IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON ARTIFICIAL INTELLIGENCE TECHNOLOGIES FOR ELECTRIC POWER SYSTEMS

Received July 12, 2019, accepted August 9, 2019, date of publication August 21, 2019, date of current version September 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2936816

Multivariate Gaussian-Based False Data Detection Against Cyber-Attacks

YU AN, (Student Member, IEEE), AND DONG LIU, (Senior Member, IEEE)
Key Laboratory of Control of Power Transmission and Conversion, Ministry of Education, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Corresponding author: Dong Liu (dongliu@sjtu.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFB0903000, and in part by the National Natural Science Foundation of China under Grant 51677116.

ABSTRACT Modern distribution power system has become a typical cyber-physical system (CPS), where reliable automation control process is heavily depending on the accurate measurement data. However, the cyber-attacks on CPS may manipulate the measurement data and mislead the control system to make incorrect operational decisions. Two types of cyber-attacks (e.g., transient cyber-attacks and steady cyber-attacks) as well as their attack templates are modeled in this paper. To effectively and accurately detect these false data injections, a multivariate Gaussian based anomaly detection method is proposed. The correlation features of comprehensive measurement data captured by micro-phasor measurement units (μPMU) are developed to train multivariate Gaussian models for the anomaly detection of transient and steady cyber-attacks, respectively. A k-means clustering method is introduced to reduce the number of μPMUs and select the placement of μPMUs. Numerical simulations on the IEEE 34 bus system show that the proposed method can effectively detect the false data injections on measurement sensors of distribution systems.

INDEX TERMS Cyber-physical system, cyber-attack, anomaly detection, distribution grid, machine learning.

I. INTRODUCTION

Information technology plays a critical role in the automatic control of modern distribution power systems, which consist of a large number of computation systems, local sensors, and communication networks. With the increasing and deep interaction between physical flow and cyber flow, the modern distribution system becomes a typical cyber-physical system (CPS) [1], [2]. The core of CPS in power grids is to achieve high-sensitivity awareness and real-time automation of physical processes through the integration and coordination of 3C (Computation, Communication, and Control) technology [3]–[5]. The integration of cyber networks and physical systems has significantly enhanced the efficiency and reliability of distribution system operations. However, the strong interdependence between cyber networks and physical power grids can bring more potential risks through complex communication links, which are significantly vulnerable to cyber security threats [6]. In addition, both the privatization of energy industries and the standardization of communication technologies facilitate opponents to tamper cyber networks in power system. The existence of cyber threats in CPS, such as cyber components failures, security risks, and cyber-attacks, can readily lead to the abnormal operation or even cascading outages of the entire power grid by propagating from a single point failure in cyber networks. For example, the "Ukrainian Blackout" [7] in 2015 was a typical cyber-physical cascading failure caused by malicious cyber-attacks, which aroused widespread concerns in the cyber security of power system operations.

The existing researches investigate the countermeasures against cyber-attacks from different approaches, such as vulnerability assessment [8], anomaly detection [9]–[11], and attack mitigation [12], [13]. From the perspective of anomaly detection, the identification methods of corrupted measurement data have widely attracted interests of worldwide scholars. Game theories are introduced to quantify the security risk in cyber networks [9], [14]. Petri net is leveraged to describe the information flows among the components in CPS of power systems [10], [15]. Also, some statistical methods are investigated in the detection of false measurement data. For example, Cui *et al.* [11] developed a machine learning

The associate editor coordinating the review of this article and approving it for publication was Mingjian Cui.

based anomaly detection method for tampered load forecasting data. Wang *et al.* [16] established a distributed framework based on the deep auto-encoder to detect the manipulated data. Esmalifalak *et al.* [17] utilized both supervised and unsupervised learning methods for the detection of stealthy attack. Among these statistical methods, comprehensive variables of essential measurement data are the key feature for successfully identifying CPS anomalies. However, compared with the transmission-level grid, the real-time condition in distribution grids is not sufficiently monitored by distribution system operators (DSOs). To acquire situational awareness on the CPS of distribution grids, there is a growing need in the enhancement of sensors.

Recently, micro-phasor measurement units (μPMUs) have been widely developed to capture operational states in the distribution-level systems. Compared with the traditional supervisory control and data acquisition (SCADA) system with minute-level power flow sampling rates, μPMUs can capture the voltage and current phasors with significantly higher sampling rates [18], e.g., 30 Hz in [19], 50 Hz in [20], 60 Hz in [21], 100 Hz in [22], and 120 Hz in [23]. The applications of μPMUs and PMUs have been discussed in plenty of literatures [18], [23]–[25]. Cui *et al.* [18] proposed a novel event detection methodology using huge amount of PMU data. Jamei *et al.* [23] established an abnormal behavior detection framework based on optimal placement of μPMUs in distribution grid. Gomez *et al.* [24] trained a support vector machines classifier to predict post-fault transient stability status based on the transient data acquired from PMU. Li and Yang applied an ensemble of OS-extreme learning machine with binary Java based feature selection to predicting the transient stability status of power system by using PMU data [25]. However, there are few applications for the anomaly detection of CPS in distribution systems by using μPMUs data as the analytical and statistical basis.

In this paper, we mainly focus on two types of false data injection attacks (i.e., transient attacks and steady attacks) in CPS of distribution grids. A multivariate Gaussian based anomaly detection method is proposed to identify abnormal CPS events by using variables of measurements acquired from μPMUs. Firstly, the templates of two types of false data injection cyber-attacks are introduced and modeled in Section II. In Section III, a multivariate Gaussian based machine learning method is developed for the CPS anomaly detection considering both transient and steady cyber-attacks. Section IV provides a method for the selection of μPMU placement to reduce the number of μPMU devices in distribution systems. Case studies as well as numerical results are discussed in Section V. Conclusions are summarized in Section VI.

II. CYBER-ATTACK TEMPLATE

According to the IEEE PES Distribution automation working group [26], the modern distribution system is automatically controlled to enable DSOs to remotely monitor, coordinate, and operate local distribution devices and components in a real-time mode. Thus, the opponents may mislead the DSOs into inaccurate operational decisions by stealthily injecting false data into local measurement sensors. In this paper, two types of false data injection attacks (i.e. transient attacks and steady attacks) are investigated. Inspired by cyber-attack templates mentioned in [11], the templates of transient attack and steady attack are considered as step attack and ramping attack, respectively.

