Point-On-Wave-based Anomaly Detection and Categorization in Low Inertia Power Systems

He Yin, *Senior Member, IEEE,* Yuru Wu, *Student Member, IEEE,* Wenpeng Yu, *Member, IEEE* Jin Tan, *Senior Member, IEEE,* Andy Hoke, *Senior Member, IEEE,* Cameron J. Kruse, Brad W. Rockwell, Yilu Liu, *Fellow, IEEE*

***Abstract*—Inverter-based resources (IBRs), such as Photovoltaic (PV) and battery storage systems, are widely deployed to achieve carbon-free power systems. However, the anomalies, such as waveform distortions, and wide-band oscillations, caused by the IBRs have brought challenges to situational awareness, especially in the low inertia power systems. The anomalies such as generation trips and load shedding also have higher frequency deviation and voltage magnitude drops caused by the changing of inertia. To achieve effective anomaly identification, this paper proposed a point-on-wave (POW) based algorithm utilizing the real-time POW measurements from synchronized measurement units (SMUs). Four SMUs are specially designed and deployed on Hawaii islands to receive instantaneous POW measurements. Then, different physical characteristics, as well as statistical features are extracted from POW measurements to filter the anomalies. The anomaly identification approach based on the random forest is developed and deployed into the FNET/GridEye system considering the trade-offs among accuracy, computational burden, and deployment cost. To verify the performance of the proposed algorithm, different experiments are carried out with collected field test data. The result demonstrates that the performance of the proposed POW-based anomaly categorization algorithm can reach 94.54% which has comparable performance among benchmarking algorithms.**

***Index Terms*—Anomaly identification, point-on-wave measurement, synchronized measurement unit, low inertia power systems**

I. Introduction

THANKS to the fast development of distributed energy resources (DERs) and smart inverter technologies, measurement dynamics widely exist in distribution level power grids [1]. These dynamics have brought challenges to accurate anomaly detection, categorization, and situational awareness [2]. Some dynamic phenomena, including large voltage magnitude changes, fast frequency drops, and waveform distortions, are caused by the low inertia in the power system [3]. The synchronized measurement units (SMUs) deployed in distributed grid contribute to the dynamic information collection of the sensors [4]. However, the successful situational awareness suffer from varied data anomalies.

In this regard, situational awareness, such as disturbance identification and localization in a low inertia power system, is more complicated than that in a large-scale power system which places a requirement on the high-density synchrophasor and waveform-related feature measurements. Furthermore, the Point on the Wave (POW) measurements [5] can acquire panorama features during the anomaly, which enables multiple critical applications for Inverter Based Resources (IBR)-heavy feeders, including protection system analysis, capturing electromagnetic transients and sub/super- synchronous resonance, DER and flexible AC transmission system (FACTS) model improvement, as well as power quality studies. Additionally, POW measurements are also able to provide robustness and algorithm-free synchrophasor measurements. Considering the abovementioned advantages, effective anomaly identification solutions can be further developed to achieve situational awareness in the low inertia power system.

Most of the existing real-time anomaly detection algorithms are based on the synchrophasors from SMUs. By collecting real-time frequency streaming, the wide-area monitoring system is able to detect and categorize anomalies such as generation trips, line outages, load shedding, and low-frequency oscillations in the past decades [6]. In addition to the frequency measurements, the root mean square (RMS) values of voltage magnitude are also widely used in anomaly identification algorithms. For example, [7] monitors the load switching statuses by using voltage magnitude RMS values. An off-line anomaly categorization algorithm is proposed using voltage magnitude measurements while considering cyberattacks [8]. Moreover, POW measurements have been converted into RMS values of voltage magnitudes in [9] to identify and categorize the anomaly within an operational smart power grid. High-dimensional synchrophasors can provide real-time anomaly identification in [10]. In addition, harmonics and other power quality factors are also utilized for anomaly detection such as using Stockwell transform [11] and its related features [12]. To identify the anomaly in a low inertia power system, a straightforward methodology is developed by checking the waveform difference between adjacent cycles. However, off-nominal frequency measurements can trigger this POW-based detector with complicated real-world cases.

Currently, anomaly categorization algorithms primarily belong to data-driven-based methods. To achieve a fast anomaly diagnosis, multiple data-driven-based anomaly categorization algorithms have been proposed by using artificial intelligence methodology. For instance, [13] proposed a task-data-driven and physic-based model targeting real-time dynamic anomaly identification. Neural-network-based algorithms, including deep neural network [14] and convolution neural network [15], have been proposed to categorize anomalies targeting generation trips and load shedding due to their strong learning ability. Additionally, a k-nearest neighborhood (kNN) based algorithm [16] has been proposed by using time and frequency domain features to categorize anomalies. However, the neural network approaches have an enormous demand for the number of samples. And the black-box property lacks

the participation of physical mechanisms thus limiting their applications.

In contrast, data-driven based methods that combine physical properties and learning capabilities provide low-cost and easy-to-implement solutions, such as decision-tree-based [17] and random-forest-based algorithms [18]. For instance, the POW-based detector named a delta-delta-based anomaly detector is designed to accurately detect anomalies in a low inertia island power system [19]. Importantly, only 2% of the POW measurements during anomalies have been analyzed by power utilities [9] because its ultra-high data resolution limits its efficiency. Therefore, the physical-based methods could effectively filter the anomalies from multiple SMUs.

Taking the panoramic features from real-time POW measurements, the waveform information can be fully utilized for situational awareness in distribution-level power systems. The first national-level wide-area measurement system, FNET/GridEye, which is currently operated by the University of Tennessee, Knoxville (UTK) and Oak Ridge National Laboratory (ORNL) [20], had upgraded its SMU by using POW-based Universal Grid Analyzers (UGAs) [21]. Four POW-based UGAs have been deployed on Kauai island which contains a typical low inertia island power grid [22]. The average proportion of DERs is higher than 38% and the maximum instantaneous value can reach 60% recently in 2021 [23]. Overall, the efficient filtering and identification of the anomalies are critical as a result of the high data density of the POW measurement.

