T SVM training samples be $\{x_p, y_p\}$ where, $p = 1$ to T, $x_p \in R^n$ represents independent variables and $y_p \in (-1, +1)$. A hyper plane is determined by

$$f(x) \Rightarrow wx + b = 0 \tag{4.7}$$

where, w is vector of coefficients and b is a constant. The perpendicular distance from hyperplane to origin is $\frac{|b|}{||w||}$. If the output is $y_p = +1$, then the training samples satisfy the following constraint,

$$y(wx + b) \geq 1 \tag{4.8}$$

The optimal hyperplane is obtained when the margin is the maximum. The Lagrange function is used to solve the constraint optimization problem for maximum margin. The support vectors are identified if Lagrange is not equal to 0. The optimization problem is given by:

$$Maximize: \sum_{p=1}^{l} \alpha_p - \frac{1}{2} \sum_{p=1}^{l} \sum_{q=1}^{l} \alpha_p \alpha_q y_p y_q (\phi(x_p)\phi(x_q)) \tag{4.9}$$

Subject to

$$\sum_{p=1}^{l} \alpha_p y_p = 0 \qquad and \qquad 0 \leq \alpha_p \leq C \qquad for \qquad p = 1, 2, ...T \tag{4.10}$$

where, $\phi(x_p)\phi(x_q)$ is the kernel function and is represented as $K_f(x_p x_q)$ and C is the cost parameter. The value of kernel function is equivalent to the inner product of two vectors x_p and x_q in the feature space, $\phi(x_p)$ and $\phi(x_q)$. The appropriate selection of kernel functions optimizes the performance of SVM. In this research work, the radial basis kernel function is utilized. The various types of kernel functions used by SVM are as follows:

- Linear Kernel K_f (x, y) = x. y
- Polynomial Kernel K_f (x, y) = (γ (x. y) + γ)^{degree} $\quad with \quad \gamma > 0$
- Radial basis kernel K_f (x, y) = exp($-\gamma$ ||x. y||^2) $\qquad with \quad \gamma > 0$
- Sigmoid kernel K_f (x, y) = tanh(γ (x. y) + γ)

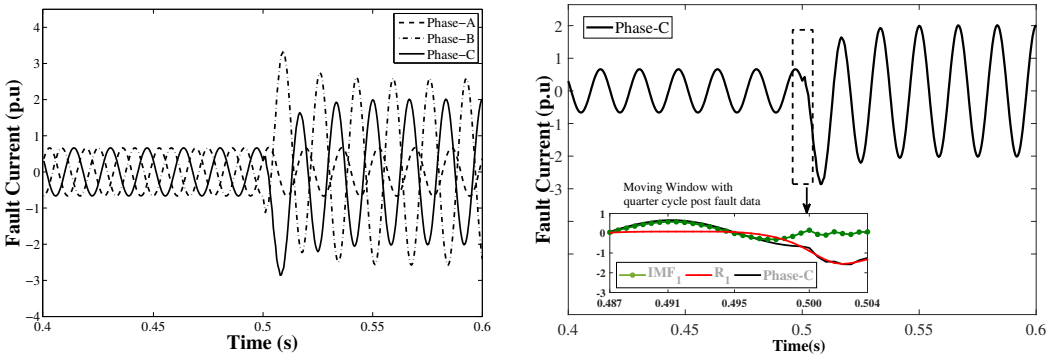
where, γ is the kernel parameter. By solving the dual Lagrangian, the nonlinear decision function is obtained as $f(x) = sgn(\sum_p^1 \alpha_p y_p k(x, x_p) + b)$
where, x is the test vector, b is the constant found from the primal constraints and is computed by $\alpha(y_p(w.x_p + b)-1) = 0$, $p = 1, ..., T$, such that α_p is not zero and sgn is a signal function [Gururajapathy *et al.*, 2017].

4.3 TEST SYSTEM CONSIDERED

The proposed algorithm is tested on both the IEEE 13 bus and 34 bus systems using the same configurations as described in the previous chapter without any change. The case studies involve faults at all the buses of IEEE 13 bus system expect measuring bus and seven fault locations in IEEE 34 bus system. All the fault types at various incidence angles, fault resistances, non faulty transients in the presence of noise upto 20 dB SNR are considered.

4.4 PROPOSED ALGORITHM WITHOUT DG

The threshold based algorithm is discussed here with its application on IEEE 13 bus system, all the measurements are in per unit values. Three phase currents are measured at substation bus (650) while all the faults are simulated at 0.5 s. Only phase-C current is chosen for demonstration of the proposed algorithm and same method is adopted for other phase currents as well. A two phase to ground fault (BCG) at bus 611 is simulated to obtain three-phase currents as shown in Fig. 4.2(a). The phase-C current is decomposed using EMD over a moving window of one cycle to obtain first level IMF_1 and R_1. The decomposition of the signal enclosed by the moving window with quarter cycle post-fault data is magnified in Fig. 4.2(b). Then residue R_1 is chosen for further analysis to derive the instantaneous frequency. The total harmonic distortion (THD) obtained using Fourier transform on IMF_1 and R_1 over the moving window is shown in Fig. 4.3. It clearly



(a) (b)

Figure 4.2: Fault at Bus 611, Fault incidence=150^0 (a) BCG Fault measured at bus 650 (b) Decomposition of phase-C current at bus 611.

indicates that the THD with R_1 is less than 1 indicating predominantly the fundamental frequency component, whereas the THD using IMF_1 was always greater than 1, indicating higher frequency content.

$$THD = \frac{\sqrt{I_2^2 + I_3^2 + \cdots + I_n^2}}{I_1} \tag{4.11}$$

Thus the EMD process filters out the high frequency transients from the signal leaving behind the residue with fundamental or low frequency content. During ground faults the unwanted transients are injected into the healthy phase through the system neutral. This leads to false tripping of non faulty phase if the features are extracted from IMF which contains high frequency component. In order to avoid this adverse effect the residue is used instead of IMF. Hilbert transform is now applied on R_1 to obtain the instantaneous frequency and its standard deviation. The standard deviation of instantaneous frequency (SDIF) at each sample is defined as Fault index. This fault index is compared with a threshold for detection of the fault. The number of phases exceeding the threshold would eventually classify the type of the fault. If fault index of only one phase current exceeds the threshold then we classify the fault as LG fault. Thus, an N-phase fault occurs

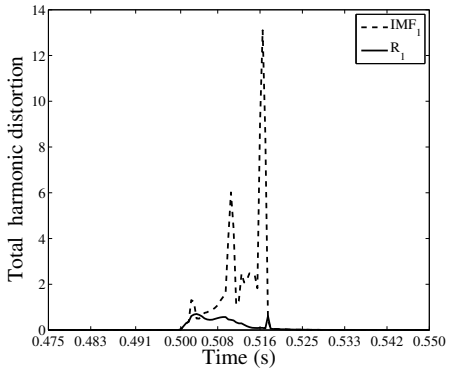


Figure 4.3: Total harmonic distortion of first level IMF and residue.