A. STEP ATTACK

The representative template of transient cyber-attack is a step attack, which modifies the measurement values in a specific duration multiplied by a parameter p_s:

$$\dot{M}_t = (1 + p_s) \times M_t, \quad \text{for } t_s \le t \le t_e \tag{1}$$

where $\dot{M}_t$ is the tampered measurement value at time t. M_t is the original measurement value at time$t$. t_s and t_e are the start and end time of one transient cyber-attack, respectively.

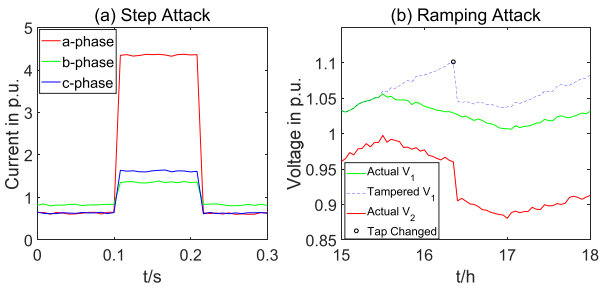

FIGURE 1. Examples of (a) transient cyber-attack and (b) steady cyber-attack.

The aim of transient cyber-attacks is to simulate a ''system fault'' by injecting step overcurrent or low-voltage data into corresponding measurement sensors, as shown in Fig. 1a. The duration of transient cyber-attack is so short (e.g., 0.1s∼0.2s) that the SCADA cannot detect this anomaly event. However, if the step parameter is selected sophisticatedly and implemented on the critical sensors of protection system, the fault detector would be deceived with the tampered transient measurements to alarm a fault. Under this situation, the distribution automation system would be misguided to isolate ''fault area'' by tripping the protective relays, which would cause unnecessary outages in distribution systems.

B. RAMPING ATTACK

The steady cyber-attack can be represented as a slowly ramping attack, which modifies the measurement values to rise (or decline) gradually with time:

$$\dot{M}_t = \begin{cases} [1 + (t - t_s)p_r] \times M_t, & \text{for } t_s \le t \le t_e \\ [1 + (t_e - t_s)p_r] \times M_t, & \text{for } t > t_e \end{cases} \tag{2}$$

where p_r is the ramping parameter. $p_r > 0$ denotes an up-ramping cyber attack, while $p_r < 0$ represents a down-ramping attack.

Unlike transient cyber-attacks, steady cyber-attacks may not show distinct transient features. Instead, they inject consistent and slowly ramping false data into measurement sensors, which are more challenging to be detected by visualization. Due to the false data tampered by cyber attackers, the control center in distribution systems may send incorrect operational decisions from DSOs. For instance, as shown in Fig. 1b, if the value of voltage V_1 is manipulated by a ramping attack, the control center would be deceived under the ''over voltage'' condition. The transformer tap would be changed to lower the system's voltage, which may further decrease the voltage V_2 to the low-voltage condition.

III. ANOMALY DETECTION METHOD

In this section, we assume that it is unrealistic for opponents to implement successful cyber intrusion in each variable of measurements simultaneously. The most economical way is to manipulate some critical measurement data that can influence the estimation of automatically controlled system. Under this circumstance, the correlations between variables of comprehensive measurements captured by μPMUs can be investigated to detect the anomalies. In this section, a machine learning based anomaly detection method using the multivariate Gaussian model is proposed to detect transient and steady cyber-attacks according to different correlation features. Two design methods of features for transient and steady cyber-attack detection are illustrated as follows.

Algorithm 1 Greedy Search Pseudo-Code for Threshold ε Selection

$\hat{x}_{CV}$ = Feature vectors in cross-validation set;
$\hat{y}_{CV}$ = Anomaly flags matched with $\hat{x}_{CV}$;
$bestF_1 = 0$; // Initialization of F_1 score
$\varepsilon = 0$; // Initialization of threshold
Given $\hat{x}_{CV}$, calculate $\hat{p}_{CV} = p(\hat{x}_{CV})$;
$stepsize = (\max(\hat{p}_{CV}) - \min(\hat{p}_{CV}))/100000$;
for $eps = \min(\hat{p}_{CV}) : stepsize : \max(\hat{p}_{CV})$ **do**
  $\hat{y}_{PR} = \hat{p}_{CV} < eps$;
  given $\hat{y}_{PR}$ and $\hat{y}_{CV}$, calculate F_1;
  if $F_1 > bestF_1$ **then**
    $bestF_1 = F_1$;
    $\varepsilon = eps$;
  end if
end for

A. MULTIVARIATE GAUSSIAN MODEL

In probability theory and statistics, the multivariate Gaussian distribution is a generalization of the one-dimensional Gaussian distribution to higher dimensions. Compared with one-dimensional Gaussian model, multivariate Gaussian model can capture the correlations between variables from different dimensions by formulating and calculating a covariance matrix. In the field of applications, the multivariate Gaussian model has been widely used in the abnormal signal detection [27], [28]. The basic idea is to train a multivariate Gaussian model of the samples with multiple features, and judge whether a new sample is homogenous or abnormal by measuring its probability. As shown in Fig. 2, an example of two-dimensional Gaussian distribution is presented. The anomaly event is flagged when the joint probability is below a threshold. It should be noted that the algorithm defaults to Gaussian distribution for all features. If the samples do not present a Gaussian distribution on a certain feature, they can be converted to a Gaussian distribution, such as log transformation, square root transformation, reciprocal transformation, etc. If the distribution of a feature is too complicated to be converted to a Gaussian distribution, especially when the distribution of the feature has multiple extremums, it can not be trained in the multivariate Gaussian model.