To achieve effective anomaly identification and further improve situational awareness capability, this paper proposes a real-time POW-based anomaly detection algorithm. The physical information such as the rate of change of frequency (RoCoF) and RMS are first calculated, and then it combines the power density threshold and data-driven-based method to achieve real-time anomaly categorization. The primary contributions are summarized below

1) A real-time POW-based anomaly detection algorithm is proposed to capture waveform distortions. The algorithm is implemented based on the recursive form to reduce the calculation burden. Importantly, the proposed algorithm has been deployed on the FNET/GridEye system for real-time anomaly detection purposes.
2) Four POW-based SMUs are deployed in Hawaii islands and one-year POW measurements are collected for anomaly categorization in a low inertia power system. An anomaly categorization algorithm based on random forest and the Bayesian optimization are designed considering POW, frequency, RMS values of voltage magnitudes, and power spectral densities.
3) A POW-based anomaly library has been built up targeting low-inertia and high PV-penetration power grids. The anomalies have been confirmed with local utility companies on Hawaii islands.
4) Multiple experiments have been carried out to verify the performance of the proposed anomaly categorization algorithm, where the results demonstrate its profound performance compared with other advanced algorithms.

TABLE I
SPECIFICATIONS OF THE POW-BASED UGA

Parameter names	Values
POW reporting rate	1440 Hz
Frequency reporting rate	10 Hz
Synchrophasor reporting rate	10 Hz
GPS information	satellite number, locations
Communication protocol	IEEE C37.118.2
Frequency accuracy	$\pm$0.00015 Hz
Phase angle accuracy	$\pm$0.005 degree

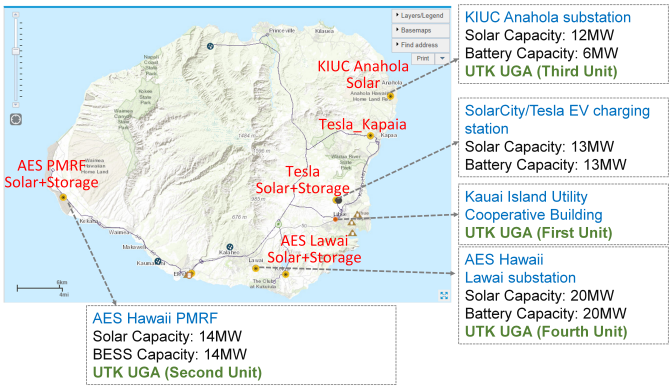

Fig. 1. The UGA deployment layout on Kauai Island. Source: https://kiuc.coop/renewable-portfolio

The rest of the paper is organized as follows. Section II presents POW-based SMU designation. Then, the POW-based anomaly detection and categorization algorithms are presented in Section III and IV, respectively. The experiment is presented in Section V. The conclusion is drawn in Section VI.

II. POW-BASED SMU DEPLOYMENT

The POW-based UGA is a specially designed SMU that targets measuring the high-resolution synchronized POW [21], where its indexes are listed in Table I. The specifications of the POW-based UGA show a strong capability of monitoring waveform distortions with a time resolution of 0.69ms and an oscillation frequency ranging from DC to 720 Hz. Notice that the frequency and synchrophasors are also available from POW-based UGAs to fulfill the IEEE C37.118.2 [24] requirement.

To detect anomalies from IBRs, four POW-based UGAs have been deployed on Kauai Island. As illustrated in Fig. 1, four POW-based UGAs are deployed in the Anahola substation, Kauai Island Utility Cooperative (KIUC) building, the Lawai substation, and the AES Hawaii Pacific Missile Range Facility (PMRF), respectively, where AES Hawaii and KIUC are local utility companies. The reason for choosing these

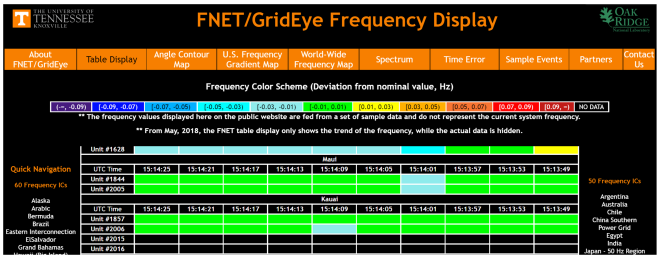

Fig. 2. The POW-based UGA data streaming from FNET public website. Source: https://fnetpublic.utk.edu/tabledisplay.html

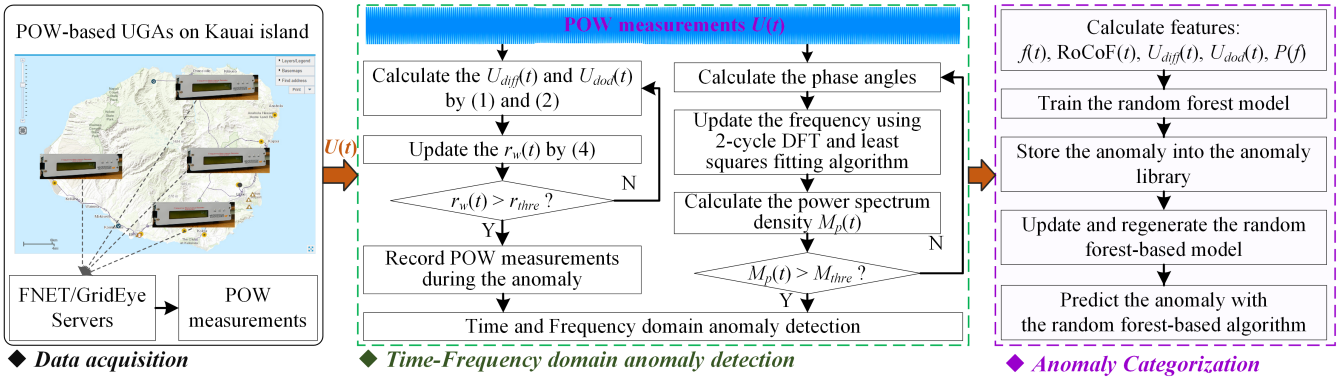

Fig. 3. The overall framework of the proposed anomaly detection and categorization algorithm.

locations is that there exist multiple IBRs aside from these locations. It also demonstrates that the UGA is close to solar utilities and battery storage systems, such as Tesla electric vehicle (EV) charging stations. The waveform distortions, harmonics, and oscillations aside from IBRs such as battery energy storage systems and photovoltaic systems are sensitive to the dynamic fluctuations for both power generations and loads.

