IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Fault Protection in Microgrid Using Wavelet Multiresolution Analysis and Data Mining

SHAZIA BALOCH, SUNIL SRIVATSAV SAMSANI (Student Member, IEEE), AND MANNAN SAEED MUHAMMAD (Senior Member, IEEE)

Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon 16419, Korea

Corresponding author: Mannan Saeed Muhammad (email: mannan@skku.edu)

ABSTRACT The protection problems in microgrid effect the reliability of the power system caused due to high distributed generator penetrations. Therefore, fault protection in microgrid is extremely important and needs to be resolved to enhance the robustness of the power system. This manuscript proposes a combined signal processing and data mining-based approach for microgrid fault protection. In this study, first the multiresolution decomposition of wavelet transform is employed to preprocess the voltage and current signals to compute the total harmonic distortion of the voltage and current. Then, the statistical indices of standard deviation, mean, and median of the total harmonic distortion and the negative sequence components of active and reactive power are used to collect the input data. After that, all the available data is provided to the random forest-based classifier to evaluate the efficiency of the proposed scheme in terms of the detection, identification, and classification of faults. This study used different aspects for data collection by simulating various fault and no-fault cases for both looped and radial configurations under grid-connected and islanded modes of operation. The simulations were performed on a standard medium voltage microgrid using MATLAB/SIMULINK, whereas the analysis for testing and training of the random forest were conducted in Python. It is recognized that the proposed method performed better than support vector machines and decision tree that are reported in the literature. The results further demonstrate that the proposed method can also detect simultaneous faults, and it is also effective against measurement noise.

INDEX TERMS Data-mining, discrete wavelet transform, feature extraction, fault protection, multiresolution analysis, random forest.

I. INTRODUCTION

THE advancement in distributed generators (DGs) has played a vital role in resolving the problems of conventional power systems over the past several decades. The development of these DGs offer a viable solution against the nonavailability and exhaustion of fossil fuels, rapid load growth, environmental pollution, and expensive petroleum products and gases. DGs met with great success to address a number of technical, regulatory, and contemporary problems of the conventional power system [1], [2].

This new technology has shifted the conventional power system to low voltage active distribution networks called microgrid. A microgrid is made up of a collection of DGs, energy storage units, and various loads, which are controllable with the monitoring and protection devices [3], [4], [5]. They are typically connected to the utility through a circuit breaker at the point of common coupling (PPC) at distribution or sub-transmission voltage levels. The microgrid operates in synchronization with the utility under normal operating conditions. However, if the main grid undergoes troubles such as voltage fluctuations or frequency variations, it will completely disconnect from the utility and continue to operate as an islanded to supply the critical loads. The DGs in a microgrid are of a much smaller capacity, unlike the large generators in traditional systems, because of their low energy density and dependence on geographical conditions of a region [6], [7], [8]. They are located near to customers to provide electric and heat loads with the proper voltage level and frequency with negligible line losses, and to prevent network congestion. They bring technical, economic, and environmental benefits to the modern power system with improved reliability by keeping the power on, when the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3088900, IEEE Access

TABLE 1: Abbreviations used in manuscript.

Abbreviations	
DSP	Digital signal processing
DWT	Discrete wavelets transform
DGs	Distributed generators
DT	Decision tree
DERS	Distributed energy resources
MRA	Multiresolution analysis
WT	Wavelets transform
FT	Fourier transform
PMU	Phasor Measurement Unit
ANN	Artificial neural network
PCC	Point of common coupling
THD	Total Harmonic Distortions
IIDGs	Inverter-interfaced distributed generators
SVM	Support vector machines
DL	Distribution line
WPT	Wavelet packets transform
CB	Circuit breaker
DNN	Deep neural network
RF	Random forest
*M*iB	Mebibyte
A	Approximation coefficient
$\psi_j(t)$	Mother wavelet
f(t)	Original signal
D	Detailed coefficient
$\phi_j(t)$	Scaling function
$a_{M,k}$	Low frequency approximation coefficients
$d_{j,k}$	High frequency detailed coefficients
f_s	Sampling frequency
$\hat{\Phi}$	Predicted fault
$\hat{\overline{\Phi}}$	Predicted no-fault
Φ	Actual fault
$\overline{\Phi}$	Actual no-fault

normal supply is unavailable. They also provide numerous advantages against the load demand, reduced capital costs, and the reduction of the environmental pollution and global warming concerns by producing clean power through low carbon utilization technologies [9].

However, the fault feeding and control characteristics of DGs are different from the synchronous generators. Therefore, the conventional power system undergoes considerable changes when the microgrid is implemented, giving rise to protection problems [10]. The two main protection issues that microgrid must address are the bidirectional fault current flow and low fault current. Bidirectional power flow in feeders is in one direction for a fault on the system, and in the opposite direction for faults at the DER. Another problem associated with microgrid is the variable and large difference of the fault current levels when undergoing grid-isolated or grid-connected operations [11], [12], [13], [14]. This level may significantly drops after the disconnection of the main grid. Another reason for the low fault current is because of the large percentage of inverter-interfaced DGs (IIDGs) in microgrid, which produces an insufficient fault current due to their fault current limiting functions. Therefore, the performance of conventional over-current directional relays may fail to observe and respond to the fault, resulting in a failure or misdetection of a fault. This turns out to be even more problematic when the IIDGs penetration is high [15], [16].

Faults cause downtime of the whole power system and can severely damage the expensive equipment if not detected and cleared quickly. Selective phase tripping can be achieved through accurate fault detection, fault location, and fault phase information [17]. Due to low fault currents in an islanded mode, the traditional over-current protection schemes may lose the selectivity [18]. Hence, building a smarter and controlled protection scheme for operating a microgrid requires much attention towards planing and designing [19].

The literature on microgrid protection shows a variety of approaches to address protection issues. An artificial neural network (ANN) and discrete wavelet transform (DWT) protection technique is presented in [20]. A similar approach is provided in [21], using a combined decision tree (DT) and a wavelet transform (WT). An intelligent WT and deep neural networks based fault detection scheme is introduced in [22]. Another intelligent fault protection strategy for microgrids based on a convolutional neural network is reported in [23]. In [24], a seamless islanding and grid synchronization-based communication approach is presented for the protection and control of a microgrid. The authors in [25] presented a squaring and low-pass filtering method based on an autocorrelation function. Another WT and Renyi entropy technique for an islanded microgrid is presented in [26]. In [27] and [28], the authors described combined signal processing and machine learning schemes for radial distribution grid. A three stage approach based on real time data for protection of microgrid is introduced in [29]. A dual setting directional over-current relay based communication assisted scheme is proposed in [30]. The authors in [31] presented a quadrature and zero sequence components based adaptive protection scheme. In [32], an integrated impedance angle based PMU protection scheme is introduced using the positive sequence components. A centralized control approach, based on a hybrid hardware-software co-simulation platform to provide protection between the physical and cyber parts is proposed in [33]. A current-only polarity comparison inverter based microgrid protection method is provided in [34]. To address the issues of protection coordination, a non-pilot protection strategy of an inverter-dominated microgrid is introduced in [35]. A multiagent system-based co-simulation platform for microgrid protection is presented in [36]. In [37], the authors provided an isolated microgrid protection scheme based on real-time adaptive differential features.