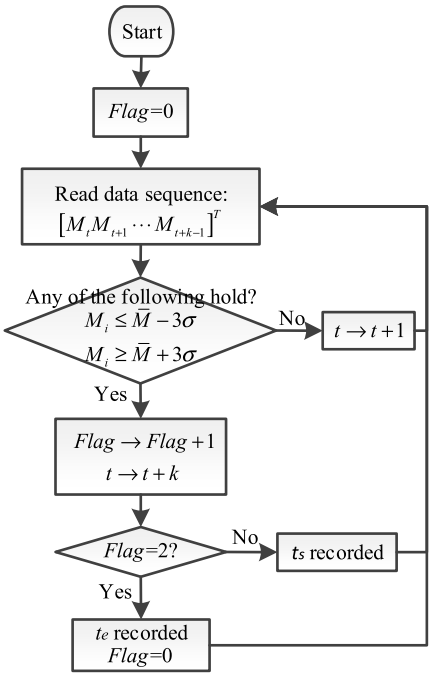

FIGURE 2. Flowchart of transient event detection.

Given a training set $\{\hat{x}^{(1)}, \hat{x}^{(2)}, \cdots, \hat{x}^{(N_T)}\}$, a multivariate Gaussian model $p(\hat{x})$ is fitted by:

$$p(\hat{x}) = \frac{1}{(2\pi)^{\frac{N_T}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\hat{x} - \hat{x}_m)^T \Sigma^{-1} (\hat{x} - \hat{x}_m)\right) \tag{3}$$

$$\hat{x}_m = \frac{1}{N_T} \sum_{i=1}^{N_T} \hat{x}^{(i)} \tag{4}$$

$$\Sigma = \frac{1}{N_T} \sum_{i=1}^{N_T} (\hat{x}^{(i)} - \hat{x}_m)(\hat{x}^{(i)} - \hat{x}_m)^T \tag{5}$$

where N_T is the number of training set. $\hat{x}$ is the feature vector. $\hat{x}_m$ is the mean vector of the training set. Σ is the covariance matrix. Given a new example $\hat{x}$, we can compute $p(\hat{x})$ by using (3). An anomaly event would be flagged if $p(\hat{x}) < \varepsilon$, where ε is the threshold used for anomaly detection.

A greedy algorithm is applied to select the threshold ε. First, we calculate $p(\hat{x}_{CV})$ through a cross-validation set $\{(\hat{x}_{CV}^{(1)}, y_{CV}^{(1)}), (\hat{x}_{CV}^{(2)}, y_{CV}^{(2)}), \cdots, (\hat{x}_{CV}^{(N_{CV})}, y_{CV}^{(N_{CV})})\}$, where N_{CV} is the number of cross-validation set. $\hat{x}_{CV}$ is the feature vector in the cross-validation set. y_{CV} is the anomaly flag that $y_{CV} = 1$ for abnormal features and $y_{CV} = 0$ for normal features. For each $\varepsilon \in [\min(p(\hat{x}_{CV})), \max(p(\hat{x}_{CV}))]$, we can evaluate the performance of ε through a matching set $\{(p(\hat{x}_{CV}^{(1)}), y_{CV}^{(1)}), (p(\hat{x}_{CV}^{(2)}), y_{CV}^{(2)}), \cdots, (p(\hat{x}_{CV}^{(\gamma)}), y_{CV}^{(\gamma)})\}$ by calculating F_1 score:

$$F_1 = \frac{2 \cdot f_p \cdot f_r}{f_p + f_r} \tag{6}$$

$$f_p = \frac{n_{tp}}{n_{tp} + n_{fp}} \tag{7}$$

$$f_r = \frac{n_{tp}}{n_{tp} + n_{fn}} \tag{8}$$

where f_p is the precision metric. f_r is the recall metric. n_{tp} is the number of true positives that indicate $p(\hat{x}_{CV}) = 1$ and $y_{CV} = 1$. n_{fp} is the number of false positives that indicate $p(\hat{x}_{CV}) = 1$ and $y_{CV} = 0$. n_{fn} is the number of false negatives that indicate $p(\hat{x}_{CV}) = 0$ and $y_{CV} = 1$. For different $\varepsilon \in [\min(p(\hat{x}_{CV})), \max(p(\hat{x}_{CV}))]$, we can get different F_1 scores. Theε with a maximum F_1 score is finally used as the threshold for the anomaly detection.

B. TRANSIENT CYBER-ATTACK DETECTION

Based on the high sampling rate of μPMU, the transient event on measurements can be readily captured and detected. In this section, an anomaly detection method is proposed for transient cyber-attacks, which aims to simulate system faults by injecting transient measurements. In this study, the monitored measurements are the magnitudes of three-phase voltage, three-phase current, three-phase active power, and three-phase reactive power. There are 12 types of measurement data monitored for cyber-attack detection.

Let $[M_1 M_2 \cdots M_{N_S}]^T$ denotes a sequence of captured measurements, where N_S is the number of captured samples in a data sequence. A transient event is alarmed if there exists a data point $i \in \{1, 2, \cdots, N_S\}$ for which any of the following holds:

$$\begin{aligned} M_i &\le \bar{M} - 3\sigma \\ M_i &\ge \bar{M} + 3\sigma \end{aligned} \tag{9}$$

where $\bar{M}$ is the mean value of data sequence. σ is the sample standard deviation of the data sequence.

The flowchart of a transient event detection is presented in Fig. 2. At each instant of time, the readings (a data sequence with a predefined window) are calculated according to (9). Once a transient event is found, the start time t_s and the end time t_e are recorded, and the data sequence from t_s to t_e is leveraged to develop the transient feature. The fluctuation ΔM in this data sequence is determined by:

$$\Delta M = \max\left\{\left|M_i - M_j\right|\right\}, \quad \text{for } i, j \in [t_s, t_e] \tag{10}$$

where $|.|$ denotes the absolute value. For a three-phase measurement, the maximum fluctuation in the three-phase measurement is regarded as a transient feature.

The transient feature vector $\hat{x}_T$ in this paper is a four-dimensional vector, which includes transient voltage feature, transient current feature, transient active power feature, and transient reactive power feature. To train a multivariate Gaussian model, all of the system faults (including single phase to ground fault, 2-phase to ground fault, phase-phase fault, and 3-phase fault) at different locations as well as some transient cyber-attack templates are simulated to acquire transient features $\hat{x}_T$. A suspicious transient cyber-attack is alarmed if $p(\hat{x}_T) < \varepsilon$.