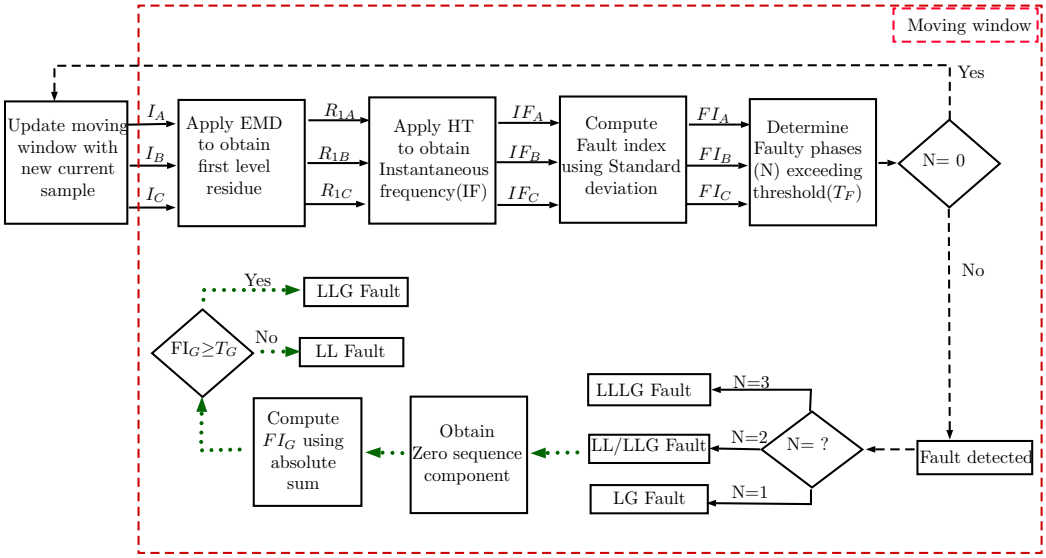


Figure 4.4: Block diagram of proposed algorithm for fault detection and classification.

if N number of phases exceed the fixed threshold. However, the LL and LLG faults cannot be distinguished by the number of faulty phases itself as both type of faults have the same number of faulty phases. Once the type of the fault is classified as two phase fault, further investigation is done to confirm whether the ground is involved or not. For this purpose, the zero sequence component is computed and a ground fault index is proposed. The absolute sum of the zero sequence component over the moving window is defined as ground fault index (FI_G). This index is compared with a threshold (T_G) obtained from quarter cycle post-fault data to decide whether the fault is LL or LLG fault. The block diagram of the proposed algorithm is shown in Fig. 4.4.

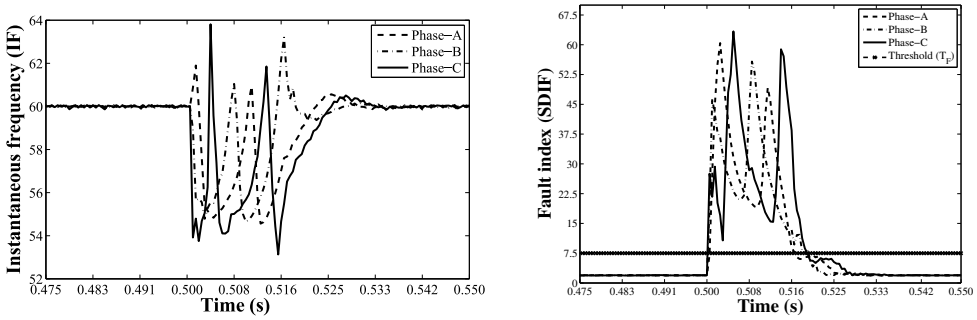
It is observed that the LL faults have lesser ground fault index than the LLG faults and can be easily classified with a fixed threshold (T_G=0.01). This threshold is set after comparing ground fault indices of all the cases of LL and LLG faults occurring at various fault locations, incidence angles and fault resistance values.

4.5 RESULTS AND DISCUSSION

The fault indices are obtained for all types of faults with 6 fault incidence angles, and at 12 fault locations excluding the substation bus. The maximum fault indices of faulty phases and healthy phases in a quarter cycle duration are designated as FPI_{SDIF} and HPI_{SDIF} respectively. A threshold value of 7.5 is set between $min(\text{FPI}_{SDIF})$ and $max(\text{HPI}_{SDIF})$ for this algorithm after analyzing all the cases considered in the following subsections.

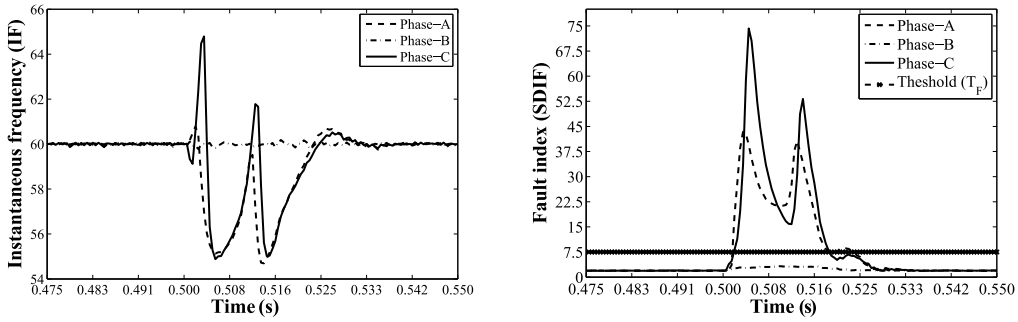
Variation of fault type

All the categories of faults considered are successfully classified using the proposed algorithm. ABC, ABCG, AB, BC, CA, ABG, BCG, CAG, AG, BG and CG type of faults are simulated at all the buses. The instantaneous frequency and standard deviation of phases A, B and C for ABCG fault at bus 611 is shown in Figs. 4.5(a) and (b) respectively. The fault index in Fig. 4.5(b) clearly shows that all the phases are above the threshold indicating a three-phase to ground fault. CAG fault simulated at the same bus is illustrated in Figs. 4.6(a) and (b). It can be noted in 4.6(b) that only the fault indices of phases A and C are above the threshold, whereas the healthy



(a) (b)

Figure 4.5: ABCG Fault at 611 (a) Variation of Instantaneous frequency (b) Variation of fault index (SDIF).



(a) (b)

Figure 4.6: CAG Fault at 611 (a) Variation of Instantaneous frequency (b) Variation of fault index (SDIF).

phase-B is clearly below the fixed threshold. It is difficult to distinguish CAG fault from CA fault as the number of faulty phases are same, therefore the zero sequence components are utilized in such cases as shown in Fig. 4.7. The fault index for LLG fault is greater than the threshold. A similar observation can be made for CG fault as noted in Fig. 4.8(a). The fault index variation is illustrated in Fig. 4.8(b), where only phase-C fault index exceeds the threshold. The difference in the time of fault occurrence and fault detection is defined in this work as FDT. It is quite evident that the FDT is less than 1/4th of the cycle for the proposed algorithm. This indicates that the algorithm is capable of detecting and classifying faults using less than quarter cycle post-fault data.

Variation of fault location

The fault current characteristics change with the location of the fault. The distribution line parameters in the fault current path also influence the fault transients. The variation of FPI and HPI with fault location can be observed in Fig. 4.9. It can be noticed that the FPI is above the threshold of 7.5 for all the fault locations considered. Thus it is proved that the algorithm is not affected by changes in fault locations.