The above-mentioned schemes protected the microgrid effectively, however, some studies did not consider both the grid-connected and islanded modes of operation [25], [36] and some did not consider both looped and radial configurations [28]. Moreover, some of the studies relied on IIDGs [34], [35]. The major drawback with an adaptive scheme is the computational burden [31], [37], and most of the previous studies did not take simultaneous faults into account. This paper introduces a fault protection scheme for microgrid using a combined signal processing and data mining approach. The statistical indices of total harmonic distortion (THD) and

negative sequence components are extracted to collected the input data-set. The obtained data-set is fed to the random forest (RF) classifier to build the data mining model to protect the microgrid against the fault. Additionally, the proposed scheme investigated for the validation of simultaneous faults and measurement noise. The proposed scheme is evaluated on MATLAB/SIMULINK and the analysis for testing and training of the RF is performed using Python. The contributions of this research are:

- To develop a multiresolution decomposition of the DWT and RF-based microgrid protection scheme.
- To compute the THD using the detailed coefficient of a three level multiresolution decomposition.
- To collect the data by extracting the standard deviation, mean, and median of THD and negative sequence components, and use this data-set to detect, identify, and classify the faults utilizing the RF classifier .
- To investigate the robustness of the proposed scheme against the measurement noise and simultaneous faults.
- To investigate the capability of the microgrid for both the grid-connected and islanded modes of operation and looped and radial configurations.

The remainder of the manuscript is structured into four sections. Section II introduces the proposed protection scheme with detailed descriptions of DWT and MRA decomposition. Section III outlines the designed model of the test system under study. Section IV is devoted to the findings and discussions. Finally, Section V reports the conclusion.

II. PROPOSED SCHEME

The traditional approaches for fault protection are prone to network disturbances caused by noise. Recent researches have focused on data mining for microgrid protection because data mining techniques, as opposed to the hard threshold, use a soft criterion for fault detection and have an outstanding capability of handling data with noise [38], [39], [40]. Hence, the proposed scheme employs the multiresolution analysis (MRA) of a DWT and RF. In the proposed method, the sampled current measurements are preprocessed by MRA of a DWT. Simplified schematic of the feature extraction is given in Fig. 1. Here, the detailed coefficients of the MRA of the DWT are used to compute the THD of voltage and current. Further, the negative sequence components of active and reactive power are extracted. After that, standard deviation, mean, and median of the THD and the negative sequence components are collected and fed to the RF for building the data mining model to detect, identify, and classify faults in a microgrid. To collect the data, a series of simulations is performed from different aspects. Data for fault events is collected by performing the simulations by changing the fault resistance, fault line, fault inception angle, and fault type. The measurement points are located on the buses ($B_2 - B_6$) for each distribution line in the power network. Additional data is collected by varying the load, capacitor switching, and DGs outages for no-fault cases. These simulations are conducted considering both modes under looped and radial configurations of the microgrid. Simulation events are depicted in Tables 2 and 3 of the manuscript.

A. WAVELET TRANSFORM (WT)

A Fourier transform (FT) is suitable for studying steady state problems and does not take into account the direct information of an oscillating signal. WT is a popular linear transformation signal processing technique developed in the $1980s$. WT is the advancement of FT. WT has the ability to decompose a signal into specific time-frequency domain characteristics that extract the hidden fault information of a signal [41]. WTs have several power system applications and are widely used to detect and classify power system disturbances. Generally, WT is divided into continuous wavelet transform (CWT), DWT, and wavelet packet transform [42], [43], [44], [45].

Since this study mainly deals with DWT, the following subsection gives the details of DWT.

B. DISCRETE WAVELET TRANSFORM (DWT)

The DWT is an extended version of the WT. DWT is easier to implement than CWT [20]. By using a scaling function $\phi_j(t)$ and the mother wavelet $\psi_j(t)$, a time series signal $f(t)$ can be decomposed into approximation and detailed coefficients as:

$$\phi_{jk}(t) = 2^{-\frac{j}{2}}\phi(2^{-j}t - n), \tag{1}$$

$$\psi_{jk}(t) = 2^{-\frac{j}{2}}\psi(2^{-j}t - n), \tag{2}$$

where, $n \in Z$, j and k are integers and the basis function is scaled by a factor of 2^j and translated by n units of time. The scaling function is associated with a low-pass filter, with filter coefficients $H = h(n)$. The wavelet function is associated with a high-pass filter, with filter coefficient $G = g(n)$ [21], [46], given as:

$$\phi(t) = \sum_n h(n)\sqrt{2}\phi(2t - n). \tag{3}$$

$$\psi(t) = \sum_n g(n)\sqrt{2}\phi(2t - n). \tag{4}$$

1) Wavelet multiresolution analysis (MRA) decomposition

MRA is a very useful implementation of DWT. Multiresolution signal decomposition theory was first introduced by Mallat as a mathematical model [20], [47]. The multiresolution decomposition of a time-varying signal $f(t)$ at level M is expressed as:

$$\begin{aligned} f(t) &= \sum_k a_{M,k}\frac{\psi}{\sqrt{2^M}}\left(\frac{t}{2^M} - k\right) \\ &+ \sum_k^M \sum_k d_{j,k}\frac{\psi}{\sqrt{2^j}}\left(\frac{t}{2^j} - k\right) \\ &\cong A_m(t) + \sum_j D_j(t), \end{aligned} \tag{5}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3088900, IEEE Access

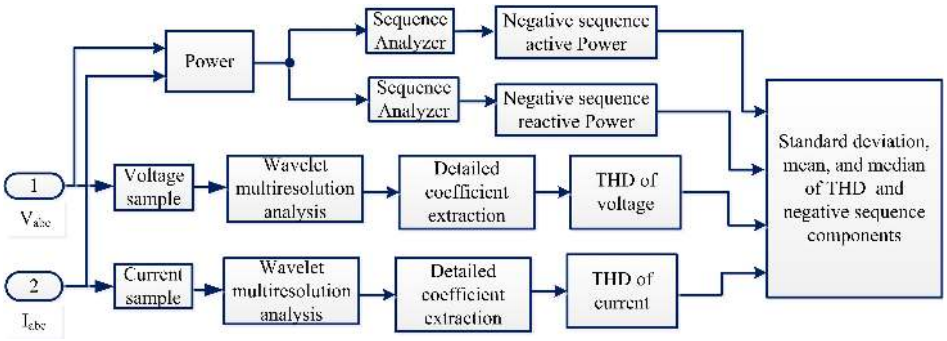

FIGURE 1: Block diagram of proposed scheme for feature extraction.