C. STEADY CYBER-ATTACK DETECTION

In the steady operation, the voltage values of buses may be tempered and modified by the opponent, which aims to mislead operational decisions of the control center. An anomaly detection method for steady cyber-attacks on voltage is proposed in this section to alarm when the suspicious voltage value is detected.

Under n different network conditions, the voltage of bus$i$ can be represented as a n-dimensional vector $\hat{v}_i = [V_{i1} V_{i2} \cdots V_{in}]^T$, where V_{ij} is the voltage value of bus i under network condition $j(j = 1, 2, \cdots n)$. Assume two of these buses are matched into a pair set $(\hat{v}_A, \hat{v}_B)$, where $\hat{v}_A = [V_{A1} V_{A2} \cdots V_{An}]^T$ and $\hat{v}_B = [V_{B1} V_{B2} \cdots V_{Bn}]^T$ denote n-dimensional voltage vectors for bus A and bus B, respectively. A linear correlation between bus A and bus B can be built by using $\{(V_{A1}, V_{B1})(V_{A2}, V_{B2}) \cdots (V_{An}, V_{Bn})\}$ as training set. Therefore, the voltage value of bus B can be predicted by the monitored voltage value of bus A with the linear regression model:

$$\dot{V}_{\mathrm{B}} = V_{\mathrm{A}} a + b \tag{11}$$

where $\dot{V}_B$ is the predicted voltage value of the bus B. V_A is the monitored voltage value of bus A. a and b are the linear regression parameters for the pair set. Assume there are m linear regression models, a m-dimensional steady feature $\hat{x}_S$ can be established based on the prediction errors, given by:

$$\hat{x}_S = \hat{v}_{PB} - \hat{v}_B \tag{12}$$

where $\hat{v}_{PB} = [\dot{V}_B^{(1)} \dot{V}_B^{(2)} \cdots \dot{V}_B^{(m)}]^T$ is the vector of predicted voltage values of bus B in m pair sets. $\hat{v}_B = [V_B^{(1)} V_B^{(2)} \cdots V_B^{(m)}]^T$ is the vector of monitored voltage values of bus B in m pair sets. $\hat{x}_S$ is a m-dimensional feature vector that is used for steady cyber-attack detection. The number of pair sets is determined by the placement of μPMUs. In this paper, the adjacent buses equipped with μPMUs are matched into a pair set. The placement selection of μPMU is illustrated in Section IV.

IV. PLACEMENT SELECTION FOR μPMU

It is challenging to place μPMU at each bus in a large-scale distribution grid due to the cost. Thus, it is necessary to

select specific buses for the placement of limited μPMUs. For a radial distribution network, the voltage profiles at different nodes are generally determined by the node position, load fluctuations, and renewables fluctuations, etc. Therefore, the voltage profiles at the same branches or areas would be similar. In this section, a k-means clustering method is leveraged to divide the distribution network into multiple areas, where the voltage profiles are similar. The bus with centroid voltage in each cluster set is placed at least one μPMU to represent the characteristics of this area.

Given a distribution grid with L buses, voltages of L buses under n different load conditions are observed as $(\hat{v}_1\hat{v}_2\cdots\hat{v}_L)$, where each $\hat{v}_i = [V_{i1}V_{i2}\cdots V_{in}]^T$ is a n-dimensional voltage vector for bus i. The observation sets $(\hat{v}_1\hat{v}_2\cdots\hat{v}_L)$ can be divided into multiple clusters by using k-means clustering method, which can partition L bus voltage vectors into K voltage cluster sets $\boldsymbol{C} = \{C_1C_2\cdots C_K\}$ by minimizing the within-cluster sum of squares:

$$\min_{\mathbb{C}} \sum_{i=1}^{K} \sum_{\hat{v}\in C_i} \left\|\hat{v} - \hat{\mu}_i\right\|_2^2 \tag{13}$$

$$\hat{\mu}_i = \frac{1}{n_i} \sum_{\hat{v}\in C_i} \hat{v} \tag{14}$$

where $\hat{\mu}_i$ is the centroid voltage vector in the i-th voltage cluster set C_i (i= 1, 2, $\cdots K$). K is the number of voltage cluster sets. $\hat{v}$ is the n-dimensional voltage vector in voltage cluster set C_i. n_i is the number of voltage vectors in cluster set C_i. $\|\cdot\|_2^2$ denotes the squared Euclidean distance.

There is at least one μPMU for each voltage cluster set C_i. The bus with the most similar voltage profile as the centroid voltage vector is selected as the centroid bus for the placement of μPMU. This objective can be achieved by selecting the minimal squared Euclidean distance between the bus's voltage vector and the centroid voltage vector:

$$j = \arg\min_{\hat{v}_j\in C_i} \left\|\hat{v}_j - \hat{\mu}_i\right\|_2^2 \tag{15}$$

where $\hat{v}_j$ is the voltage vector of bus j in cluster set C_i. Then bus j will be selected for the placement of μPMU.

To estimate the number of cluster sets in the k-means clustering method, set a cost function of K:

$$W_K(K) = \sum_{i=1}^{K} \sum_{\hat{v}_j\in C_i} \left\|\hat{v}_j - \hat{\mu}_i\right\|_2^2 \tag{16}$$

where $W_K(K)$ represents the pooled within-cluster sum of squares around K centroids. K is the number of cluster sets. $\hat{\mu}_i$ is the centroid voltage vector in the i-th voltage cluster set C_i. $\hat{v}_j$ is the voltage vector of bus j in cluster set C_i. C_i is the i-th voltage cluster set that is divided by the k-means clustering method.

By plotting the curve of W_K with K as the abscissa, the elbow point should be the number of clusters. In this paper, 500 load conditions in an unbalanced distribution grid (IEEE 34 bus system) are considered and simulated through

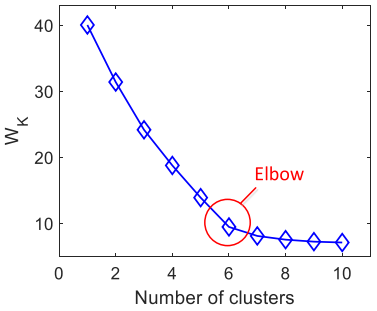

FIGURE 3. Find the elbow point of the pooled within-cluster sum of squares around centroids based on different numbers of the clusters in k-means clustering method.