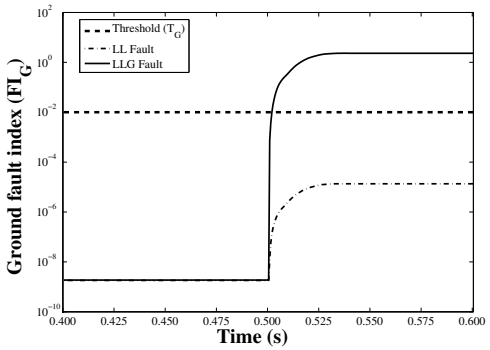
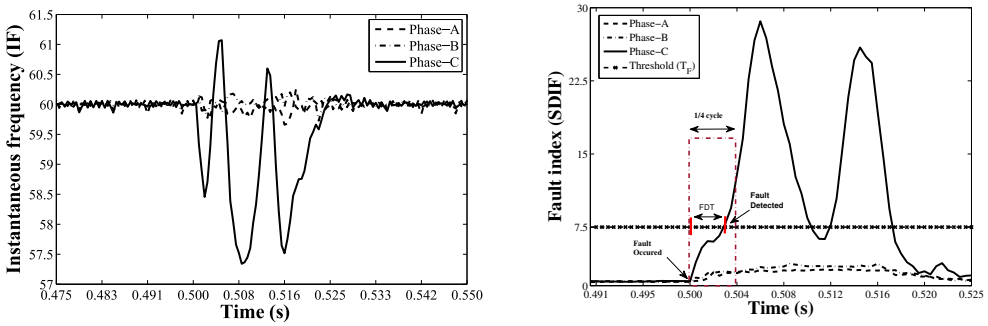


Figure 4.7: Variation of ground fault index with time.



(a) (b)

Figure 4.8: CG Fault at 611 (a) Variation of Instantaneous frequency (b) Variation of fault index (SDIF).

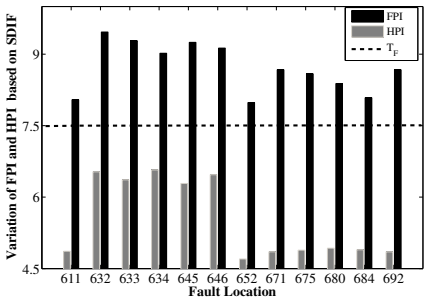


Figure 4.9: Variation of FPI and HPI with fault location.

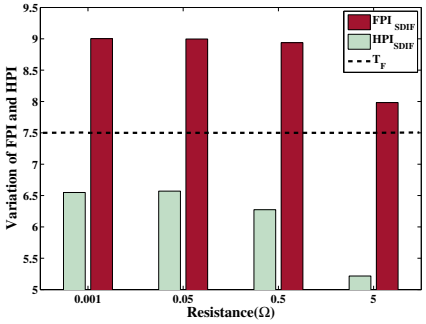


Figure 4.10: Variation of FPI and HPI with fault resistance.

Table 4.1: Fault index values for various fault incidence angles

Fault →	LLLG Fault (ABCG)			LLG Fault (BCG)			LL Fault (BC)			LG Fault (BG)		
FIA ↓	A	B	C	A	B	C	A	B	C	A	B	C
0^0	**32.575**	**30.674**	**58.731**	2.414	**25.791**	**25.608**	1.980	**23.166**	**26.151**	4.620	**18.368**	3.558
30^0	**34.830**	**21.354**	**31.702**	3.031	**14.733**	**24.905**	1.980	**12.745**	**19.648**	4.417	**12.150**	3.880
60^0	**30.710**	**57.072**	**32.774**	3.054	**23.004**	**27.034**	1.980	**28.560**	**23.933**	3.249	**11.488**	2.954
90^0	**21.199**	**31.012**	**35.271**	2.446	**30.602**	**28.840**	1.979	**32.139**	**25.480**	2.044	**18.372**	2.140
120^0	**60.294**	**32.422**	**28.981**	2.789	**32.539**	**26.771**	1.953	**34.329**	**25.839**	2.156	**20.428**	2.525
150^0	**30.247**	**34.240**	**20.970**	2.056	**32.796**	**21.733**	1.939	**32.056**	**22.186**	3.226	**21.468**	3.714

Variation of fault resistance

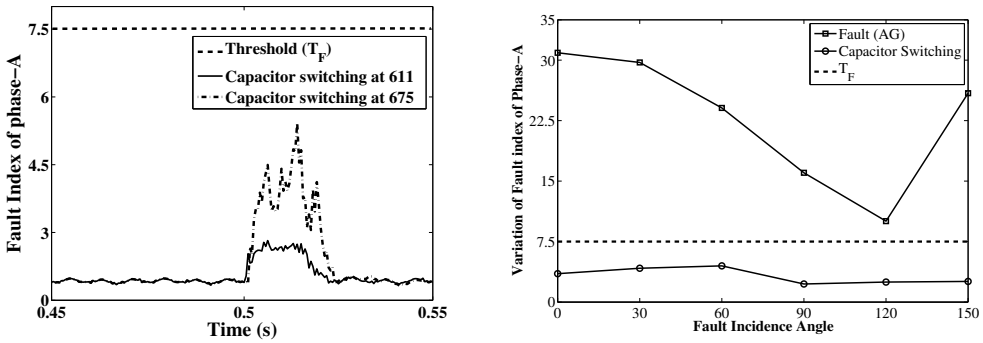
The proposed algorithm is tested with variations of fault resistance. For an IEEE 13 bus system, a fault resistance value of 3.46 Ω (Z_{Base}), results in minor change in the fault current amplitude. Hence, it is difficult to detect a fault [Paz *et al.*, 2010]. However the proposed algorithm is found to be effective upto 5 Ω which is almost 45% greater than the system base impedance. The variation of FPI and HPI with changes in fault resistance is plotted in Fig. 4.10. It is clearly demonstrated that the algorithm is immune to changes in the fault resistance.

Variation of fault incidence angle

As the fault transients are dependent on the fault incidence angle, the proposed algorithm is tested for faults occurring at 6 inception angles. All the faults are simulated with fault incidence angles of $(0^0 - 150^0)$ in steps of 30^0. The variation of the maximum value of the fault index with the fault incidence angle for different types of faults at bus 634 are shown in Table 4.1. ABCG, BCG, BC and BG faults are tabulated with only the faulty phases highlighted to show that the fault indices are above the threshold for faulty phases and below the threshold for healthy phases. Clearly the proposed algorithm is not affected by the changes in the fault incidence angle.

Capacitor switching

Although capacitor switching is performed to improve the voltage profile, it can also introduce transients. The proposed algorithm successfully differentiates the capacitor switching transient from a faulty transient. The capacitor switching events are simulated at buses 611 and 675 at the same instant of fault. It is observed that fault indices for phases during capacitor switching are lower than the prescribed fault threshold. The variation of phase-A at both the locations is shown in Fig. 4.11(a). The other phases also follow a similar phenomenon. As the



(a) (b)

Figure 4.11: Capacitor switching (a) Variation of phase-A fault index (b) variation with fault incidence angle.

capacitor switching is also dependent on the incidence of the event, the variation of the fault index of phase-A is shown for different incidence angles in Fig. 4.11(b). The phase-A fault index corresponding to fault (AG) at bus 675 is above the threshold and that of capacitor switching is below the threshold for all incidence angles, thus proving that the algorithm is not affected by switching of compensating equipment.