The real-time POW measurements of these four UGAs will be streamed back to the FNET/GridEye servers. The FNET/GridEye servers provide online frequency monitoring and application services [25]. The four POW-based UGAs on Kauai island are illustrated in Fig. 2, where the color bars represent the frequency measurement interval from the units [26]. Meanwhile, the POW measurements will be calculated into anomaly detectors and categorization triggers.

III. POW BASED ANOMALY DETECTION ALGORITHM

A. Anomaly detection and categorization framework

The overall framework of the proposed anomaly detection and categorization algorithm is illustrated in Fig. 3, where the framework contains three stages.

1) *Data acquisition*: POW measurements are collected by UGAs and then will be firstly streamed back to the FNET/GridEye server with data preprocessing including data conditioning and time alignment.
2) *Anomaly detection*: In the time domain, the periodic waveform detector is designed to compare the integration of the two adjacent cycles' voltage magnitude values. Then the time and frequency domain anomalies are detected based on the threshold.
3) *Anomaly categorization*: The physical features are extracted including the POW, frequency, RMS values of voltage magnitudes, and power spectral densities. Then, the anomaly categorization algorithm based on the random forest is designed.

B. Time domain based Periodic waveform detector

To achieve anomaly detection, the location of the distinctive waveform is required to be detected in the time domain first. Here, the phase and voltage magnitude step changes cost by DERs are first detected. Its basic idea is to calculate the voltage magnitude difference, U_{diff}, between two adjacent cycles, which can be calculated as

$$U_{diff}(t) = U(t+N) - U(t) \tag{1}$$

where t is the sampling time stamp, N is the number of sample per cycle, U_{t+N} and U_t are the voltage magnitude values at timestamp $t+N$ and t.

The magnitude difference will be obtained according to the periodical delta value. However, the voltage magnitude difference between two adjacent cycles depends on a fixed sampling instant so the off-nominal frequency cases will have serious influences on this detector.

To improve the performance of the periodic waveform detector, a delta-of-delta method is utilized by calculating the waveform measurement differences twice. Thereafter, the off-nominal frequency and periodic characteristics of the $U_{diff}(t)$ can be removed. The delta-of-delta voltage at timestamp t can be calculated as,

$$U_{dod}(t) = U_{diff}(t+N) - U_{diff}(t) \tag{2}$$

Then the periodic waveform detector $r_w(t)$ can be expressed as,

$$r_w(t) = \sum_{n=0}^{N-1} |(U_{dod}(t+n))| \tag{3}$$

To achieve real-time detection, the $r_w(t)$ will be calculated once new sampling data come in by using a recursive method,

$$r_w(t+1) = r_w(t) - |(U_{dod}(t))| + |(U_{dod}(t+n+1))| \tag{4}$$

Based on the detector $r_w(t)$, a threshold is designed to detect anomalies based on the distribution of $r_w(t)$. Since it is the first round of anomaly detection, a top 5% is chosen as the threshold value for filtering the anomalies in this research.

C. Frequency domain based Oscillation detector

In addition to the time domain detection, the oscillations and harmonics generated from DERs can be another sort of interesting anomaly. As a result of the streaming POW measurements, arbitrary frequency estimation algorithms can be utilized in the server rather than in the SMU. In this paper, an ultra-fast frequency estimation algorithm with a low calculation burden is deployed on the server side for real-time frequency estimation and a high reporting rate.

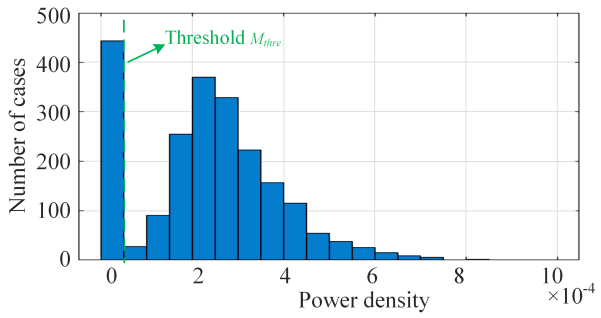

Fig. 4. The power density of the oscillation cases.

A 2-cycle DFT-based frequency estimation algorithm is used to estimate the frequency with a recursive solution. First of all, the phase angle at timestamp t, φ_t, are calculated by using the POW measurements through Discrete Fourier Transform based methods. Then the frequency at timestamp t can be calculated by using least squares fitting technology with a quadratic polynomial function as $f(t)$. The calculation burden can be further reduced by deploying this algorithm on the FNET/GridEye server, which is also discussed in [21]. This process is illustrated in Fig. 3.

Finally, by using a five-second frequency measurement window, a fast Fourier transform (FFT) is utilized to calculate the power spectral density of the frequency measurements in a frequency spectrum, which is denoted as $M_p(t)$. Similarly, the power density larger than M_{thre} at a single frequency will be recorded where M_{thre} is designed as 5% largest oscillation magnitude anomalies.

Here, the power density calculated from one-day data is illustrated in Fig. 4. It provides insight into how to select the suitable threshold M_{thre}. In the last step, the anomalies captured from both time and frequency domains will be gathered and utilized for anomaly categorization purposes.