TABLE 2: Fault events simulating conditions.

Parameters	Fault Events	Counts
Operating modes	Grid-connected or islanded	2
Topologies	Radial or loop	2
Fault types	AG, BG, CG, ABG, BCG, CAG, AB, BC, CA, ABC, ABCG	11
Fault resistance (Ω)	0.01-100	10
Fault inception angle	$0°, 30°, 60°, 90°$	4
Fault line	$DL_1, DL_2, DL_3, DL_4, DL_5$	5
Total fault cases	8800	

TABLE 3: No-fault events simulating conditions.

Parameters	Counts
Operating modes (grid-connected or islanded)	2
Topologies (radial or loop)	2
Capacitor switching at load buses and PCC	6
Sudden load changes	6
DG_1 and DG_3 out	2
Total no-fault cases	288

TABLE 4: Detailed Coefficient frequency range.

Detailed coefficients	Frequency Range in (kHz)
D_1	1.8-0.9
D_2	0.9-0.45
D_3	0.45-0.225

where $a_{M,k}$ and $d_{j,k}$ are the low frequency approximation and high frequency detailed coefficients of the original signal at level M.

The DWT can be considered as a filter bank. Decomposition is performed by passing the sampled signal $f(t)$ through a low-pass filter $h(n)$ and a high-pass filter $g(n)$ to obtain the approximation and detailed coefficients. In the first step, the signal is decomposed into detailed coefficient (D_1) with frequency band of $(\frac{f_s}{2} - \frac{f_s}{4})$ kHz and approximation coefficient (A_1) with frequency band of $(\frac{f_s}{4} - 0)$ kHz. Then a factor of 2 is used to down sample the output of both filters. To repeat the procedure for further decomposition, the detailed coefficient is sent to the second stage to produce new coefficients. In the second decomposition step, D_2 collects the information between $(\frac{f_s}{4} - \frac{f_s}{8})$ kHz and A_2 is computed between $(\frac{f_s}{8} - 0)$ kHz. Similarly, D_3 captures the information between $(\frac{f_s}{8} - \frac{f_s}{16})$ kHz, and A_3 between 0 and $\frac{f_s}{16}$, where, f_s is the original signal sampling frequency. Decomposition of input signal is illustrated in Fig. 2. The process is repeated for the detailed coefficients until the desired level of detailed coefficients is retrieved [48], [47]. The number of decomposition steps is influenced by the sampling frequency which is 3.6 kHz for the proposed scheme. Information related to the detailed coefficients obtained using the wavelet decomposition is mentioned in Table 4.

III. TEST SYSTEM UNDER STUDY

The details of the test system of a microgrid are explained in this section. The PCC circuit breaker in the model is used to connect the microgrid and the utility, and it is also used to change the modes between grid and islanded. The base values used in this study are $25k$V, $15M$VA, and 60Hz. The test system is comprised of three IIDGs, (DER_1 and DER_3 each of $3M$VA and DGR_2 of $2M$VA) and one $7M$VA synchronous-based DG (DER_4). It comprises of five $20k$m line length distribution lines ($DL_1 - DL_5$). The purpose of Circuit Breaker CB loop-1 and Circuit Breaker CB loop-2 is to switch the microgrid between looped and radial configurations. There are six loads, one load is connected to each bus with the values given in Table 5. A $120/25k$V Dyn transformer is used to interconnect the microgrid and the main grid. All DG sources are connected through a $0.630/25k$V transformer. The test system model of the microgrid is presented in Fig. 3.

IV. RESULTS AND DISCUSSIONS

The complete flowchart of the proposed fault protection algorithm is shown in Fig. 4. To evaluate the capability of the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3088900, IEEE Access

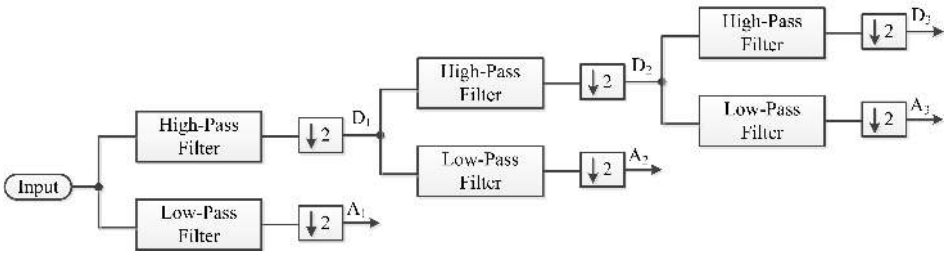

FIGURE 2: Decomposition of detailed coefficients.