4.6 PROPOSED ALGORITHM WITH DG

Though the threshold-based algorithm discussed could detect and classify faults without DG, it involves more thresholds to detect islanding and switching transients in the case of DG. Therefore an approach to considering machine learning is proposed in this section to avoid multiple thresholds in the presence of DG and noise. For an IEEE 13 bus system, all the types of faults are considered at each bus exempting the substation at 0.5 s in the presence of solar DG. The islanding event is simulated with a three-phase breaker that disconnects the utility grid from the rest of the distribution system. The switching transients involving transformer at bus 634, feeder at bus 692, and capacitor at buses 611 and 675 are also simulated at 0.5 s. The three-phase currents measurement is done at the substation or bus 650 over one cycle's moving window. The sampling frequency of 1.92 kHz considered for a system frequency of 60 Hz results in 32 samples per cycle. The three-phase currents measured after a two-phase to the ground (ABG) fault occurred with an inception angle of 30^0 at bus 671 is shown in Fig. 4.12. These fault currents are then decomposed using empirical mode decomposition to obtain the first level residue (R_1) for each phase current. The phase-A current in the enclosed moving window of one cycle containing a quarter cycle PFD is shown decomposed into first level IMF and residue using EMD in Fig. 4.13.

The high-frequency content associated with noise or injection of harmonics through the grounded path is removed from the signal before applying HT. HT is then applied on R_1 to obtain four instantaneous attributes or feature vectors, namely, Instantaneous frequency (IF), standard deviation of instantaneous frequency (SDIF), Instantaneous Amplitude (IA), and Standard deviation of Instantaneous Amplitude (SDIA). The variation of the four HHT features extracted for phase-A for the above fault case at an SNR of 20 dB is illustrated in Fig. 4.14.

All the features exhibit significant change after the occurrence of the fault. A similar approach is undertaken for all the current signals associated with islanding and switching transients considered. To avoid the thresholding issues associated with the events considered, the DT-based ML model is designed to classify the disturbances into four classes of fault, islanding, no-fault

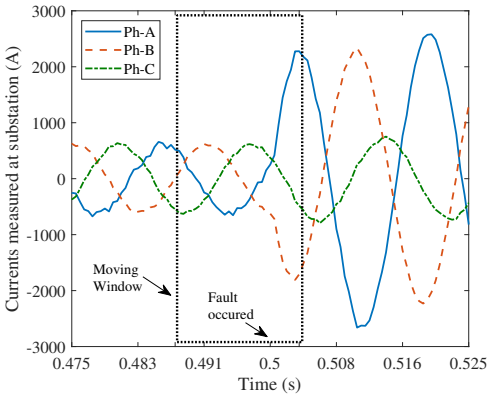


Figure 4.12: Fault currents during ABG fault at 671.

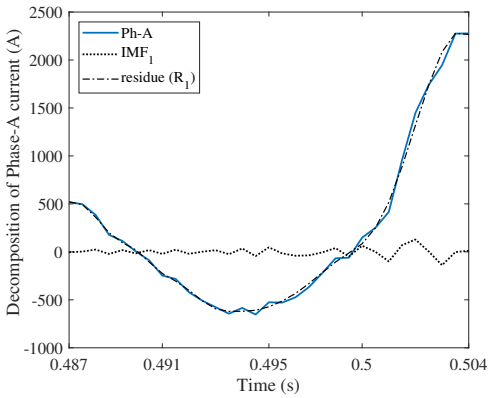


Figure 4.13: EMD decomposition of Phase-A current in moving window.

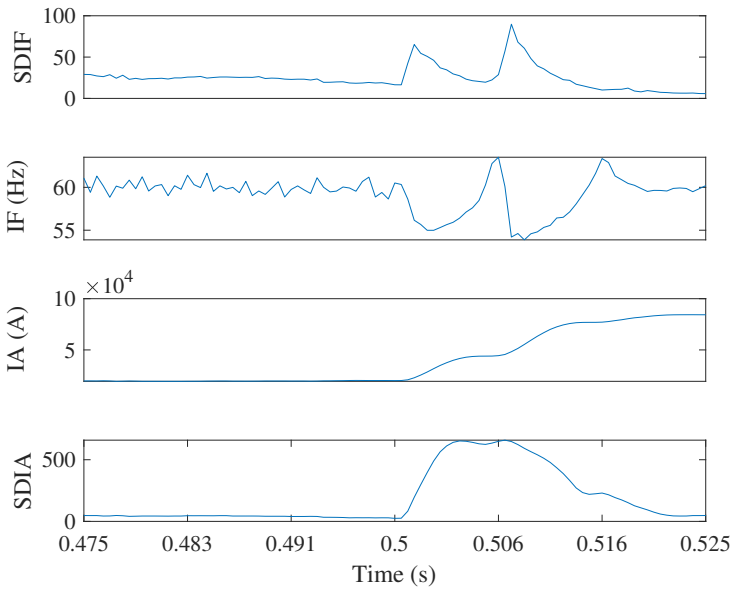


Figure 4.14: Variation of HHT features for Phase-A during fault with time.

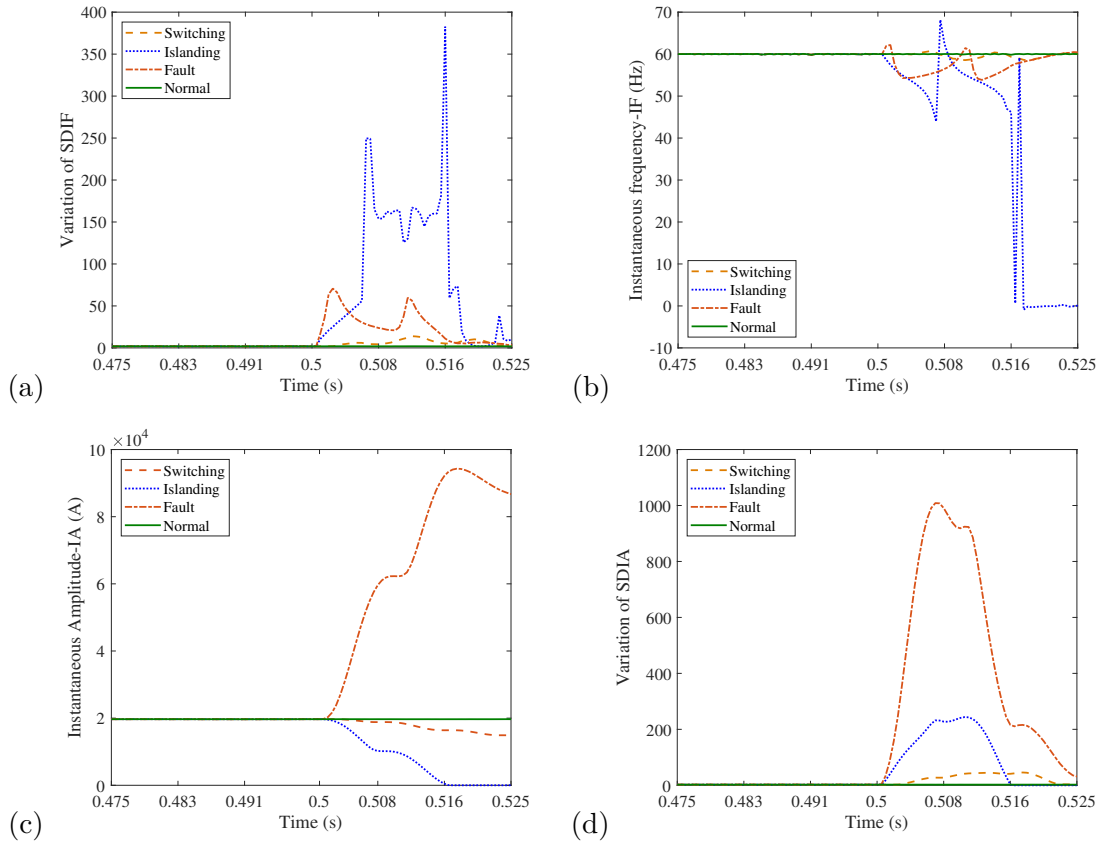


Figure 4.15: Fault index variation for features considered (a) SDIF (b) IF (c) IA (d) SDIA.