OpenDSS [29], [30] for the bus's voltage clustering. Therefore, the voltages on the buses are 500-dimensional vectors. These voltage vectors can be divided into K clusters based on the k-means clustering method. With different number of the clusters, the function $W_K(K)$ in (16) is plotted in Fig. 3. As the increase of the number of the cluster sets, the value of the cost function decreases, and the elbow point is found when $K = 6$. Therefore, the IEEE 34 bus system is divided into 6 areas, as shown in Fig. 4. According to (15), the centroid buses are bus 808 for area 1, bus 816 for area 2, bus 820 for area 3, bus 854 for area 4, bus 890 for area 5, and bus 834 for area 6. It should be noted that if the budget allows, the proposed method would perform better when multiple μPMUs are placed in the same area. In this paper, each area is placed with only one μPMU at each centroid bus.

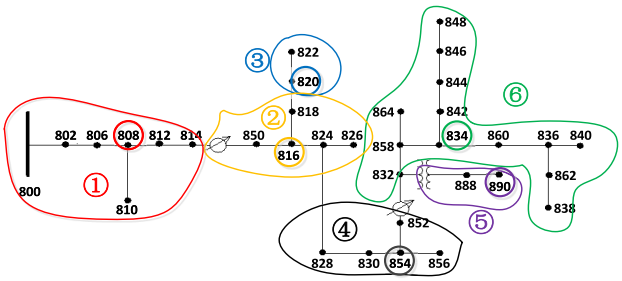

FIGURE 4. Topology of the IEEE 34 bus system.

V. NUMERRICAL TEST

The proposed method is tested on IEEE 34 bus system as the Fig. 4, where 6 μPMUs are placed on bus 808, bus 816, bus 820, bus 854, bus 890, and bus 834, respectively. The numerical simulation is conducted on a 64-bit computer with an Intel Core i5 CPU of 2.30GHz and 8GB RAM. The proposed algorithms are implemented in the MATLAB platform.

A. TRANSIENT ATTACK DETECTION

In this section, 10 types of faults, including single phase (a, b, and c) to ground fault, phase to phase (a-b, b-c, and c-a) fault, 2 phase (ab, bc, and ca) to ground fault, and 3 phase fault, are simulated at each node and each distribution lines of the IEEE 34 bus system. Therefore, a multivariate Gaussian model $p(\hat{x}_T)$ is derived from the transient features,

such as maximum current fluctuation, maximum current fluctuation, maximum active power fluctuation, and maximum reactive power fluctuation, based on 400 different system fault conditions. In cross-validation, 100 system fault conditions and 100 transient false data injection conditions are simulated to select the threshold. For each transient false data injection condition, the measurements of current, voltage, active power and reactive power under the normal condition are randomly selected to be injected with simulated false data under the fault condition. The threshold is determined as 0.99×10^{-3}.

A transient even is firstly captured by the μPMUs, which then input the transient features to the trained multivariate Gaussian model $p(\hat{x}_T)$ to identify whether it is a system fault or cyber-attack. For instance, a single phase (a-phase) to ground fault is simulated on line 824-828 at 0.1s and cleared at 0.2s. The transient event captured by μPMU 1 at bus 808 is shown in Fig. 5. With the captured start time t_s and the captured end time t_e, the measurement sequence between t_s and t_e is leveraged to establish the transient features $\hat{x}_T$. By comparing $p(\hat{x}_T)$ and the threshold, the anomaly would be flagged.

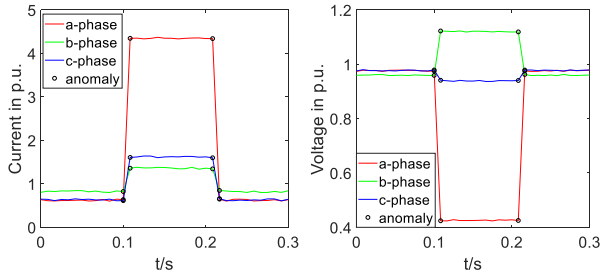

FIGURE 5. Anomalies captured by μPMU under a single phase (a-phase) to ground fault.

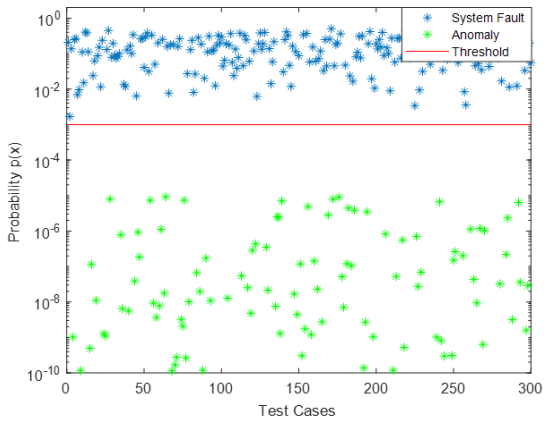

FIGURE 6. Test results of transient cyber-attack detection.

To validate the effectiveness of the trained multivariate Gaussian model $p(\hat{x}_T)$, 200 system fault conditions as well as 100 transient false data injection conditions are simulated. The test result is shown in Fig. 6. According to the calculation results of multivariate Gaussian model, there is a significant distinction in the value of probability $p(\hat{x}_T)$ between system fault conditions and cyber attacked conditions. In Fig. 6, the anomalies that represent cyber attacked conditions have extreme small values of probabilities (between 0.99×10^{-5} to 0.99×10^{-5}), while the most probabilities of system fault conditions are around 0.99×10^{-3} to 1.

B. STEADY ATTACK DETECTION

In this section, 500 cases of IEEE 34 bus system under different load conditions are considered and simulated based on the OpenDSS platform. The voltage values of 6 centroid buses under 500 network conditions are presented in Fig. 7. As shown in Fig. 7, there are many scenarios where the voltages of centroid buses exceed the upper boundary of 1.1 p.u. or drop below the lower boundary of 0.9 p.u. Thus, the opponent can inject false voltage data into measurement sensors of centroid buses to pretend a low-voltage condition or over-voltage condition, which may mislead the control center to change the transformer tap and influence the system's voltage under normal operations.