At last, all anomalies triggered by both the time-domain-based periodic waveform detector and the frequency-domain-based oscillation detector will be logged and then transferred into the anomaly library. Through confirming with the local utility companies, the anomaly labels will be tagged by comparing the anomaly library to their outage lists.

IV. POW Based Anomaly Categorization Algorithm

A. Extracted physical features

After the anomaly detection, the 5-second POW measurements window is collected to achieve anomaly categorization. Based on the real-time POW measurements, different typical physical features from both time and frequency domains are extracted.

In this section, five types of physical features are calculated. First of all, the $U_{diff}(t)$ and $U_{dod}(t)$ are utilized as two waveform features focusing on the waveform distortion features. To distinguish the typical events including generation trip, line outage, and load shedding anomalies, the frequency, $f(t)$, and RoCoF, $RoCoF(t)$ are extracted since obvious frequency drop or increase can be observed during these anomalies. The RoCoFs are calculated by using two adjacent frequency measurements as a result of the fast frequency changes.

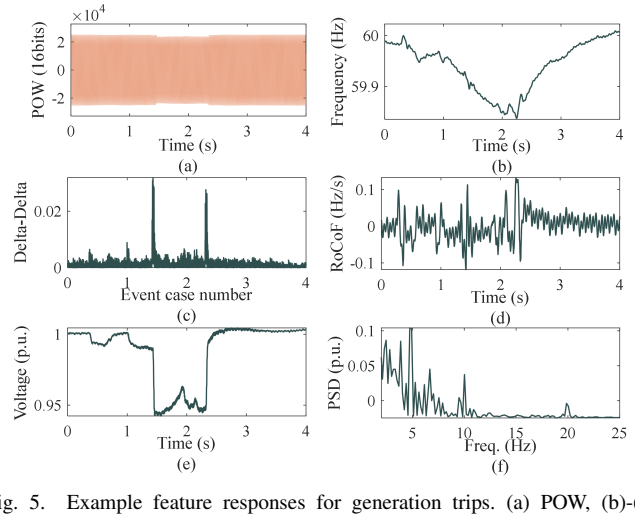

Fig. 5. Example feature responses for generation trips. (a) POW, (b)-(f) frequency, delta-delta, RoCoF, RMS of voltage, power spectral density (PSD).

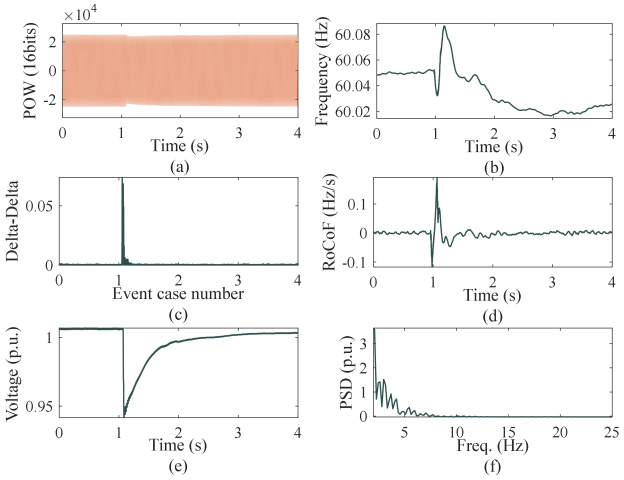

Fig. 6. Example feature responses for faults. (a) POW, (b)-(f) frequency, delta-delta, RoCoF, RMS of voltage, PSD.

Additionally, considering that the voltage magnitude $V_{RMS}(t)$ of the low inertia grid will drop or increase heavily during the anomaly including faults, generation trip, and load shedding, the root mean square (RMS) values are further extracted. Notice that this is not observable in a large power system due to the voltage regulators and capacitors deployed. Thus, the $V_{RMS}(t)$ is used as a feature targeting a low inertia power system. Besides, to avoid misclassification, the frequency spectrum $P(f)$ will be used to represent the feature of oscillations and harmonics in the frequency domain.

B. Feature Visualization Analysis

To achieve the automatic anomaly categorization in the FNET server, the existing anomalies should be labeled manually first according to the outage list from the local utility companies. Based on the collected POW measurements, the anomalies include oscillations, generation trips, load sheddings, voltage magnitude drops, waveform distortions, faults, switching, and others are observed.

Six example anomalies from POW, frequency, POW delta-delta, RoCoF, RMS value of voltage, and power spectral density responses are illustrated in Fig. 5 to Fig.10, where obvious features can be observed. Fig. 5 gives the features in

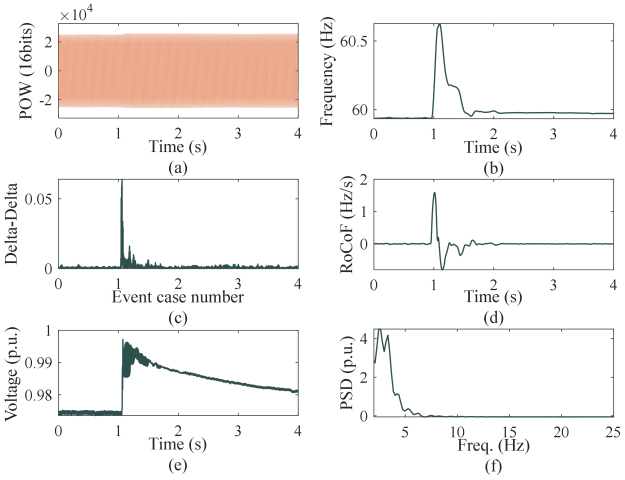

Fig. 7. Example feature responses for load shedding. (a) POW, (b)-(f) frequency, delta-delta, RoCoF, RMS of voltage, PSD.