(normal), and switching transients. The variation of each feature corresponding to the four classes is shown in Fig. 4.15. Only Phase-A feature vectors corresponding to ABCG fault at 671, islanding, and switching transient associated with feeder energizing (FDE) are demonstrated here. It can be observed from Fig. 4.15(a) that the SDIF exhibits a sharp rise with a short duration for fault, whereas the magnitude is higher for islanding. The instantaneous frequency is calculated precisely as system frequency of 60 Hz before fault as shown in Fig. 4.15(b), but after 0.5 s, its variation is high in the case of islanding when compared to a fault and switching transient. Also, the transients exist for a longer duration in the case of islanding than in other cases. The IA, as expected, was very high in the case of a fault, whereas the IA reduced to zero within a cycle for islanding case, as depicted in Fig. 4.15(c). After the FDE event, IA reduced, but it reached a steady state after a few cycles. The SDIA has shown a sharp rise and higher magnitude in the case of a fault. The islanding event exhibits higher SDIA than FDE, as shown in Fig. 4.15(d). PFD of quarter cycle, half cycle, and full cycle after 0.5 s is then utilized to obtain fault indices, which are then fed to a DT to detect, classify and locate the faulty zone. The proposed algorithm is demonstrated through a flow chart in Fig. 4.16.

4.7 RESULTS AND DISCUSSION

4.7.1 Islanding and Fault Detection On IEEE 13 Bus System

The data simulated for the ML model's design is enumerated in Table 4.2. The model is trained with 70% of generated data, and 30% of the unused data is used for testing. The proposed

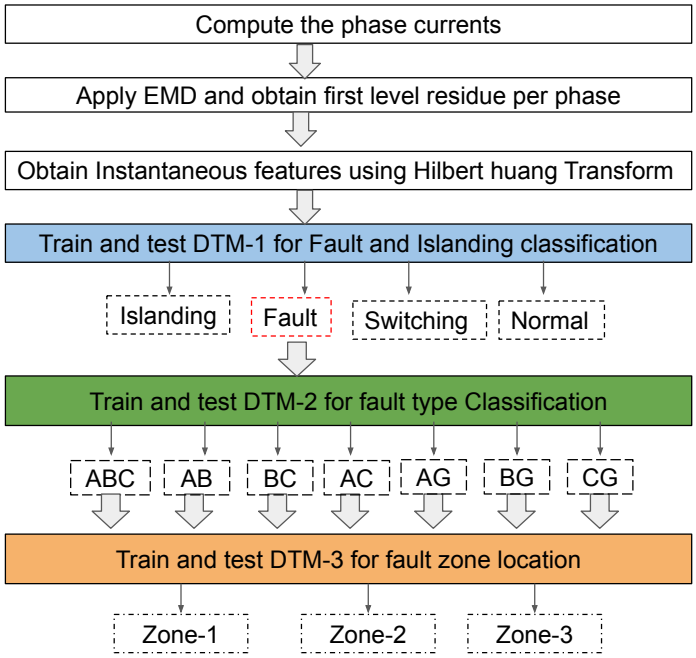


Figure 4.16: Flow Chart of the Proposed Algorithm.

Table 4.2: Data set for DT model

Event	**Dataset Simulated**	**Total cases generated at each DG capacity**
Fault	FIA=0-150 deg in steps of 30 deg (6 cases) DG =200 kW, 400 kW, 600 kW and 800 kW (4 cases) Fault locations (all except bus 650)-12 cases Fault types (ABC/ABCG, AB/ABG, BC/BCG, CA/CAG, AG, BG, CG) -11 fault types classified into 7 cases Noise levels (20 dB, 30 dB, 40 dB)-100 cases per level	6x12x7x100=50400
Islanding	FIA=0-150 deg in steps of 30 deg (6 cases) DG =200 kW, 400 kW, 600 kW and 800 kW (4 cases) Noise levels (20 dB, 30 dB, 40 dB)-50 cases per level	6x50=300
Switching transient	Transformer energization at 634 Transformer de-energization at 634, Load switching at 652, Feeder on/OFF at 692, Capacitor switching at 611 and 675 (7 cases) FIA=0-150 deg in steps of 30 deg (6 cases) DG =200 kW, 400 kW, 600 kW and 800 kW (4 cases) Noise levels (20 dB, 30 dB, 40 dB)-50 cases per level	7x6x50=2100
No fault	Pre fault data of 11 types of faults, at 6 FIA, 12 locations 10 cases per noise level	11x6x12x10=7920

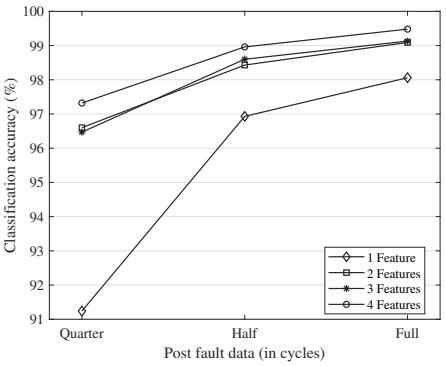


Figure 4.17: Variation of accuracy with PFD.

detection algorithm's effectiveness is illustrated with respect to the number of features used for training. The accuracy of the all DT models is nothing but the percentage ratio of the sum of all true predicted cases (P_{true}) to the sum of all test cases.

$$Accuracy(\%) = \frac{\sum P_{true}}{\sum Testcases}.100 \tag{4.12}$$

Variation of PFD

Any protection algorithm's speed depends on the PFD considered for training the ML model. Thus the proposed algorithm is tested for various PFD of quarter cycle, half cycle, and full cycle. It can be observed that the accuracy of fault detection is higher for full-cycle data and slightly less for half cycle. The results are presented by fixing the size of DG at 200 kW at 20 dB SNR. The features determine the algorithm's complexity; the more the number of features, the more the DT's processing time. Thus the effect is observed with the number of features in Table 4.3, as an array one feature [FT1], two features [FT1 FT2], three features [FT1 FT2 FT3], and four features [FT1 FT2 FT3 FT4]. When the model is fed with an increased array of information, the accuracy with quarter cycle PFD also increased. The accuracy with the quarter cycle is 91.24% with one feature, and it increased to 97.32% with all four features. The accuracy is as high as 99.48% with a full cycle PFD. The accuracy with three features is marginally better than obtained with two features. The variation is shown in Fig. 4.17.