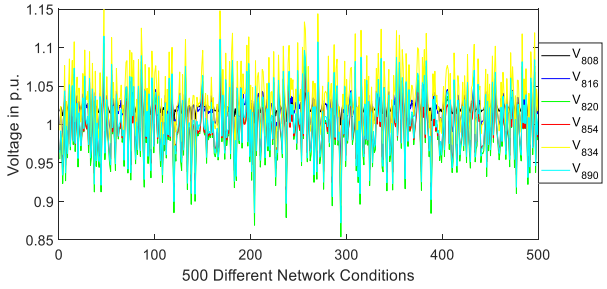

FIGURE 7. Voltage values of centroid buses under 500 network conditions.

To detect the abnormal voltage data that is tampered with false data, we developed 5 linear regression models to create the correlations between the voltages of adjacent centroid buses ($V_{808} - V_{816}$, $V_{816} - V_{820}$, $V_{816} - V_{854}$, $V_{854} - V_{834}$, and $V_{854} - V_{890}$) based on 500 load conditions. The linear correlations between voltages of centroid buses are shown in Fig. 8. The good fit results indicate that the voltage value of a centroid bus can be predicted by the adjacent centroid bus with a linear regression model, which can be leveraged to predict the voltages and detect the anomalies based on the prediction errors. The steady features are established from the prediction errors calculated from (17), (18)

$$\begin{aligned}
\dot{V}_{816} &= -2.0299 + 2.9790V_{808} \\
\dot{V}_{820} &= -0.8581 + 1.8124V_{816} \\
\dot{V}_{854} &= -0.3435 + 1.3097V_{816} \\
\dot{V}_{834} &= -0.3847 + 1.4246V_{854} \\
\dot{V}_{890} &= -0.4265 + 1.3547V_{854} \\
x_1 &= \dot{V}_{816} - V_{816} \\
x_2 &= \dot{V}_{820} - V_{820}
\end{aligned} \tag{17}$$

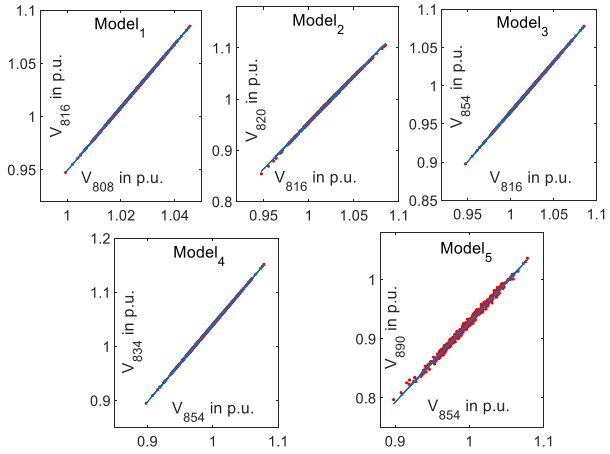

FIGURE 8. Results of 5 linear regression models.

$$\begin{aligned} x_3 &= \dot{V}_{854} - V_{854} \\ x_4 &= \dot{V}_{834} - V_{834} \\ x_5 &= \dot{V}_{890} - V_{890} \end{aligned} \tag{18}$$

where $\hat{x}_S = [x_1 x_2 x_3 x_4 x_5]^T$ is the prediction error vector as well as the feature vector for steady cyber-attack detection. V_{808}, V_{816}, V_{820}, V_{854}, V_{834}, and V_{890} are the monitored voltage values of bus 808, bus 816, bus 820, bus 854, bus 834, and bus 890, respectively. $\dot{V}_{816}$, $\dot{V}_{820}$, $\dot{V}_{854}$, $\dot{V}_{834}$, and $\dot{V}_{890}$ are the predicted voltage values of bus 816, bus 820, bus 854, and bus 890, respectively.

There are 500 feature vectors developed based on simulation results of the IEEE 34 bus system under different load conditions, and 400 of these feature vectors are leveraged to train a multivariate Gaussian model. The other 100 feature vectors of normal load conditions with another created 100 cyber-attack conditions are used for cross-validation. The cyber-attack conditions are created based on the attack template of ramping attacks, where the deviation range of tampered voltage data is randomly changed between 0.01 to 0.1. With 100 abnormal features and 100 normal features in cross-validation test, 1.99×10^{-8} is chosen as the threshold to identify the steady cyber-attacks.

A sensitivity test for the selected threshold to flag an anomaly is implemented in this section. The result is shown in Fig. 9. The voltage data of the centroid bus (V_{816}, V_{820}, V_{854}, V_{834}, and V_{890}) is manipulated and increased gradually from 0 to 0.02. With the increase of the deviation between tampered voltage data and predicted voltage value, the probabilities calculated by $p(\hat{x}_S)$ are reduced. As shown in Fig. 9, with the increasing deviation of tampered voltage data, the value of $p(\hat{x}_S)$ is decreasing severely. The most sensitive buses are V_{816} and V_{854}, because they are used most frequently in the linear regression models and influence more features in multivariate Gaussian model. The V_{890} is the least sensitive because the prediction of V_{890} based on V_{854} is the most inaccurate among 5 linear regression models, as shown in Fig. 8. This is because the bus 854 and bus

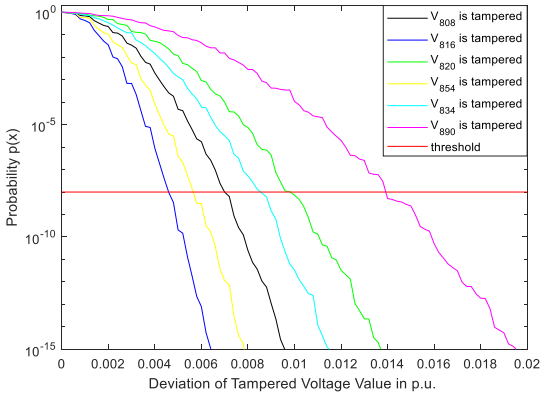

FIGURE 9. Results of sensitivity test for tampered voltage value with different degrees of deviation.