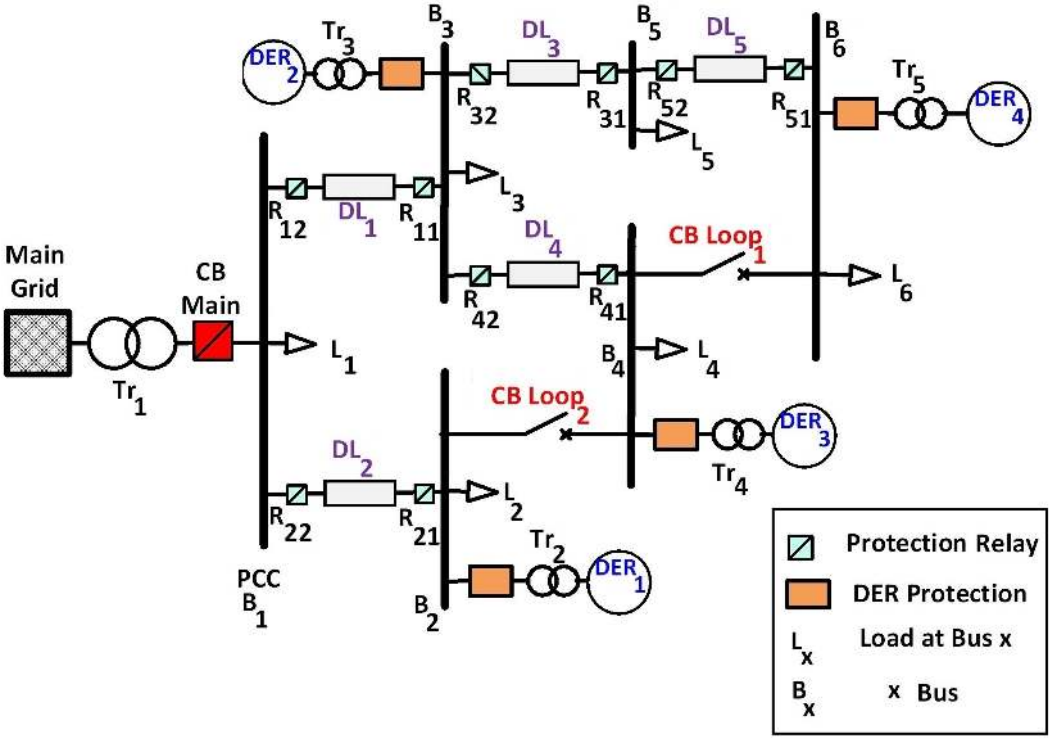

FIGURE 3: Microgrid test system.

TABLE 5: Loads data of the proposed test system.

Load	P (*M*W)	Q (*M*VAR)
L_1	3.0	1.00
L_2	3.0	1.00
L_3	4.0	1.50
L_4	1.0	0.75
L_5	1.0	0.75
L_6	1.0	0.50

proposed fault protection strategy, accuracy, precision, and recall are considered as the measurement indices.

1) *Accuracy* is used to compute the reliability of the proposed scheme for the total number of predicted and actual fault and no-fault events, given by:

$$A = \frac{(\hat{\Phi} + \hat{\bar{\Phi}})_T}{(\Phi + \bar{\Phi})_T}, \tag{6}$$

where, $\hat{\Phi}_T$ and Φ_T are the total number of predicted and actual faults, whereas, $\hat{\bar{\Phi}}_T$ and $\bar{\Phi}_T$ are the total number of predicted and the total number of actual no-fault events.

2) *Precision* describes the relationship between the predicted and actual fault events. It is given as follows:

$$P = \frac{\hat{\Phi}_T}{\Phi_T}, \tag{7}$$

where, $\hat{\Phi}_T$ and Φ_T denote the total number of predicted faults and the total number of actual fault events respectively.

3) *Recall* gives the relationship between the total number of predicted and actual no-fault events given as:

$$R = \frac{\hat{\bar{\Phi}}_T}{\bar{\Phi}_T}, \tag{8}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3088900, IEEE Access

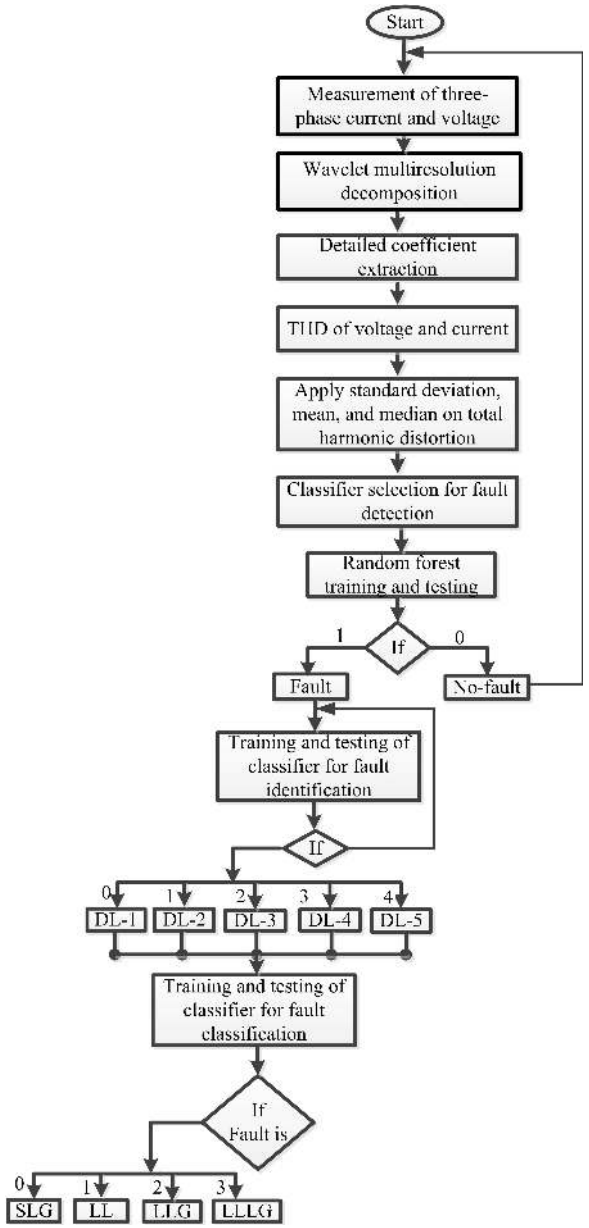

FIGURE 4: Flowchart of the proposed fault protection technique.

where, $\hat{\overline{\Phi}}_T$ and $\overline{\Phi}_T$ represent the total number of predicted no-fault and the total number of actual no-fault events, respectively.

A. DATA MINING MODEL FOR FAULT DETECTION

To build the data mining model for fault detection for the proposed scheme, the obtained data-set is divided into two parts with a proportion of 75% for training and 25% for testing. Values of 1 and 0 are assigned for fault and no-fault events, respectively. Once the training is done, the rest of the data-set is used for testing to investigate the performance through the RF to distinguish between fault and no-fault events. In this study, the total data-set consists of $9,088$ cases,

		Predicted	
	Total cases (2272)	No-fault	Fault
Actual	No-fault (69)	67	2
	Fault (2203)	0	2203

FIGURE 5: Confusion matrix for fault detection.

TABLE 6: Fault detection accuracy comparison.