Variation of noise level

The proposed algorithm is tested for its robustness with the current signals contaminated with noise. The noise in the signal distorts the current signal similar to that of CT saturation and thus affects the algorithm's robustness. It is established that the proposed algorithm is not affected by noise up to 20 dB SNR in the presence of solar DG. The algorithm is tested by adding random noise with variations in standard deviations to obtain three noise levels of 40 dB, 30 dB, and 20 dB. The SNR is computed using equation (4.13).

$$SNR(dB) = 10 * log_{10}\frac{E_{Signal}}{E_{Noise}} \tag{4.13}$$

where, E_{Signal} and E_{Noise} are the energies of the signal and the random noise respectively. The variation of accuracy with noise using quarter cycle PFD at 200 kW is shown in Fig. 4.18. It is evident that the accuracy decreased with an increase in the noise level.

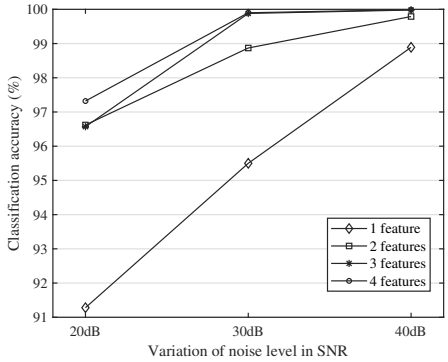


Figure 4.18: Variation of accuracy with noise.

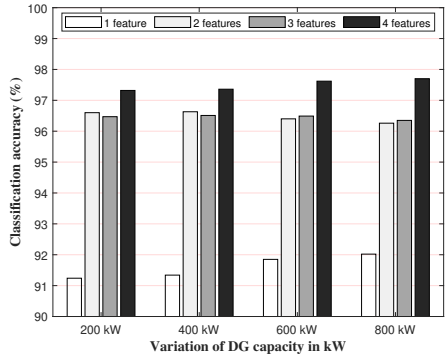


Figure 4.19: Variation of accuracy with DG.

Variation of DG capacity

The algorithm is tested for various levels of renewable energy capacity installed. The performance in the presence of DG is obtained for 200 kW, 400 kW, 600 kW, and 800 kW. As the DG size increased, the power mismatch also reduced. However, the effectiveness of the algorithm is not affected by the increase in the DG capacity. The variation of accuracy with DG is shown in Fig. 4.19. It can be observed that the accuracy is less than 92% with one feature, but it increased over 96% with an addition of another two features. The accuracy was above 97% when all the four features are fed to the DTML-1.

4.7.2 Fault Classification In IEEE 13 Bus System

After the fault detection, it is imperative to classify the type of fault in the diagnosis sequence. For this purpose, the fault class's fault indices are fed as input to the model DTML-2 to classify the faults into seven classes 3Ph-ABC, 2Ph-AB, 2Ph-BC, 2Ph-CA, 1Ph-AG, 1Ph-BG, and 1Ph-CG. The two-phase and two-phase to ground faults are classified as one class in this work. i.e., AB and ABG type of faults are represented as 2Ph-AB. Similarly, ABC and ABCG are grouped under 3Ph-ABC. The fault classification is tested for various penetration levels of solar DG in the presence of noise.

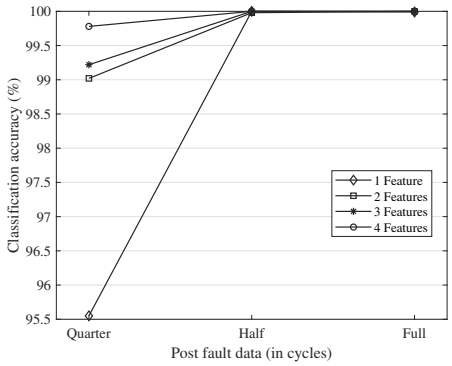


Figure 4.20: Variation of accuracy with PFD.

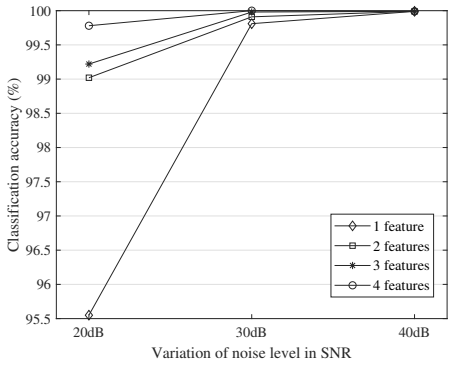


Figure 4.21: Variation of accuracy with noise.

Variation of PFD

The results are presented by fixing the size of DG at 200 kW and noise at 20 dB SNR. The proposed fault classification algorithm is tested for a quarter cycle, half cycle, and full cycle PFD. It can be observed that the accuracy of fault classification is nearly 100 % for full cycle data and half cycle data. The quarter cycle PFD with one feature resulted in an accuracy of 95.5%. The accuracy was over 99% for two and three features. By increasing the number of features to four, the accuracy was above 99.5%. The variation of accuracy with PFD is shown in Fig. 4.20.

Variation of noise level

The algorithm is tested for various noise levels similar to that of the fault detection model DTML-1 by fixing the DG size as 200 kW and PFD as a quarter cycle. The algorithm has 100% accuracy for 40 dB but slightly less than 100% for 30 dB. However, the accuracy is affected due to the number of features for quarter cycle PFD. Its accuracy is 95.5% with only one feature as input to DTML-2 and increased with an increase in the number of input features. The variation of accuracy with noise level is shown in Fig. 4.21.

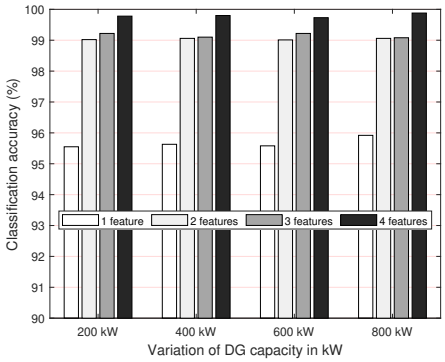


Figure 4.22: Variation of accuracy with DG.

Table 4.3: Features computed for decision tree.

Feature	Fault index
FT1	Max(abs[Standard deviation of IF])
FT2	Max(abs[Instantaneous Frequency IF])
FT3	Max(abs[Standard deviation of IA])
FT4	Max(abs[Instantaneous Amplitude IA])

Variation of DG capacity

A similar approach as that of DTML-1 was undertaken to test the fault classification algorithm DTML-2 for DG capacity variations. The PFD is fixed as a quarter cycle and noise level at 20 dB SNR. When only one feature was fed to the DTML-2, the fault classification accuracy was consistent between 95.5% to 96%. When the number of features increased, the accuracy was also over 99% at all the penetration levels. With all the four features fed to the model, the accuracy was over 99.5%, even with a DG size of 800 kW. Thus the proposed DTML-2 is not affected by the size of DG. The variation with DG capacity is shown in Fig. 4.22.