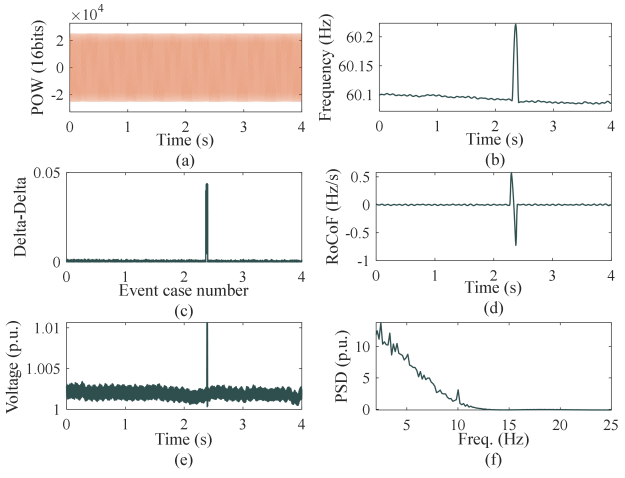

Fig. 8. Example feature responses for waveform distortions. (a) POW, (b)-(f) frequency, delta-delta, RoCoF, RMS of voltage, PSD.

a typical generation trip, where an obvious voltage magnitude drop during the anomaly can be observed from both POW and $V_{RMS}(t)$ responses. This magnitude drop is clearly captured by the POW delta-delta response as two spikes. From the frequency and RoCoF responses, the frequency dropping and recovering during the anomaly can be captured. Fig. 6 presents the behavior of an example fault. There is only one voltage magnitude drop with a short duration. A small frequency spike is captured by the frequency and RoCoF responses.

The load shedding in Fig. 7 clearly shows that both the voltage magnitude and frequency increase which behaves in the opposite direction as the generation trip. The waveform distortion is illustrated in Fig. 8(d), since there is a phase shift, obvious spikes can be observed in the POW delta-delta, frequency, and RoCoF responses which can be easily missed from the voltage magnitude response. As shown in Fig. 9, a switching anomaly usually has a short voltage magnitude drop, where no obvious frequency change can be captured. Besides, the oscillations caused by the droop coefficient design from the diesel generators are confirmed by KIUC. As illustrated in Fig. 10, obvious oscillations in 3.3 Hz, 6.6 Hz, 9.9 Hz, and 20 Hz can be observed from the power spectral density response. Notice that the highest oscillation that can be detected by the proposed algorithm is 720 Hz.

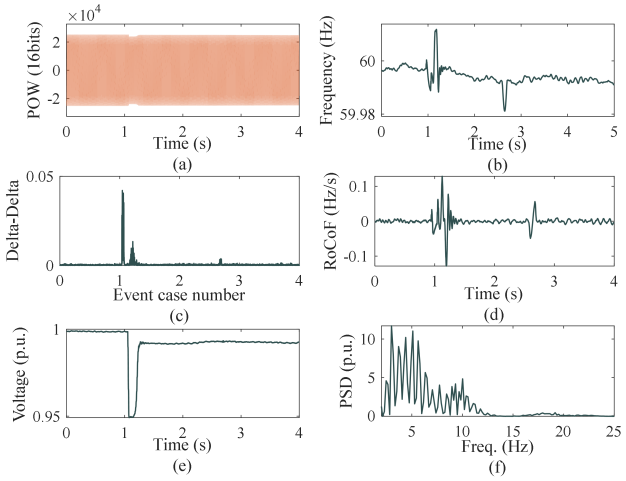

Fig. 9. Example feature responses for switching. (a) POW, (b)-(f) frequency, delta-delta, RoCoF, RMS of voltage, PSD.

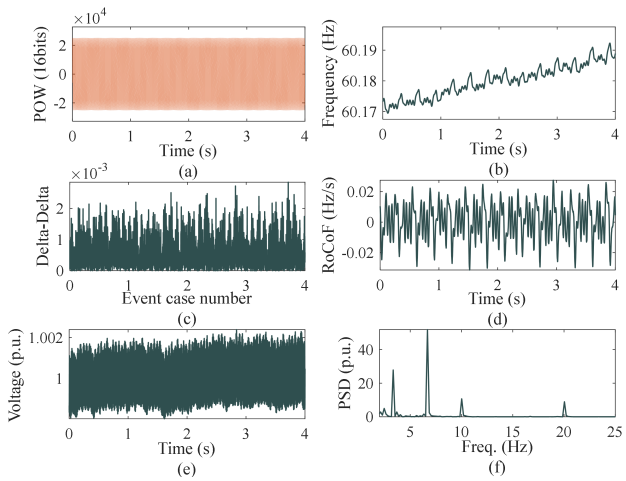

Fig. 10. Example feature responses for oscillations. (a) POW, (b)-(f) frequency, delta-delta, RoCoF, RMS of voltage, PSD.

C. POW-based Anomaly Library

To achieve accurate identification of anomalies, a POW-based anomaly library is established, where the collected data set information is listed in Table II. The anomalies collected within one-year or three months are utilized to perform the training data set. It should be noted that all the cases have been confirmed by the power utility company KIUC on Kauai island.

Moreover, the anomalies have multiple labels from both time and frequency domain aspects. For example, a waveform distortion anomaly can also be an oscillation anomaly (denoted as WO). In contrast, a waveform distortion anomaly without oscillation is denoted as WNO. Other anomalies are switching with/without oscillations (SO/SNO), fault with/without oscillations (FO/FNO), generation trips with/without oscillations (GO/GNO), load shedding with/without oscillations (LO/LNO), and others with/without oscillations (OO/ONO). In the anomaly categorization process, these twelve categories will be utilized.

To capture the feature of the abovementioned measurements, 24 features are available which can be classified into frequency-related, voltage-related, POW-related, and frequency spectrum-related, as listed in Table III. For each fea-

TABLE II
ANOMALY LABELS FOR CATEGORIZATION

Labels for anomalies	Number of anomalies			
	Oscillation		Non-oscillation	
Switching	SO	13	SNO	2295
Waveform distortions	WO	316	WNO	385
Others	OO	23	ONO	533
Generation trips	GO	17	GNO	47
Faults	FO	9	FNO	119
Load shedding	LO	3	LNO	7

TABLE III
DESIGNED ANOMALY FEATURE SETS

Feature sets	Features
All features	$f(t)$, $RoCoF(t)$, $V_{RMS}(t)$, $U_{diff}(t)$, $U_{dod}(t)$, and $P(f)$
Frequency related	$f(t)$ and $RoCoF(t)$
Voltage related	$V_{RMS}(t)$
POW related	$U_{diff}(t)$ and $U_{dod}(t)$
Frequency spectrum related	$P(f)$

ture, the mean, variance, skewness, and kurtosis are calculated.