890 belong to the different voltage levels, which cause a weaker correlation compared with other pair sets. However, the detection performance is still good as the anomaly of V_{890} is flagged when the tampered deviation exceeds 0.014 p.u., within which the tampered voltage value cannot mislead the voltage control. Thus, the detection sensitivity of a specific bus can be improved when a more accurate linear regression model is developed.

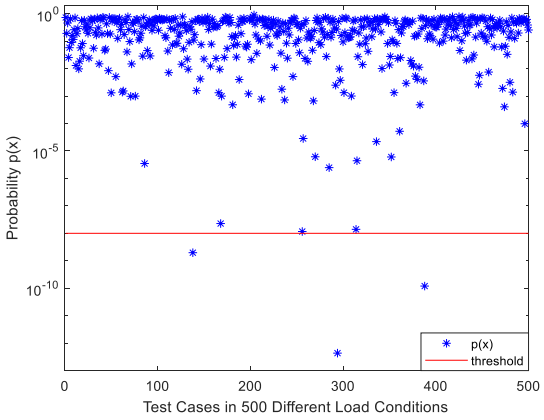

FIGURE 10. Detection results under 500 different normal conditions.

Fig. 10 shows the results under 500 normal operation conditions. As seen in this figure, 3 representative false positive results are marked. Compared with the results of sensitivity test in Fig. 9 and the results of normal condition test in Fig. 10, it can be observed that the detection is more accurate but less sensitive with the decrease of the threshold. Thus, the threshold should be determined experientially to obtain a better compromise between detection accuracy and detection sensitivity.

C. DISCUSSION

The proposed method can detect the false data by training one classifier (one multivariate Gaussian model), because the multivariate Gaussian model can capture the correlations between variables from different dimensions by itself with

TABLE 1. Nomenclature.

Notation	Description	Notation	Description	Notation	Description
$\hat{x}_T$	Transient feature vector	M_t	Original measurement	N_T	Number of training set
$\hat{x}_S$	Steady feature vector	$\dot{M}_t$	Tampered measurement	N_{CV}	Number of cross validation set
y	Anomaly flag	ΔM	Fluctuation in data sequence	N_S	Number of samples in sequence
Σ	Covariance matrix	$\hat{v}$	Voltage vector	n	Number of network conditions
ε	Anomaly threshold	$\hat{\mu}$	Centroid voltage vector	m	Number of voltage pair sets
f_p	Precision metric	V	Voltage measurement	L	Number of buses
f_r	Recall metric	$\dot{V}$	Predicted voltage measurement	K	Number of cluster sets
n_{fp}	Number of false positives	t_s	Start time of cyber-attack	a, b	Linear regression parameters
n_{fn}	Number of false negatives	t_e	End time of cyber-attack	$C\{\cdot\}$	Cluster set
n_{tp}	Number of true positives	p_s	Step attack parameter	$\lvert\cdot\rvert$	Absolute value
$p(\cdot)$	Probability value	p_r	Ramping attack parameter	$\lVert\cdot\rVert_2^2$	Squared Euclidean distance

the covariance matrix. In MATLAB, the training of a multivariate model with 1000 10-dimensional training samples costs less than 5 seconds. Concretely, in this paper, there are 600 4-dimensional samples (400 training samples and 200 cross-validation samples) of the transient cyber-attacks, and 600 5-dimensional samples (400 training samples and 200 cross-validation samples) of the steady cyber-attacks. The proposed approach is computationally effective to be used in practical applications as the model training can be accomplished by off-line calculation with historical data.

In the result of the transient cyber-attacks detection, both the accuracy and precision perform well as the distinction of the probabilities between system fault conditions and cyber-attack conditions is significant (as shown in Fig. 6). The reason is that we assume it unrealistic for opponents to implement successful cyber intrusion in each variable of measurements simultaneously. If any of the monitored measurement is not tampered, the feature of the normal operation fluctuations would lead to very small probability value in the multivariate Gaussian distribution. However, in the result of the steady cyber-attacks detection, the feature is established based on the prediction error, which can be influenced by some extraordinary network conditions. Generally, the threshold should be determined experientially to obtain a better compromise between the detection accuracy and the detection sensitivity. In this paper, the threshold is determined by maximizing the F_1 score in (6). From the comparison between Fig. 9 and Fig. 10, when the threshold is set higher, the detection is more sensitive to both the anomalies and the normal conditions with higher prediction errors. Therefore, an improved regression model with lower prediction errors can contribute to the detection accuracy based on the multivariate Gaussian model.

VI. CONCLUSION

This paper studies the detection of false data injected in the measurement sensors of CPS in distribution systems. A multivariate Gaussian based anomaly detection method is developed to identify transient and steady cyber-attacks. The comprehensive measurements in distribution systems can be acquired by μPMUs for statistical analysis. The multivariate features are established by formulating correlations among different variables of measurement data. A cyber-attack event is flagged when its probability in the multivariate Gaussian distribution is below a selected threshold. The threshold is determined by maximizing the F_1 score. The k-means clustering method is used to divide the distribution system into areas with similar voltage profiles, which are leveraged to formulate the linear regression models between the adjacent centroid voltage buses placed with the μPMUs. The numerical test is simulated on the IEEE 34 bus system. The test result has shown the proposed approach is sensitive to the false data injection and is computationally effective to be used in practical applications. In this paper, there is a trade-off between the detection accuracy and the detection sensitivity when selecting the threshold. In the future studies, the regression model should be improved to decrease prediction errors, which can contribute to more accurate regression model will contribute to the detection accuracy as well as the detection sensitivity.

APPENDIX

The nomenclature used in this paper is given in Table 1.

REFERENCES

[1] S. Xin, Q. Guo, H. Sun, B. Zhang, J. Wang, and C. Chen, "Cyber-physical modeling and cyber-contingency assessment of hierarchical control systems," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2375–2385, Sep. 2015.

[2] W. Liu, Q. Gong, H. Han, Z. Wang, and L. Wang, "Reliability modeling and evaluation of active cyber physical distribution system," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 7096–7108, Nov. 2018.