4.7.3 Fault Location In IEEE 13 Bus System

After successful fault classification, the next step is to locate the fault. All the fault indices presented in Table 4.3 which are previously utilized for DTML-1 and DTML-2, are used for the formation of fault location DT based on the data pertaining to the seven fault types. Each of which decides the fault location as one of the three fault zones. The distribution system is divided into three zones as zone-1 (632, 633, 634, 645, and 646) closer to the substation, Zone-2 (611, 652, and 684) having underground feeders, and zone-3 (671, 675, 692, and 680) containing the DG. The case studies used for fault location involves all the fault types at six fault incidence angles of 0^0, 30^0, 60^0, 90^0, 120^0 and 150^0 at each location. The randomly generated cases at each location per noise level are 4200.

Effect of PFD

The fault location algorithm is tested with the variation of PFD. The zone of the fault location is identified using the fault-specific decision trees. It can be observed that the quarter cycle PFD resulted in an overall accuracy of 87.69% with all the four features from each phase fed

Table 4.4: Fault location accuracy with PFD.

Fault	PFD		
	Quarter	Half	Full
3Ph-ABC	95.18	98.42	99.21
2Ph-AB	91.52	90.60	94.39
2Ph-BC	92.26	92.82	97.31
2Ph-CA	90.09	90.83	96.52
1Ph-AG	76.34	88.24	94.21
1Ph-BG	83.98	90.50	94.72
1Ph-CG	84.49	88.05	93.42
Overall	**87.69**	**91.20**	**95.68**

Table 4.5: Fault location accuracy with noise.

Fault	Noise		
	40 dB	30 dB	20 dB
3Ph-ABC	98.93	98.75	98.42
2Ph-AB	99.02	97.03	90.60
2Ph-BC	99.07	98.28	92.82
2Ph-CA	99.35	98.56	90.83
1Ph-AG	98.19	91.94	88.24
1Ph-BG	98.14	95.23	90.50
1Ph-CG	98.10	94.16	88.05
Overall	**98.69**	**96.28**	**91.20**

to the DT as input. The half cycle PFD has an average accuracy of 91.20%. The accuracy of fault location improved with the PFD of full-cycle at 95.68%. It can also be confirmed that the accuracy is more excellent for three-phase faults when compared to two-phase faults. The accuracy with single-phase faults is less for all the PFD when compared to other types of faults. The accuracy obtained for all the PFD considered at 20 dB SNR with 200 kW DG is presented in Table 4.4.

Effect of Noise

The fault location accuracy at each noise level using half cycle PFD and with 200 kW capacity is enumerated in Table 4.5. It can be ascertained that the accuracy is high for a 3Ph-ABC model but is least for 1P fault models at all the noise levels. The 2Ph-AB, 2Ph-BC and 2Ph-CA faults have better performance with noise. The 1Ph fault modules have shown less accuracy when compared to other faults due to the transients in the healthy phases. Also, with an increase in the noise level, the accuracy decreased for all types of models. The overall accuracy was as good as 98.69% for a noise level of 40 dB and decreased to 91.20% with 20 dB SNR.

Effect of DG

The fault location algorithm needs to be tested for variations of the DG size to establish the adaptability of the algorithm. For this purpose, the algorithm is tested with an increase of DG size in steps of 200 kW. It can be observed from Table 4.6, the variation of DG size from 200 kW to 800 kW does not affect the fault location accuracy. The accuracy using quarter cycle information is 87.69%, 87.81%, 87.93%, and 87.94% for the DG capacities of 200 kW, 400 kW, 600 kW, and 800 kW, respectively. The accuracies are almost constant with an error margin of 0.25%. As PFD increased to half cycle, the overall accuracy improved to 91.46% for 800 kW. The error margin for half cycle PFD is 0.44%. With full-cycle PFD the accuracies are 95.68%, 94.94%, 95.58%, and

Table 4.6: Variation of accuracies with DG size.

Fault Model	Quarter Cycle				Half Cycle				Full Cycle			
	200 kW	**400 kW**	**600 kW**	**800 kW**	**200 kW**	**400 kW**	**600 kW**	**800 kW**	**200 kW**	**400 kW**	**600 kW**	**800 kW**
3Ph-ABC	95.18	95.04	95.83	94.86	98.42	98.70	98.75	98.37	99.21	98.98	99.35	99.02
2Ph-AB	91.52	90.60	91.62	90.74	90.60	92.87	92.22	92.54	94.39	95.04	96.15	96.99
2Ph-BC	92.26	93.51	94.39	93.10	92.82	93.19	92.31	91.62	97.31	95.41	97.26	96.99
2Ph-CA	90.09	90.69	88.47	89.86	90.83	91.11	93.00	91.01	96.52	95.83	96.11	95.46
1Ph-AG	76.34	76.57	78.98	79.62	88.24	87.12	87.26	88.70	94.21	93.56	94.16	95.27
1Ph-BG	83.98	83.51	83.42	83.33	90.50	88.75	87.36	89.39	94.72	94.81	94.07	94.39
1Ph-CG	84.49	84.72	82.82	84.07	88.05	89.76	89.53	88.61	93.42	90.92	91.99	91.06
Overall Acc	**87.69**	**87.81**	**87.93**	**87.94**	**91.20**	**91.64**	**91.49**	**91.46**	**95.68**	**94.94**	**95.58**	**95.60**

Table 4.7: Comparison of fault location accuracy with DT and SVM at 20 dB SNR for various DG values.

Fault Model	Quarter Cycle-DT					Quarter Cycle-SVM				
	No DG	**200 kW**	**400 kW**	**600 kW**	**800 kW**	**No DG**	**200 kW**	**400 kW**	**600 kW**	**800 kW**
3Ph-ABC	86.98	95.18	95.04	95.83	94.86	93.68	97.97	98.07	97.87	97.50
2Ph-AB	81.55	91.52	90.60	91.62	90.74	90.73	95.95	96.41	96.37	95.38
2Ph-BC	88.59	92.26	93.51	94.39	93.10	96.24	97.23	96.94	97.53	96.70
2Ph-CA	77.22	90.09	90.69	88.47	89.86	88.16	95.95	95.85	95.80	96.16
1Ph-AG	53.91	76.34	76.57	78.98	79.62	68.07	90.47	89.83	91.38	91.64
1Ph-BG	75.15	83.98	83.51	83.42	83.33	87.11	95.67	96.83	95.43	96.09
1Ph-CG	72.89	84.49	84.72	82.82	84.07	87.31	95.03	94.00	94.49	94.58
Overall Acc	**76.61**	**87.69**	**87.81**	**87.93**	**87.94**	**87.32**	**95.47**	**95.41**	**95.55**	**95.44**

95.60% for the DG levels in incrementing order with an error margin of 0.74%. The maximum error margin is less than 1%, which proves that the variations of DG sizes have a negligible effect on the algorithm.

In order to improve the fault location accuracy, the extracted features over a quarter cycle are fed to SVM in the presence of DG and without DG, which involves higher fault resistance. The fault location accuracies for various DG values for an IEEE 13 bus system are shown in Table 4.7. It can be inferred that the fault location accuracy with DT suffers compared to that of SVM. Therefore it can be concluded that for the cases considered, the fault location accuracy possible using DT is 82.23%. The fault location accuracy with the DT model without DG is 76.61%, and the average fault location accuracy for all DG capacities is 87.85%. The overall accuracy possible with DT with DG and without DG is 82.23%. However, with SVM, the fault location accuracy improved to 87.32% without DG, and the average accuracy with DG is over 95%. The overall accuracy possible with SVM with DG and without DG is 91.38%. The usual trend is observed even with SVM in that if the fault is three-phase, the fault location accuracy is also better compared to single-phase faults. Therefore for better fault location accuracy, it is recommended to employ the HHT and SVM combination.