D. Random forest Based Anomaly categorization

Considering the online deployment requirement of anomaly categorization, a practical and field-deployment-based algorithm is chosen in this paper. Currently, the frequency and phase-angle-based anomaly categorization algorithm deployed on FNET/GridEye is a decision tree [27], which is one of the typical threshold-based algorithms. However, it still suffers from features with strong correlations.

To overcome this problem, a random-forest-based algorithm is introduced in this paper. Compared with the decision tree, the proposed random-forest-based algorithm utilizes the decision tree as a basic categorization tool and thus upgrades it by using integrated learning algorithms.

The random forest-based algorithm implemented in this research consists of multiple decision trees and voting technology, of which the structure is illustrated in Fig. 11. Here, all designed anomaly feature sets are listed in Table III. All features are also classified into different combinations.

In the next step, multiple data subsets are generated as S_1 to S_N, where N is the number of trees. Given the data subsets, decision trees T_1 to T_N, are generated. The validation data sets are used to calculate the validation accuracies, A_1 to A_N which are used to vote for the random forest-based anomaly categorization results. In this regard, a robust forest is built up by using voting-based methods.

The overall framework of the random-forest-based algorithm is illustrated in Fig. 3. First, the features and anomaly labels of the anomalies are generated as inputs to the random-forest-based algorithm. Moreover, the parameters in the random-forest-based algorithm, including the minimum leaf number, the minimum predictor number, and the tree number are optimized by using Bayesian Optimization. Importantly, since the algorithm is deployed on the FNET/GridEye system, the anomaly library can be enriched gradually from the captured anomalies.

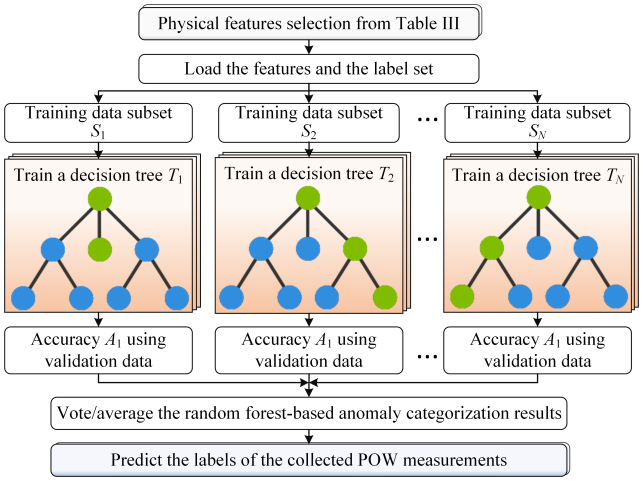

Fig. 11. The flow chart of the random forest-based algorithm.

E. Visualizations of the anomalies and features

To verify the categorization ability of the random-forest, the exacted features and the learned features are visualized using the t-Distributed Stochastic Neighbor Embedding (t-SNE). The t-SNE is a dimensional reduction method, and it can be used to reduce the dimensions of the input data into two dimensions.

The results are illustrated in Fig. 12. The results from Fig. 12(a) demonstrate that the features from different categories overlap with each other, indicating that it is challenging to identify the anomalies. In contrast, the features of Fig. 12(b) show that almost all the features have been differentiated from each other. This means that different types of anomalies can be effectively identified by using the proposed method.

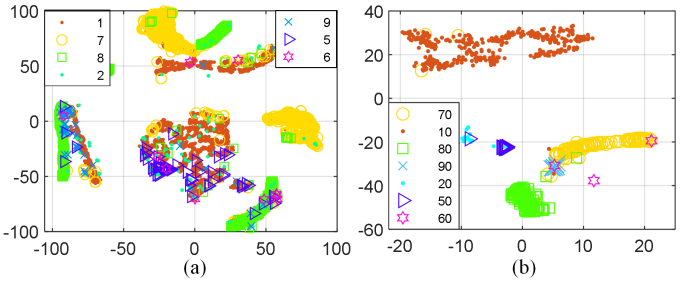

Fig. 12. Visualizations of the anomalies and features of after training based on t-SNE. Each number denotes one class and number i corresponding to $i * 10$. (a) features before training, (b) features after learning.

V. EXPERIMENT RESULTS

A. Parameter fine tuning

To further verify the performance of the proposed method, different experiments are carried out. To avoid over-fitting, 70%, 15%, and 15% of all the data sets are randomly selected for training, validation, and testing, respectively. Notice that the data set distributions are the same for all anomaly categories. During the optimization, the ten-fold cross-validation loss is set as the objective function while the above-mentioned parameters of the random forest algorithm are set as variables. 30 iterations in total are run for each round.

The validity of the selected feature sets is critical for the performance of the proposed methods. To fully demonstrate the performance of both the parameter optimization and feature

TABLE IV
PERFORMANCE COMPARISON UNDER DIFFERENT FEATURE SETS AND OPTIMIZATION

Selected feature sets	Performance	
	Before optimization	After optimization
Frequency related	91.71%	91.92%
Voltage related	87.39%	87.46%
POW related	91.54%	91.64%
Frequency spectrum related	83.97%	84.21%
All features	94.23%	94.54%

selection, the average accuracies with multiple combinations between feature selection and optimization are given in Table IV. It shows that the accuracy is higher while using the frequency-related features. Besides, the $U_{diff}(t)$ and $U_{dod}(t)$ related cases are the second highest compared with other feature selection cases. Additionally, it can be clearly observed that by using parameter optimization, the overall performance has been improved to 94.54% when all features are selected. The primary reason is that more features contain sufficient information.