Data mining method	Accuracy %	Precision %	Recall %
SVM	99.42	95.59	93.85
DT	99.47	100	83.56
Random forest	99.91	100	97.10

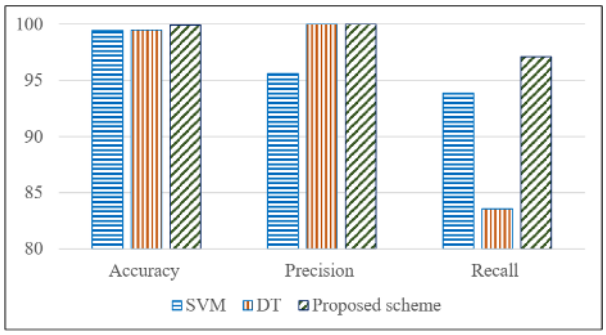

FIGURE 6: Comparison of accuracy measurement of RF with SVM and DT for fault detection.

where $8,800$ are fault events and 288 are no-fault cases, which are obtained as discussed in section II. Fig. 5 gives the details of the confusion matrix of actual versus predicted fault and no-fault cases. The proposed scheme classifier selected a total of 2272 cases randomly, out of which 69 were no-fault cases and $2,203$ were fault cases. The proposed scheme detected all of the $2,203$ fault cases correctly, whereas, out of 69 no-fault cases it detected 67 cases as no-faults, and only 2 cases as fault, with an overall accuracy of 99.91%. To verify this method, RF is compared with SVM and DT. Table 6 shows the performance of the proposed scheme with SVM and DT, and clearly shows better results than the other methods. It is observed that the proposed scheme has an accuracy of 99.91% (with $100.\%$ precision, and 97.10% recall) for fault detection, as compared to SVM which has an accuracy of 99.42% (with 95.59% precision and 93.85% recall), whereas DT has an accuracy of 99.47% (with 100% precision and 83.56% recall). The bar-chart representation of accuracy measurement of the proposed scheme with SVM and DT can be found in Fig. 6. The proposed scheme takes half cycle to detect the fault in microgrid.

B. PERFORMANCE OF PROPOSED SCHEME FOR FAULT IDENTIFICATION

Once the fault is detected, the next step is to identify the faulty section. For fault identification in the microgrid, training is done by assigning values from $0-4$ to all of the five distribution lines $(DL_1 - DL_5)$, respectively. After that,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3088900, IEEE Access

		Predicted				
	Total cases (2197)	DL-1	DL-2	DL-3	DL-4	DL-5
Actual	DL-1	440	5	0	0	5
	DL-2	4	451	3	0	0
	DL-3	2	3	425	2	0
	DL-4	0	0	3	394	7
	DL-5	3	0	2	4	444
	Accuracy	98.04%				

FIGURE 7: Confusion matrix for fault identification.

		Predicted				
	Total cases (2200)	SLG	LL	LLG	LLLG	Accuracy %
Actual	SLG	613	0	5	0	99.19
	LL	0	579	0	0	100
	LLG	9	0	581	0	98.48
	LLLG	4	0	0	409	99.03
	Overall accuracy	99.18%				

FIGURE 8: Confusion matrix for fault classification.

testing is done by feeding the input data to RF by considering the same cases as in Table 2. Fig. 7 shows the confusion matrix for fault identification of the proposed scheme. It can be seen that the proposed scheme can identify the faults in a microgrid with 98.04% accuracy.

C. DATA MINING MODEL FOR FAULT CLASSIFICATION

After fault detection and identification, RF is used to classify the fault. For fault classification 75% of the data-set is used for training and 25% is used for testing, similar to that for fault detection. The model generated for fault classification uses eleven fault types, single-line-to-ground faults (AG, BG, CG), line-to-line faults (AB, BC, CA), line-to-line-to-ground faults (ABG, BCG, CAG), three-phase faults (ABC), and three-phase-to-ground faults (ABCG). While training the input data-set, the SLG faults (AG, BG, CG) are assigned with a value 0. Similarly, a value of 1 is considered for LL (AB, BC, CA) faults, 2 is assigned for LLG (ABG, BCG, CAG) faults, and for LLL/LLLG faults the value of 3 is assumed. The confusion matrix for fault classification is presented in Fig. 8. It can be seen from the figure that the proposed scheme considered 2200 fault cases randomly. It classified 613 SLG fault cases correctly, with 5 misdetections, leading to 99.19% accuracy. The figure also shows that the proposed scheme classified all 579 LL faults cases correctly with 100% accuracy. For LLG faults, the proposed scheme detected 581 cases correctly, with 98.48% accuracy. Finally, for LLL/LLLG faults, 409 cases were classified correctly with 99.03% accuracy. However, the overall accuracy of the proposed scheme for fault classification is 99.18%. A comparison of the proposed scheme with SVM and DT is presented in Table 7. It can be seen that the accuracy of the proposed scheme is 99.18% compared to SVM and DT with accuracies of 95.36% and 96.09%, respectively. The bar-chart of the proposed fault classification in shown in Fig. 9. After fault detection, identification, and classification, a trip signal is generated to isolate the faulty part from rest of the system in one cycle.

TABLE 7: Comparison of fault classification accuracy with SVM and DT.

Data mining method	Accuracy %
SVM	95.36
DT	96.09
Proposed method	99.18

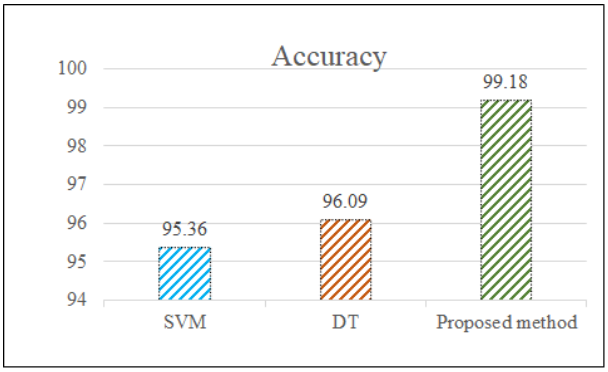

FIGURE 9: Comparison of accuracy measurement for fault classification.

D. PERFORMANCE OF PROPOSED SCHEME FOR GRID-CONNECTED AND ISLANDED MODE

The proposed scheme was also investigated for accuracy to differentiate between grid-connected and islanded modes of operation. While building the data mining model, a value of 1 is assigned for grid-connected and 0 is assumed for grid-isolation, considering the same cases as in Tables 2 and 3. It was found that the proposed scheme can effectively differentiate between grid-connected and islanded modes with 100% accuracy.

E. SIMULTANEOUS FAULTS DETECTION

Simultaneous faults are when two or more faults occur at the same time at different locations, and the protective relays may malfunction due to such faults. Therefore, it is necessary that the protection systems take into account simultaneous faults precisely while protecting the microgrid. Hence, the proposed scheme also considered simultaneous faults, and their detection is performed by considering the same cases as discussed in Section II. The proposed scheme considered two simultaneous faults F_1 (ABCG) on DL_1, and another fault F_2 (ABG) on DL_3 of the microgrid. Training and testing were performed in the same manner as for fault detection. While building the data mining model a value of 0 is assigned for no-faults and a value of 1 is assigned for simultaneous faults. The proposed scheme can effectively differentiate the simultaneous faults with 100% accuracy. DT also has an accuracy of 100% for simultaneous faults, whereas, SVM has an accuracy of 94.29%

F. IMPACT OF NOISE

The robustness of the proposed scheme was also evaluated under measurement noise. The simulations were conducted

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3088900, IEEE Access

TABLE 8: Effect of noise on fault detection under measurement noise.