[3] K. Wan, D. Hughes, K. L. Man, T. Krilavicius, and S. Zou, "Investigation on composition mechanisms for cyber physical systems," *Int. J. Des., Anal. Tools Circuits Syst.*, vol. 2, no. 1, pp. 30–40, Aug. 2011.

[4] I. Yahyaoui, *Advances in Renewable Energies and Power Technologies*, vol. 1. Amsterdam, The Netherlands: Elsevier, 2018, pp. 480–482.

[5] Y. Liu, Y. Peng, B. Wang, S. Yao, and Z. Liu, "Review on cyber-physical systems," *IEEE/CAA J. Autom. Sinica*, vol. 4, no. 1, pp. 27–40, Jan. 2017.

[6] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber-physical system security for the electric power grid," *Proc. IEEE*, vol. 100, no. 1, pp. 210–224, Jan. 2012.

[7] C.-W. Ten, K. Yamashita, Z. Yang, A. V. Vasilakos, and A. Ginter, "Impact assessment of hypothesized cyberattacks on interconnected bulk power systems," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4405–4425, Sep. 2018.

[8] A. Srivastava, T. Morris, T. Ernster, C. Vellaithurai, S. Pan, and U. Adhikari, "Modeling cyber-physical vulnerability of the smart grid with incomplete information," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 235–244, Mar. 2013.
[9] A. Farraj, E. Hammad, A. A. Daoud, and D. Kundur, "A game-theoretic analysis of cyber switching attacks and mitigation in smart grid systems," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1846–1855, Jul. 2016.
[10] C.-W. Ten, J. Hong, and C.-C. Liu, "Anomaly detection for cybersecurity of the substations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 865–873, Dec. 2011.
[11] M. Cui, J. Wang, and M. Yue, "Machine learning based anomaly detection for load forecasting under cyberattacks," *IEEE Trans. Smart Grid*, to be published. doi: 10.1109/TSG.2018.2890809.
[12] Y. Liu, H. Xin, Z. Qu, and D. Gan, "An attack-resilient cooperative control strategy of multiple distributed generators in distribution networks," *IEEE Trans. Smart Grid*, vol. 7, no. 6, pp. 2923–2932, Nov. 2016.
[13] A. Ashok, M. Govindarasu, and J. Wang, "Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid," *Proc. IEEE*, vol. 105, no. 7, pp. 1389–1407, Jul. 2017.
[14] Y. W. Law, T. Alpcan, and M. Palaniswami, "Security games for risk minimization in automatic generation control," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 223–232, Jan. 2015.
[15] X. Liu, P. Zhu, Y. Zhang, and K. Chen, "A collaborative intrusion detection mechanism against false data injection attack in advanced metering infrastructure," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2435–2443, Sep. 2015.
[16] H. Wang, J. Ruan, G. Wang, B. Zhou, Y. Liu, X. Fu, and J. Peng, "Deep learning-based interval state estimation of AC smart grids against sparse cyber attacks," *IEEE Trans. Ind. Informat.*, vol. 14, no. 11, pp. 4766–4778, Nov. 2018.
[17] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1644–1652, Sep. 2017.
[18] M. Cui, J. Wang, J. Tan, A. R. Florita, and Y. Zhang, "A novel event detection method using PMU data with high precision," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 454–466, Jan. 2019.
[19] L. Xie, Y. Chen, and P. R. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2784–2794, Nov. 2014.
[20] S. S. Negi, N. Kishor, K. Uhlen, and R. Negi, "Event detection and its signal characterization in PMU data stream," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 3108–3118, Dec. 2017.
[21] D.-I. Kim, T. Y. Chun, S.-H. Yoon, G. Lee, and Y.-J. Shin, "Wavelet-based event detection method using PMU data," *IEEE Trans. Smart Grid*, vol. 8, no. 3, pp. 1154–1162, May 2017.
[22] F. Zhang, L. Cheng, X. Li, Y. Sun, W. Gao, and W. Zhao, "Application of a real-time data compression and adapted Protocol technique for WAMS," *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 653–662, Mar. 2015.
[23] M. Jamei, A. Scaglione, C. Roberts, E. Stewart, S. Peisert, C. McParland, and A. McEachern, "Anomaly detection using optimally placed μPMU sensors in distribution grids," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3611–3623, Jul. 2018.
[24] F. R. Gomez, A. D. Rajapakse, U. D. Annakkage, and I. T. Fernando, "Support vector machine-based algorithm for post-fault transient stability status prediction using synchronized measurements," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1474–1483, Aug. 2011.
[25] Y. Li and Z. Yang, "Application of EOS-ELM with binary jaya-based feature selection to real-time transient stability assessment using PMU data," *IEEE Access*, vol. 5, pp. 23092–23101, 2017.
[26] E. Kyriakides, and P. Marios, *Intelligent Monitoring, Control, and Security of Critical Infrastructure Systems*. New York, NY, USA: Springer, 2014, pp. 1–30.
[27] D. W. J. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 58–69, Jan. 2002.
[28] Q. Fu, J.-G. Lou, Y. Wang, and J. Li, "Execution anomaly detection in distributed systems through unstructured log analysis," in *Proc. 9th IEEE Int. Conf. Data Mining*, Miami, FL, USA, Dec. 2009, pp. 149–158.
[29] R. C. Dugan and T. E. McDermott, "An open source platform for collaborating on smart grid research," in *Proc. IEEE Power Energy Soc. General Meeting*, San Diego, CA, USA, Jul. 2011, pp. 1–7.
[30] *OpenDSS Program, Through Sourceforge.Net*. Accessed: 2019. [Online]. Available: http://sourceforge.net/projects/electricdss

YU AN received the B.Sc. degree in electrical engineering from Shanghai Jiao Tong University, China, in 2014, where he is currently pursuing the Ph.D. degree in electrical engineering. His research interests include anomaly detection in cyber-physical distribution system, resilience operation of distribution system, and microgrids.

DONG LIU received the B.S. and M.S. degrees from Sichuan University, China, in 1989 and 1994, respectively, and the Ph.D. degree from Southeast University, China, in 1997. He is currently a Professor with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, China. His research interests include smart grid, active distribution network, and cyber-physical system in power grid.