4.7.4 Analysis on IEEE 34 Bus System

The proposed algorithm is tested on an IEEE 34 bus system for the exact switching and islanding case studies mentioned in section 3.5.3 for EMD based algorithm. The 11 types of faults simulated at each of the seven locations considered are clubbed into seven fault categories as mentioned in the proposed algorithm. The data set considered for the ML algorithm is mentioned below, in which 70% of data is used for training at random, and the remaining 30% of untrained data is used for testing.

Table 4.8: Fault location accuracy for IEEE 34 bus system.

Fault Model	**DT**		**SVM**	
	Without DG	With DG	Without DG	With DG
3Ph-ABC	96.82	97.06	100	99.93
2Ph-AB	88.73	93.01	98.13	98.29
2Ph-BC	92.93	94.04	99.61	99.32
2Ph-CA	89.28	90.07	98.63	99.01
1Ph-AG	85.87	89.36	96.47	97.28
1Ph-BG	90.31	93.17	98.76	98.95
1Ph-CG	95.31	96.34	99.09	99.24
Overall Acc	**91.32**	**93.29**	**98.67**	**98.86**

- Fault cases: 7 (Fault locations) x 7 (Fault types) x 6 (FIA) x 50 (cases at each noise level) = 29400.
- Islanding cases: 6 (FIA) x 50 (cases at each noise level) = 300.
- Switching cases: 6 (Non faulty transients) x 6 (FIA) x 50 (cases at each noise level) = 1800.
- Normal cases: 11 (Fault types) x6 (FIA) x 7 (Fault locations) x 10 (cases at each noise level) = 4620.

The results using PFD of the quarter cycle at 20 dB SNR are only presented here. When the DG of 400 kW is introduced, the fault detection problem is formulated as a four-class problem to classify islanding, fault, switching transient, and no fault using DT. The accuracies obtained are 95.18%, 95.89%, 96.17%, and 96.59% for one, two, three, and four features, respectively. Similar to the fault classification of the IEEE 13 bus system, the faults are classified into seven categories. The accuracies obtained are 96.63%, 99.41%, 99.49%, and 99.71% for one, two, three, and four features, respectively.

The ML algorithm is applied for the system without DG where the events are classified into three classes, namely fault, switching transient, and normal operation. The accuracies obtained at 20 dB SNR with quarter cycle PFD are 97.42%, 97.63%, 98.03%, and 98.91% for one, two, three, and four features, respectively. The fault classification accuracy is 96.90%, 99.75%, 99.76% and 99.77%.

The fault location accuracy using quarter cycle post-fault data at 20 dB SNR with DG and without DG using DT and SVM models is shown in Table 4.8. It can be observed that the fault location accuracies are 91.32% and 93.29% with DG and without DG respectively using DT. The accuracies with SVM are higher than DT and obtained as 98.67% without DG and 98.86% in the presence of DG.

4.8 PERFORMANCE COMPARISON

The proposed algorithm is better when compared to other similar work in literature in terms of the number of signals used, PFD considered, and the accuracy of classification. An SVM-based method is proposed in Baghaee *et al.* [2019] to discriminate grid faults from islanding using seven signals. However, only current signals measured at single location (substation) are needed here. The proposed algorithm is compared with the PFD utilized for islanding and fault classification on an IEEE 13 bus system. In Mishra *et al.* [2017], the islanding and faults are classified with HHT as signal processing technique and achieved an accuracy of 90% with Extreme Learning Machine (ELM), 90.4% accuracy with Bidirectional Extreme Learning Machine (BELM), and 96.06% with

weighted BELM (WBELM) using a one-cycle PFD. In Mishra and Rout [2016a], a one-cycle PFD was considered to achieve islanding classification of 91.5% with a combination of Fast discrete S-Transform (FDST) and BELM. In this work using DT, an accuracy of 97.32% is achieved with only quarter cycle PFD, making it the fastest of all. Also, in Mishra and Rout [2017], the faults are detected and classified with HHT as the preprocessor. Though the authors did not mention the post-fault information and noise level, the fault detection was achieved with an accuracy of 96.75% with Naive Bayesian Classifier (NBC), 96.15% with Support Vector Machine (SVM), and 96.99% with ELM. It reported fault classification accuracy of 91.41% with NBC, 91.91% with SVM, and 93.93% with ELM. In this work using HHT and DT, fault detection was achieved with an accuracy of 97.32% and fault classification accuracy over 99.5% at 20 dB SNR.

The reason for better accuracy is that EMD first filters the unwanted noise from the signal to obtain the first-level residue. The instantaneous features obtained through HHT on this residue level further distinguish the events for effective partitioning by DT algorithm. In a grounded distribution system, the system uncertainties during fault cause an abrupt rise in the harmonics of healthy phases along with faulty phases, causing a failure in fault detection. Therefore, removing these harmonics from the healthy phases is imperative in the form of intrinsic mode function. Thus, when the outliers are removed from the healthy phases, the DT performance improves as the cross-validation and test sample errors reduce.

The SVM provided better results for fault location due its ability to form non linear boundaries by increasing the margin between the support vectors. The fault location accuracies are better than the previous algorithm for the test systems considered. Using EMD and DT in the previous algorithm the fault location accuracies are 88.71% for IEEE 13 bus system and 98.34% for IEEE 34 bus system using half cycle PFD. But by using HHT and SVM, the accuracies obtained are 91.38% and 98.76% for the IEEE 13 bus and IEEE 34 bus system, respectively using only quarter cycle PFD.

The computational time of the proposed algorithm is obtained on a PC with Intel(R), Xeon (R) CPU E5-2618L v3 @ 2.30 GHz, with 8 GB RAM. For a BG fault case study on IEEE 13 bus system with 200 kW DG at 20 dB SNR, the features were extracted using HHT within a time of 4.8 ms over a moving window. The machine learning algorithm execution time using DT was approximately 0.25 μs for fault/islanding detection, 0.27 μs for fault classification and 1.35 ms for fault zone location. The fault location algorithm using SVM consumed a time of 12 ms.

4.9 CONCLUSION

This chapter focusses on the machine learning-based protection algorithm for a distribution system demonstrating the need for it to avoid multiple thresholds. Hilbert-Huang transform features from the first-level residue for the first time are successfully utilized for islanding and fault classification in the presence of distributed generation. The instantaneous features or attributes from the first-level residue after EMD are fed to the decision tree model to achieve it. These features enabled faster fault diagnosis with post-fault information of only a quarter cycle. The three-phase currents obtained at the substation, without any measurement at DG location, are utilized for both islanding and fault detection in an IEEE 13 bus and 34 bus network in the presence of solar renewable energy. Due to the moving window, the proposed algorithm can be utilized as an online solution using the substation's current signals. Also, the fault is located using DT and SVM and found out that the combination of HHT and SVM yields better results in the presence of noise up to 20 dB SNR. The proposed algorithm successfully classifies various nonfaulty transients like transformer energizing and de-energizing, ON/OFF feeder switching, capacitor and load switching. The proposed algorithm is not affected by the variation of DG size, thus establishing its reliability for distribution system protection in growing renewable energy penetration.