Proposed method	Accuracy %	Precision%	Recall %
Without Noise	99.91	100	97.10
40dB (SNR)	99.82	100	93.85
30db (SNR)	99.69	100	89.85

TABLE 9: Comparison of the proposed scheme with SVM and DT under noise.

Data mining method	Without Noise	40dB (SNR)	30db (SNR)
SVM	99.42	97.14	96.96
DT	99.47	88.68	98.02
Random forest	99.91	99.82	99.69

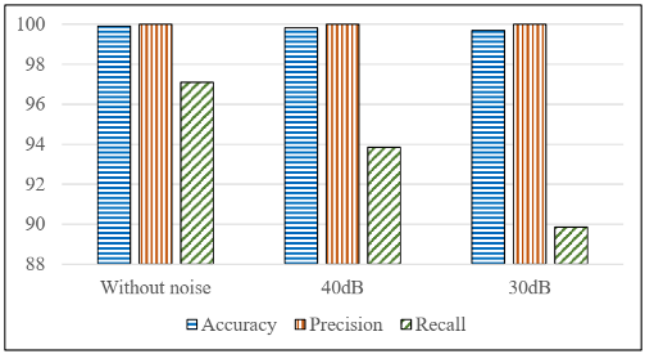

FIGURE 10: Accuracy measurement of proposed scheme for fault detection with and without noise.

by distorting the three-phase voltage measurement by adding white Gaussian noise with a 40dB and 30dB signal-to-noise ratio (SNR), [22]. The data collection is done with the distorted data by considering the cases as in Tables 2 and 3. After that, the data is fed to the RF for testing. Table 8 and Fig. 10 summarize the performance of the proposed scheme under noise. It can be seen from the table and figure that without the influence of noise the accuracy of the proposed scheme is 99.91%. The performance of the proposed scheme is also good under 40dB and 30dB noise with accuracies of 99.82% and 99.69%. Similarly, Table 9 and Fig. 11 provide the comparison of the proposed scheme accuracy with SVM and DT for fault detection under the influence of noise.

G. UNCERTAINTY OF DER

To enhance the performance of the proposed scheme, the simulations were performed with DERs uncertainty. The uncertainty is considered at 80%, 90%, and 100% of the rated capacity of the DERs. The data is collected accordingly by considering all the cases as in Tables 2 and 3 with limited simulations. The obtained data-set is trained and tested for the proposed algorithm. Fig. 12 shows the confusion matrix for fault detection under DERs uncertainty. The accuracy of the proposed scheme along with the SVM and DT are shown in Table 10 and Fig. 13. The simulation results show that the performance of the proposed scheme is efficient under DERs

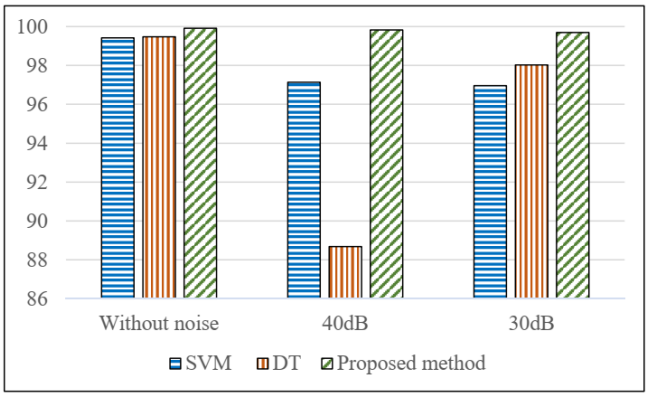

FIGURE 11: Accuracy comparison of the proposed scheme with SVM and DT under noise.

		Predicted	
	Total cases = 672	No-fault	Fault
Actual	No-fault (78)	74	4
	Fault (594)	1	593

FIGURE 12: Confusion matrix for fault detection under DERs uncertainty.

uncertainty.

TABLE 10: Fault detection accuracy comparison under DERs uncertainty.

Data mining method	Accuracy %	Precision %	Recall %
SVM	96.28	96.63	93.59
DT	99.11	99.66	94.87
Random forest	99.25	99.83	94.87

In addition, the time and space complexity are used to validate the computational complexity of the proposed scheme. The proposed scheme was implemented using the Python framework in intel i7-10700 CPU. The total response time and space complexity are tabulated in Table 11 for the proposed scheme, DT, and SVM. The training and testing time of

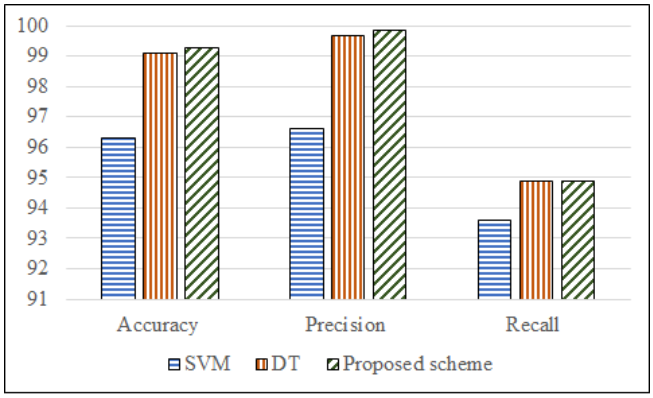

FIGURE 13: Fault detection accuracy comparison under DERs uncertainty.

TABLE 11: Time and space complexity.

Data mining model	Time complexity Training and testing (ms)	Space complexity MiB
DT	9.138	146.218
SVM	1783.597	146.347
Proposed scheme	110.818	144.789

RF is 110.818 *m*s, out of which response time or predication time is 6.46 *m*s. This makes the proposed scheme efficient having good detection, identification, and classification rates while requiring less time and space complexity. This reduces the complexity of fetching the proposed scheme into reality. In [21], [49], the authors performed the real time validation using OPAL-RT setup based on signal processing and Data mining techniques. The response time in [21] is ranging from 25 to 41.7 *m*s and less than 16 *m*s in [49], which is similar or more than our proposed scheme i.e. 6.46 *m*s. The OPAL-RT digital simulator is compatible with MATLAB/Simulink and Python applications, making it easier for the proposed scheme in interfacing with real time simulation.

Table 12 shows the comparison of the proposed schemes with other existing techniques.