To give an in-depth analysis of the results, the confusion matrix is also presented in Fig. 13. Based on both the validation and test data set with 1166 test cases, it can be concluded that the most serious confusion is between the waveform distortion with and without oscillations, which is caused by the lack of enough power spectral density information. In addition, there is some confusion about the fault without oscillation, which is because of the complicated characteristics of the fault and the similarity between fault and other anomalies such as frequency spikes and $V_{RMS}(t)$ drops.

B. Comparison to other algorithms

To verify the performance of the proposed anomaly categorization algorithm, a comparison among Decision tree-based, random forest-based, support vector machine (SVM)-based [28], kNN-based [16], Deep Neural Network (DNN)-based [14], and artificial Neural Network (ANN)-based algorithms is given in Table V. The feature sets are normalized between zero and one before conducting the experiments. The Bayesian optimization is applied to the SVM and kNN-based algorithms to improve the prediction accuracy. Meanwhile, the network structures of the Neural Network based algorithms are also optimized by using particle swarm optimization.

Among the benchmark algorithms, SVM and kNN-based algorithms are used as the basics as a result of their simple structures. The ANN and DNN are representing the Neural Network based machine learning algorithms. Meanwhile, the decision tree-based and random forest-based algorithms are tree-based machine learning algorithms. Basically, tree-based algorithms have simple structures and can be easily deployed in real-time. In contrast, neural-network-based algorithms have higher accuracies while they are usually deployed offline. For each algorithm, the accuracies are calculated under different feature sets and are discussed according to Table III.

It demonstrated that the performance of AVM and kNN-based methods based on the frequency and POW-related feature sets reach 91.88% and 92.57%, respectively. Meanwhile,

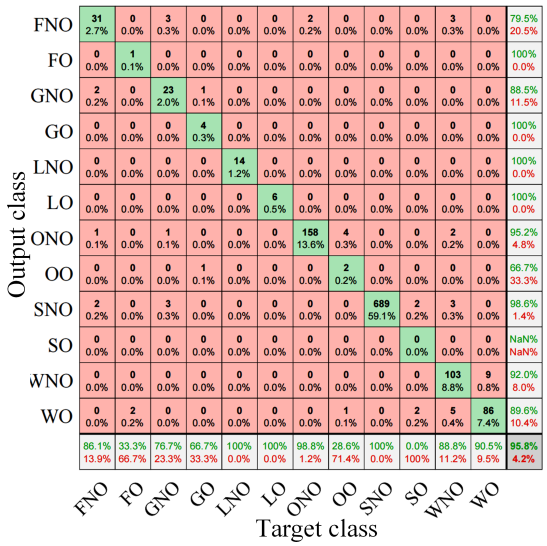

Fig. 13. The confusion matrix of the random forest-based algorithm.

the performance comparison under the frequency-related and POW-related features shows similar accuracy, indicating that the POW-related features are as important as the frequency-related features.

More combinations of the feature sets are tested to have a comprehensive comparison in the last four rows in Table III. Again, the overall performance of all algorithms basically follows the results from the previous discussion. It is worth mentioning that the overall accuracies can be improved by using POW-related features which verify that POW-related features, i.e., $U_{diff}(t)$ and $U_{dod}(t)$, can provide a complement to the frequency related features. The performance with frequency and POW-related features is 2.79% higher than that using frequency and voltage-related features. Besides, the performance under spectrum-related features is between 60.45% and 84.21%, which is the lowest one because the spectral characteristics can be greatly affected by noise.

Additionally, the random-forest-based algorithm has an accuracy as good as that of the DNN-based algorithm which is better than other basic and decision tree-based algorithms. It can be observed from Table V that with all features given, the accuracies from almost all the algorithms are the highest ones. The DNN, which is one of the most complicated algorithms, requires GPU to speed up the calculation, whose accuracy is only 0.48% higher than the random forest. Besides, from the algorithm complexity point of view, the tree-based characteristics of the random forest algorithm make it possible to be deployed in the general hardware system. Therefore, the random-forest-based algorithm is utilized as the POW-based anomaly categorization algorithm deployed in the server.

VI. CONCLUSIONS

To enhance the situational awareness capability in low inertia power systems, this paper presents POW-based anomaly detection and categorization algorithms. To detect the anomaly, a real-time POW-based anomaly detection algorithm is developed based on the delta-delta of the POW measurements on two adjacent cycles. Meanwhile, one year of POW measurements from four SMUs deployed on Hawaii islands are

TABLE V
PERFORMANCE COMPARISON AMONG DIFFERENT ANOMALY CATEGORIZATION ALGORITHMS

Selected feature sets	Performance of different algorithms					
	SVM-based [28]	kNN-based [16]	ANN-based [29]	DNN-based [14]	Decision tree-based	Random forest-based
Frequency related	88.60%	90.26%	89.88%	88.83%	89.22%	91.92%
Voltage related	78.24%	81.87%	84.99%	87.84%	84.80%	87.46%
POW related	84.77%	90.50%	88.60%	91.15%	89.43%	91.64%
Frequency spectrum related	60.45%	73.09%	67.36%	75.01%	83.63%	84.21%
Frequency and voltage related	84.97%	91.71%	90.05%	93.94%	91.09%	93.78%
Frequency and POW related	**91.88%**	**92.57%**	92.84%	93.42%	91.05%	93.64%
POW and voltage related	86.36%	82.21%	88.31%	93.21%	90.57%	93.09%
All features	87.56%	91.81%	**93.37%**	**95.02%**	**92.54%**	**94.54%**

collected by using FNET/GridEye system for a case study. The feature visualization analysis demonstrates the differences in several typical physical features. Thereafter, twelve categories of anomalies are labeled based on twenty-four features gathered from both time and frequency domains. A random-forest-based algorithm, as well as Bayesian optimization, are also developed to identify the anomalies. The performance comparison under different types of sub-feature sets reveals that all twenty-four features contribute to anomaly categorization. The comparison among some advanced data-driven-based algorithms indicates that the proposed random-forest-based algorithm has an accuracy as high as 94.54% and can achieve real-time computing.

