Outage Detection in Partially Observable Distribution Systems Using Smart Meters and Generative Adversarial Networks

Yuxuan Yuan, *Graduate Student Member, IEEE*, Kaveh Dehghanpour, *Member, IEEE*,
AQ1 Fankun Bu, *Graduate Student Member, IEEE*, and Zhaoyu Wang, *Member, IEEE*

***Abstract*—In this paper, we present a novel data-driven approach to detect outage events in partially observable distribution systems by capturing the changes in smart meters' (SMs) data distribution. To achieve this, first, a breadth-first search (BFS)-based mechanism is proposed to decompose the network into a set of *zones* that maximize outage location information in partially observable systems. Then, using SM data in each zone, a generative adversarial network (GAN) is designed to implicitly extract the temporal-spatial behavior in *normal conditions* in an unsupervised fashion. After training, an anomaly scoring technique is leveraged to determine if real-time measurements indicate an outage event in the zone. Finally, to infer the location of the outage events in a multi-zone network, a zone coordination process is proposed to take into account the interdependencies of intersecting zones. We have provided analytical guarantees of performance for our algorithm using the concept of *entropy*, which is leveraged to quantify outage location information in multi-zone grids. The proposed method has been tested and verified on distribution feeder models with real SM data.**

***Index Terms*—Generative adversarial networks, outage detection, partially observable system, smart meter, zone.**

NOMENCLATURE

AMI	Advanced metering infrastructure
BFS	Breadth-first search
GAN	Generative adversarial network
SM	Smart meter
$A \setminus B$	Elements of set A that are not in set B
$A \succ B$	A has a higher topological order than B
B_c	Candidate branch set that are potentially the location of outage event
B_g	Set of grid branches
$cos\phi_i$	Power factor of node i
D	Discriminator
G	Generator
$H(\cdot)$	Entropy function for assessing outage location information
$\mathbf{I}_{i-1,i}$	Branch current between nodes $i-1$ and i
$\mathbf{K}_{i-1,i}$	Approximate voltage drop factor
$\mathbf{l}_{i-1,i}$	Length of distribution line segment between nodes $i-1$ and i
m	Number of batch size
M	Number of branches in the system
n_D	Number of iterations for D per G iteration
$N(g)$	Set of neighboring nodes in the grid
O	Number of observable nodes
O_{end}	Number of observable nodes that do not have any observable downstream nodes
P_i	Power consumption of node i
$\Delta\mathbf{P}_s$	Outage event magnitude
$P_{X_{\Psi_i}}$	Probability density function of historical data in zone Ψ_i
S_r	Network's root node
S_g	Set of observable nodes in the partially observable distribution system
S_{o1}	Upstream observable node of zone
S_{o2}	Downstream observable node of zone
T	Length of the time window
u_k	k'th set of branches that are covered with the exact same set of zones
$U(\Psi^g)$	Undetectable branch set for the selected zone set Ψ^g
V_O	Number of zones containing the faulted branch in the system
$\lvert\mathbf{V}_i\rvert$	Voltage magnitude measurements at node i
$\Delta\mathbf{V}$	Voltage drop value in normal condition
$\Delta\mathbf{V}_o$	Post-outage voltage drop value
X_{Ψ_i}	Training dataset for zone Ψ_i
z	Noise signal with uniform distribution
z^*	Optimal solution for residual error
Z_{Ψ_i}	Set of branch in zone Ψ_i
$\mathbf{Z}_{(i-1,i),abc}$	Phase impedance matrix between nodes $i-1$ and i
α	Learning rate
$\delta_R(\cdot)$	Residual error
$\delta_D(\cdot)$	Discriminator error

AQ2 Manuscript received December 10, 2019; revised March 31, 2020 and June 4, 2020; accepted July 8, 2020. This work was supported in part by the Advanced Grid Modeling Program at the U.S. Department of Energy Office of Electricity under Grant DE-OE0000875, and in part by the National Science Foundation under Grant ECCS 1929975. Paper no. TSG-01853-2019. *(Corresponding author: Zhaoyu Wang.)*