V. CONCLUSION

This paper considered a fault protection scheme for microgrid based on signal processing and data mining. The proposed scheme preprocessed the three-phase voltage and current signals by a multiresolution analysis of discrete wavelets transform. Statistical indices of total harmonic distortion and negative sequence components are collected and fed into random forest for fault protection. The simulations were conducted in the MATLAB/SIMULINK software for both grid-connected and grid-isolated modes, and looped and radial configurations. The results demonstrates that the proposed scheme is capable to detect, identify, and classify the faults in microgrid. In addition, it can perform very well for simultaneous faults and under the influence of measurement noise. The results confirm that the proposed scheme is more accurate for fault protection as compared to decision tree and support vector machines.

REFERENCES

[1] P. K. Ray, A. Mohanty, B. K. Panigrahi, and P. K. Rout, "Modified wavelet transform based fault analysis in a solar photovoltaic system," Optik, vol. 168, pp. 754–763, 2018.

[2] S. Ram Ola, A. Saraswat, S. K. Goyal, V. Sharma, B. Khan, O. P. Mahela, H. H. Alhelou, and P. Siano, "Alienation coefficient and wigner distribution function based protection scheme for hybrid power system network with renewable energy penetration," Energies, vol. 13, no. 5, p. 1120, 2020.

[3] P. Tian and L. Zhang, "Big data mining based coordinated control discrete algorithm of independent micro grid with pv and energy," Microprocessors and Microsystems, vol. 82, p. 103808, 2021.

[4] S. Zhang, M. Pu, B. Wang, and B. Dong, "A privacy protection scheme of microgrid direct electricity transaction based on consortium blockchain and continuous double auction," IEEE Access, vol. 7, pp. 151 746–151 753, 2019.

[5] A. Shabani and K. Mazlumi, "Evaluation of a communication-assisted overcurrent protection scheme for photovoltaic-based dc microgrid," IEEE Transactions on Smart Grid, vol. 11, no. 1, pp. 429–439, 2020.

[6] I. Series, "Microgrids and active distribution networks," The institution of Engineering and Technology, 2009.

[7] B. Patnaik, M. Mishra, R. C. Bansal, and R. K. Jena, "Modwt-xgboost based smart energy solution for fault detection and classification in a smart microgrid," Applied Energy, vol. 285, p. 116457, 2021.

[8] J. Marín-Quintero, C. Orozco-Henao, W. Percybrooks, J. C. Vélez, O. D. Montoya, and W. Gil-González, "Toward an adaptive protection scheme in active distribution networks: Intelligent approach fault detector," Applied Soft Computing, vol. 98, p. 106839, 2021.

[9] S. Baloch and M. S. Muhammad, "An intelligent data mining-based fault detection and classification strategy for microgrid," IEEE Access, vol. 9, pp. 22 470–22 479, 2021.

[10] A. B. Nassif, "A protection and grounding strategy for integrating inverter-based distributed energy resources in an isolated microgrid," CPSS Transactions on Power Electronics and Applications, vol. 5, no. 3, pp. 242–250, 2020.

[11] A. Kulshrestha, O. P. Mahela, M. K. Gupta, N. Gupta, N. Patel, T. Senjyu, M. S. S. Danish, and M. Khosravy, "A hybrid fault recognition algorithm using stockwell transform and wigner distribution function for power system network with solar energy penetration," Energies, vol. 13, no. 14, p. 3519, 2020.

[12] S. Ram Ola, A. Saraswat, S. K. Goyal, S. Jhajharia, B. Khan, O. P. Mahela, H. Haes Alhelou, and P. Siano, "A protection scheme for a power system with solar energy penetration," Applied Sciences, vol. 10, no. 4, p. 1516, 2020.

[13] T. S. Aghdam, H. K. Karegar, and H. H. Zeineldin, "Variable tripping time differential protection for microgrids considering dg stability," IEEE Transactions on Smart Grid, vol. 10, no. 3, pp. 2407–2415, 2019.

[14] T. Gush, S. B. A. Bukhari, K. K. Mehmood, S. Admasie, J.-S. Kim, and C.-H. Kim, "Intelligent fault classification and location identification method for microgrids using discrete orthonormal stockwell transform-based optimized multi-kernel extreme learning machine," Energies, vol. 12, no. 23, p. 4504, 2019.

[15] D. A. Gadanayak, "Protection algorithms of microgrids with inverter interfaced distributed generation units—a review," Electric Power Systems Research, vol. 192, p. 106986, 2021.

[16] K. Allahdadi, I. Sadeghkhani, and B. Fani, "Protection of converter-interfaced microgrids using modified short-time correlation transform," IEEE Systems Journal, vol. 14, no. 4, pp. 5172–5175, 2020.

[17] O. P. Mahela, J. Sharma, B. Kumar, B. Khan, and H. H. Alhelou, "An algorithm for the protection of distribution feeder using stockwell and hilbert transforms supported features," CSEE Journal of Power and Energy Systems, 2020.

[18] Z. Liu, C. Su, H. Høidalen, and Z. Chen, "A multiagent system-based protection and control scheme for distribution system with distributed-generation integration," IEEE transactions on power delivery, vol. 32, no. 1, pp. 536–545, 2016.

[19] J. Nsengiyaremye, B. C. Pal, and M. M. Begovic, "Microgrid protection using low-cost communication systems," IEEE Transactions on Power Delivery, 2020.

[20] Y.-Y. Hong and M. T. A. M. Cabatac, "Fault detection, classification, and location by static switch in microgrids using wavelet transform and taguchi-based artificial neural network," IEEE Systems Journal, 2019.

[21] D. P. Mishra, S. R. Samantaray, and G. Joos, "A combined wavelet and data-mining based intelligent protection scheme for microgrid," IEEE Transactions on Smart Grid, vol. 7, no. 5, pp. 2295–2304, 2015.

[22] J. James, Y. Hou, A. Y. Lam, and V. O. 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.

[23] S. B. A. Bukhari, C.-H. Kim, K. K. Mehmood, R. Haider, and M. S. U. Zaman, "Convolutional neural network-based intelligent protection strategy for microgrids," IET Generation, Transmission & Distribution, vol. 14, no. 7, pp. 1177–1185, 2020.

[24] A. Vukojevic and S. Lukic, "Microgrid protection and control schemes for seamless transition to island and grid synchronization," IEEE Transactions on Smart Grid, vol. 11, no. 4, pp. 2845–2855, 2020.