REFERENCES

[1] E. Hossain, , and et al., "Analysis and mitigation of power quality issues in distributed generation systems using custom power devices," *Ieee Access*, vol. 6, pp. 16 816–16 833, 2018.

[2] A. Shahsavari, M. Farajollahi, , and et al., "Situational awareness in distribution grid using micro-pmu data: A machine learning approach," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6167–6177, 2019.

[3] H. Yin, Y. Wu, and et al., "Precise rocof estimation algorithm for low inertia power grids," *Electric Power Systems Research*, vol. 209, p. 107968, 2022.

[4] E. Heylen, F. Teng, and G. Strbac, "Challenges and opportunities of inertia estimation and forecasting in low-inertia power systems," *Renewable and Sustainable Energy Reviews*, vol. 147, p. 111176, 2021.

[5] A. F. Bastos, S. Santoso, W. Freitas, and W. Xu, "Synchrowaveform measurement units and applications," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2019, pp. 1–5.

[6] Y. Liu, L. Zhan, and et al., "Wide-area-measurement system development at the distribution level: An fnet/grideye example," *IEEE Transactions on Power Delivery*, vol. 31, no. 2, pp. 721–731, 2016.

[7] M. Lu and Z. Li, "A hybrid event detection approach for non-intrusive load monitoring," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 528–540, 2019.

[8] R. Ma, S. Basumallik, and S. Eftekharnejad, "A pmu-based data-driven approach for classifying power system events considering cyberattacks," *IEEE Systems Journal*, vol. 14, no. 3, pp. 3558–3569, 2020.

[9] A. J. Wilson and et al., "Automated identification of electrical disturbance waveforms within an operational smart power grid," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4380–4389, 2020.

[10] W. Li, M. Wang, and J. H. Chow, "Real-time event identification through low-dimensional subspace characterization of high-dimensional synchrophasor data," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4937–4947, 2018.

[11] N. Minh Khoa and L. Van Dai, "Detection and classification of power quality disturbances in power system using modified-combination between the stockwell transform and decision tree methods," *Energies*, vol. 13, no. 14, p. 3623, 2020.

[12] A. G. Shaik and O. P. Mahela, "Power quality assessment and event detection in hybrid power system," *Electric power systems research*, vol. 161, pp. 26–44, 2018.

[13] S. Brahma, , and et al., "Real-time identification of dynamic events in power systems using pmu data, and potential applications—models, promises, and challenges," *IEEE Transactions on Power Delivery*, vol. 32, no. 1, pp. 294–301, 2017.

[14] R. Yadav, S. Raj, and A. K. Pradhan, "Real-time event classification in power system with renewables using kernel density estimation and deep neural network," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6849–6859, 2019.

[15] S. Wang, P. Dehghanian, and L. Li, "Power grid online surveillance through pmu-embedded convolutional neural networks," *IEEE Transactions on Industry Applications*, vol. 56, no. 2, pp. 1146–1155, 2019.

[16] M. Biswal, S. M. Brahma, and H. Cao, "Supervisory protection and automated event diagnosis using pmu data," *IEEE Transactions on Power Delivery*, vol. 31, no. 4, pp. 1855–1863, 2016.

[17] M. Al Karim, J. Currie, and T.-T. Lie, "Dynamic event detection using a distributed feature selection based machine learning approach in a self-healing microgrid," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4706–4718, 2018.

[18] A. Zakariazadeh, "Smart meter data classification using optimized random forest algorithm," *ISA transactions*, vol. 126, pp. 361–369, 2022.

[19] L. Zhan, , and et al., "Fault-tolerant grid frequency measurement algorithm during transients," *IET Energy Systems Integration*, vol. 2, no. 3, pp. 173–178, 2020.

[20] L. Zhu and et al., "Fnet/grideye: A tool for situational awareness of large power interconnetion grids," in *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*. IEEE, 2020, pp. 379–383.

[21] H. Yin, W. Yao, and et al., "Low cost, flexible, and distribution level universal grid analyser platform: designs and implementations," *IET Generation, Transmission & Distribution*, vol. 14, no. 19, pp. 3945–3952, 2020.

[22] A. Hoke, V. Gevorgian, , and et al., "Island power systems with high levels of inverter-based resources: Stability and reliability challenges," *IEEE Electrification Magazine*, vol. 9, no. 1, pp. 74–91, 2021.

[23] Hawaiian Electric, "Hawaiian electric hits 38% renewable energy in 2021," *[Online], available at: https://www.hawaiianelectric.com/hawaiianelectrichits38renewable-energyin2021*, 2022.

[24] "Ieee standard for synchrophasor data transfer for power systems," *IEEE Std C37.118.2-2011 (Revision of IEEE Std C37.118-2005)*, pp. 1–53, 2011.

[25] X. Deng and et al., "Frequency observations and statistic analysis of worldwide main power grids using fnet/grideye," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2019, pp. 1–5.

[26] FNET/GridEye, "Fnet/grideye frequency display," *[Online], available at: https://fnetpublic.utk.edu/tabledisplay.html*, 2022.

[27] S. Liu, S. You, and et al., "Data-driven event identification in the u.s. power systems based on 2d-olpp and rusboosted trees," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 94–105, 2022.

[28] H. Erişti and Y. Demir, "Automatic classification of power quality events and disturbances using wavelet transform and support vector machines," *IET generation, transmission & distribution*, vol. 6, no. 10, pp. 968–976, 2012.

[29] W. Li, A. Monti, and F. Ponci, "Fault detection and classification in medium voltage dc shipboard power systems with wavelets and artificial neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 11, pp. 2651–2665, 2014.