The authors are with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: yuanyx@iastate.edu; wzy@iastate.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2020.3008770

$\gamma^g(b_j)$	Set of zones in the grid that include branch b_j
λ	Weight factor for combining $\delta_R(\cdot)$ and $\delta_D(\cdot)$
μ_{Ψ_i}	Sample mean of the anomaly scores for the training dataset in zone Ψ_i
ω	Number of zones in the system
Ψ_i	i'th outage detection zone
Ψ^g	Set of all selected zones for the partially observable grid
Ψ_a	Target zone containing the maximum information on the outage event
σ_{Ψ_i}	Sample variance of the anomaly scores for the training dataset in zone Ψ_i
θ_G, θ_D	Learning parameters for G and D
ζ_{Ψ_i}	GAN-based anomaly score in zone Ψ_i.

I. Introduction

OUTAGE detection is a challenging problem in power systems, especially in distribution networks where the majority of outage events take place. According to the statistical data provided by the U.S. Energy Information Administration, each customer lost power for around 4 hours on average in 2016 [1]. To decrease outage duration, and improve system reliability and customer satisfaction, distribution system operators (DSOs) deploy state-of-the-art outage management systems, using modern software tools and protection devices with bidirectional communication function. This allows DSOs to collect real-time up-to-the-second data from the network [2]. Nevertheless, use of intelligent communication-capable devices in distribution systems has not become prevalent, mostly due to budgetary limitations of utilities [3]. Hence, identification of distribution system outage events, especially for small utilities, still relies on trouble calls from customers and manual inspection. However, trouble calls alone are not a reliable data source of outage detection because customers may not make prompt calls to utilities [4]. Also, conventional expert-experience-based outage discovery methods that use customer calls are laborious, costly, and time-consuming [5].

In recent years, a number of papers have explored data-driven alternatives for outage detection. According to the type of data source, the previous works in this area can be classified into two groups: *Class I - Smart meter (SM)-based methods*: With the widespread deployment of advanced metering infrastructure (AMI), SMs provide an opportunity to rapidly detect outage events by recording the real-time demand consumption and automatically sending "last gasp" signals to the utilities. In [6], a multi-label support vector machine classification method is presented that utilizes the last gasp signals of SMs to detect and find the locations of damaged lines in fully observable networks. In [7], a hierarchical framework is developed to provide anomaly-related insights using multivariate event counter data collected from SMs. In [8], a fuzzy Petri nets-based approach is proposed to detect nontechnical losses and outage events by tracking the differences between profiled and irregular power consumption. In [9], a probabilistic and fuzzy model-based algorithm is presented to process outage data using AMI. In [10], a tree-based polling algorithm is developed to obtain information about the system conditions by polling local SMs. *Class II - non-SM-based methods*: Other data sources have been used in the literature for outage detection, as well. In [2], a hypothesis testing-based outage detection method is developed combining the use of real-time power flow measurements and load forecasts of the nodes. In [4], a social network-based data-driven method is proposed by leveraging real-time information extraction from Twitter. In [11], a new boosting algorithm is developed to estimate outages in overhead distribution systems by utilizing weather information.

Even though previous works provide valuable results, critical questions remain unanswered in this area. The limitation of most Class I models is their basic assumption that the distribution system is *fully observable*, i.e., all the nodes have measurement devices. However, this assumption does not necessarily apply to practical systems, in which large portions of customers do not own smart meters [6]. On the other hand, Class II methods are generally based on several limiting assumptions, such as availability of accurate forecasts for customer loads, availability of real-time power flow measurements, and reliability of social network data. Another difficulty in outage detection is *outage data scarcity*, which means that the size of the outage data is far smaller compared to the data in normal conditions. This issue causes a *data imbalance problem* that could hinder reliable training of supervised learning-based outage detection models [12].