[25] S. Baloch, S. Z. Jamali, K. K. Mehmood, S. B. A. Bukhari, M. S. U. Zaman, A. Hussain, and C.-H. Kim, "Microgrid protection strategy based on the autocorrelation of current envelopes using the squaring and low-pass filtering method," Energies, vol. 13, no. 9, p. 2350, 2020.

[26] Y. Wang, J. Ravishankar, and T. Phung, "Wavelet transform-based feature extraction for detection and classification of disturbances in an islanded micro-grid," IET Generation, Transmission & Distribution, vol. 13, no. 11, pp. 2077–2087, 2018.

TABLE 12: Comparison of the proposed scheme with existing techniques.

S.No	Characteristics	Existing schemes		Proposed schemes	
1.	Operating modes [26] [37]	Islanded mode only		Both modes	
2.	Configurations [27]	Radial configuration only		Both looped and radial configurations	
3.	DGs type [34] [35]	IIDGs only		IIDGs and synchronous-based DGs	
4.	Simultaneous faults	None of the previous studies considered simultaneous fault		Considered the simultaneous faults	
		Detection %	Classification %	Detection %	Classification %
5.	Accuracy [22] [9]	99.31 99.29	97.92 98.57	99.91	99.18

[27] Y. D. Mamuya, Y.-D. Lee, J.-W. Shen, and C.-C. Kuo, "Application of machine learning for fault classification and location in a radial distribution grid," Applied Sciences, vol. 10, no. 14, p. 4965, 2020.

[28] M. Shafiullah, M. A. Abido, and Z. Al-Hamouz, "Wavelet-based extreme learning machine for distribution grid fault location," IET Generation, Transmission & Distribution, vol. 11, no. 17, pp. 4256–4263, 2017.

[29] S. Teimourzadeh, F. Aminifar, M. Davarpanah, and M. Shahidehpour, "Adaptive protection for preserving microgrid security," IEEE Transactions on Smart Grid, vol. 10, no. 1, pp. 592–600, 2017.

[30] H. M. Sharaf, H. H. Zeineldin, and E. El-Saadany, "Protection coordination for microgrids with grid-connected and islanded capabilities using communication assisted dual setting directional overcurrent relays," IEEE Transactions on Smart Grid, vol. 9, no. 1, pp. 143–151, 2016.

[31] M. Singh and P. Basak, "Adaptive protection methodology in microgrid for fault location and nature detection using q0 components of fault current," IET Generation, Transmission & Distribution, vol. 13, no. 6, pp. 760–769, 2018.

[32] N. K. Sharma and S. R. Samantaray, "Pmu assisted integrated impedance angle-based microgrid protection scheme," IEEE Transactions on Power Delivery, vol. 35, no. 1, pp. 183–193, 2019.

[33] H. F. Habib, N. Fawzy, M. M. Esfahani, O. A. Mohammed, and S. Brahma, "An enhancement of protection strategy for distribution network using the communication protocols," IEEE Transactions on Industry Applications, vol. 56, no. 2, pp. 1240–1249, 2020.

[34] B. Wang and L. Jing, "A protection method for inverter-based microgrid using current-only polarity comparison," Journal of Modern Power Systems and Clean Energy, vol. 8, no. 3, pp. 446–453, 2019.

[35] H. Lahiji, F. B. Ajaei, and R. Boudreau, "Non-pilot protection of the inverter-dominated microgrid," IEEE Access, vol. 7, pp. 142 190–142 202, 2019.

[36] H. F. Habib, M. M. Esfahani, and O. A. Mohammed, "Investigation of protection strategy for microgrid system using lithium-ion battery during islanding," IEEE Transactions on Industry Applications, vol. 55, no. 4, pp. 3411–3420, 2019.

[37] A. Arunan, T. Sirojan, J. Ravishankar, and E. Ambikairajah, "Real-time adaptive differential feature-based protection scheme for isolated microgrids using edge computing," IEEE Systems Journal, 2020.

[38] S. Jamali, S. Ranjbar, and A. Bahmanyar, "Identification of faulted line section in microgrids using data mining method based on feature discretisation," International Transactions on Electrical Energy Systems, vol. 30, no. 6, p. e12353, 2020.

[39] S. C. Shekar, G. R. Kumar, and S. Lalitha, "Wavelet based transient fault detection and analysis of microgrid connected power system," Int. J. Power Syst, vol. 1, pp. 46–52, 2016.

[40] D. Jayamaha, N. Lidula, and A. D. Rajapakse, "Wavelet-multi resolution analysis based ann architecture for fault detection and localization in dc microgrids," IEEE Access, vol. 7, pp. 145 371–145 384, 2019.

[41] J.-C. Yin, A. N. Perakis, and N. Wang, "An ensemble real-time tidal level prediction mechanism using multiresolution wavelet decomposition method," IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 8, pp. 4856–4865, 2018.

[42] L. Yuan-yuan and L. Mao-jun, "Detecting methods of harmonic in power system based on wavelet transform," in 2006 International Conference on Power System Technology. IEEE, 2006, pp. 1–4.

[43] V. Thiyagarajan and N. Subramanium, "Analysis and estimation of harmonics using wavelet technique," TELKOMNIKA Indonesian Journal of Electrical Engineering, vol. 13, no. 2, pp. 305–313, 2015.

[44] O. Chaari, M. Meunier, and F. Brouaye, "Wavelets: A new tool for the resonant grounded power distribution systems relaying," IEEE Transactions on Power Delivery, vol. 11, no. 3, pp. 1301–1308, 1996.

[45] G. S. Chawda, A. G. Shaik, M. Shaik, S. Padmanaban, J. B. Holm-Nielsen, O. P. Mahela, and P. Kaliannan, "Comprehensive review on detection and classification of power quality disturbances in utility grid with renewable energy penetration," IEEE Access, vol. 8, pp. 146 807–146 830, 2020.

[46] B. Chen, Y. Li, and N. Zeng, "Centralized wavelet multiresolution for exact translation invariant processing of ecg signals," IEEE Access, vol. 7, pp. 42 322–42 330, 2019.

[47] S. K. G. MukeshThakre and M. K. Mishra, "Distribution system faults classification and location based on wavelet transform," International Journal on Advanced Computer Theory and Engineering, vol. 2, no. 4, pp. 2319–2526, 2013.

[48] W. K. Ngui, M. S. Leong, L. M. Hee, and A. M. Abdelrhman, "Wavelet analysis: mother wavelet selection methods," in Applied mechanics and materials, vol. 393. Trans Tech Publ, 2013, pp. 953–958.

[49] M. Manohar, E. Koley, and S. Ghosh, "Microgrid protection under wind speed intermittency using extreme learning machine," Computers & Electrical Engineering, vol. 72, pp. 369–382, 2018.