To address these shortcomings, in this paper, a generative adversarial network (GAN)-based method is developed to detect power outages in partially observable distribution systems by capturing the anomalous changes in SMs' measurement data distributions that are caused by outage events [13]. Compared to the previous works, the proposed method solves three fundamental challenges in outage monitoring for partially observable distribution systems: 1) Unlike supervised classifiers that can fail in case of outage data scarcity, the proposed generative model follows an *unsupervised learning* style which only relies on the operation data in normal conditions for model training. Then, a GAN-based anomaly score is defined to quantify the deviations between the learned distribution and the real-time measurements to detect potential outage events, i.e., new observations with high anomaly scores imply outage [14]. 2) Due to the temporal variability of AMI data, efficient outage detection requires capturing high-dimensional temporal-spatial relationships in measurement data. Conventional data distribution estimators are limited by the high-dimensional nature of the data. Instead of constructing a complex data likelihood function explicitly, our approach trains GANs to implicitly extract the underlying distribution of the data. Each GAN consists of two interconnected deep neural networks (DNNs) [15]. 3) Considering the partial observability of real systems, we have proposed a breadth-first search (BFS)-based mechanism to decompose large-scale distribution networks into a set of intersecting *zones* [16]. Each zone is determined by two neighboring observable nodes of the network (i.e., nodes with known voltages and demands) and contains only a subset of network branches. A separate GAN is trained in each zone using the time-series data of the two

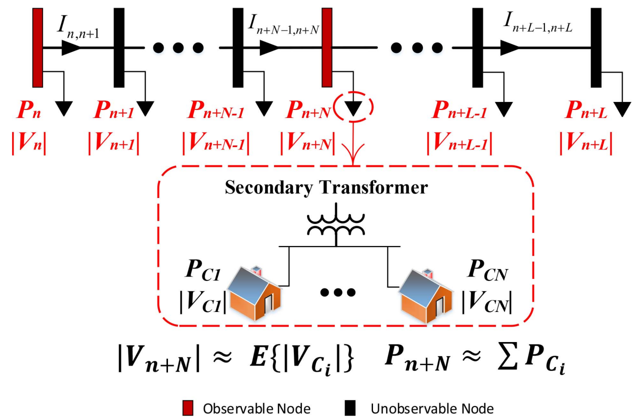

Fig. 1. Example zone in normal condition.

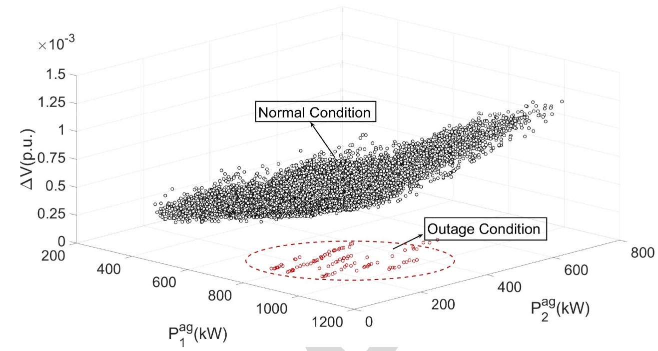

Fig. 2. Joint data distribution under normal and outage conditions.

observable nodes. Since sectionalizing networks into multiple zones can be done in more than one way depending on the choice of observable nodes, it is necessary to find the optimal set of zones. Our BFS-based approach optimizes the zone selection and anomaly score coordination process and achieves maximum outage location information. To demonstrate this, we have proposed an outage detection metric based on the information-theoretic concept of *entropy* to quantify outage location information. The proposed outage detection methodology has been tested and verified using real AMI data and network models.

II. Real Data Description and Zone Selection

A. AMI Data Description

The historical AMI data used in this paper contains several U.S., mid-west utilities' hourly energy consumption data (kWh) and voltage magnitude measurements of over 6000 customers [17]. The dataset includes around four years of measurements, from January 2015 to May 2018. Over 95% of customers are residential and commercial loads in the dataset. The hourly data was initially processed to remove bad and missing data caused by communication error.

B. Outage Detection Zone Definition

When an outage happens in a radial system, a protective device isolates the faulted area along with the loads downstream of the fault location [2]. This will cause the measurement data samples from unfaulted upstream observable nodes to deviate from the data distribution in normal condition. In this paper, we exploit this phenomenon to define an outage detection zone.

In general, two observable nodes (i.e., nodes with AMI-based measured voltage magnitudes and power consumption) can be utilized to detect an outage happening on the paths downstream of the two nodes. To show this, Fig. 1 presents a typical distribution feeder with two observable nodes, node n and node $n+N$. Given the radial structure of the feeder, the voltage drop, ΔV, between nodes n and $n+N$ can be expressed as [18]:

$$\Delta \mathbf{V} = |\mathbf{V}_n| - |\mathbf{V}_{n+N}| \approx |\sum_{i=n+1}^{n+N} \mathbf{Z}_{(i-1,i),abc} \cdot \mathbf{I}_{i-1,i}| \qquad (1)$$

where, $|\mathbf{V}_n|$ and $|\mathbf{V}_{n+N}|$ are the voltage magnitude measurements of the observable nodes, $\mathbf{I}_{i-1,i}$ and $\mathbf{Z}_{(i-1,i),abc}$ are the branch current and the phase impedance matrix between bus $i-1$ and i. Dimensions of the variables in (1) depend on the number of phases of distribution lines. For example, for a three-phase feeder $|\mathbf{V}_n|$, $|\mathbf{V}_{n+N}|$ and $\mathbf{I}_{i-1,i}$ are 3-by-1 vectors, and $\mathbf{Z}_{(i-1,i),abc}$ is a 3-by-3 matrix. The above equation can be rewritten in terms of nodal power measurements, as follows [18]:

$$\Delta \mathbf{V} \approx \sum_{i=n+1}^{n+N} \sum_{j=i}^{n+L} \mathbf{K}_{i-1,i} \otimes \mathbf{l}_{i-1,i} \otimes \frac{\mathbf{P_j}}{cos\phi_j} \qquad (2)$$

where, $n+L$ is the total length of this path, $\mathbf{K}_{i-1,i}$ $[\frac{\%drop}{kVA \cdot mile}]$ and $\mathbf{l}_{i-1,i}$ are the approximate voltage drop factor and the length of distribution line segment between nodes $i-1$ and i, $\mathbf{P}_j$ and $\cos\phi_j$ represent the nodal power consumption and the power factor at node j. Here, $\mathbf{K}_{i-1,i}$, $\mathbf{l}_{i-1,i}$, and $\mathbf{P}_j$ are 3-by-1 vectors, and $\otimes$ denotes element-wise multiplication. When outage happens at an unobservable node s downstream of node n, $n+1 \le s \le n+L$, the post-outage voltage drop value, ΔV_o, is determined as follows:

$$\Delta \mathbf{V}_o \approx \Delta \mathbf{V} + \sum_{i=n+1}^{\min(s,n+N)} \mathbf{K}_{i-1,i} \otimes \mathbf{l}_{i-1,i} \otimes \frac{\Delta \mathbf{P}_s}{cos\phi_s} \qquad (3)$$

where, $\Delta \mathbf{P}_s$ represents the outage event magnitude and has a negative value. Comparing (3) with (2), we can observe that the voltage drop value across the two observable nodes changes after an outage event downstream of any of the two nodes. These changes are almost proportional to the outage magnitude, ΔP_s. This can also be confirmed using real AMI data, as shown in Fig. 2, where P_1^{ag} and P_2^{ag} are the aggregated power consumption of the first and second observable nodes in a zone. This figure shows the perceivable gap between the joint data distribution obtained from two observable nodes under normal and one specific outage condition, in three dimensions. Given that an outage event anywhere downstream of the two nodes will lead to deviations from their underlying joint measurement data distribution in normal operations, we define an outage detection zone as follows.

Definition 1: In a radial network, an outage detection zone, Ψ_i, is defined as $\Psi_i = \{S_{o1}, S_{o2}, Z_{\Psi_i}\}$ where S_{o1} and S_{o2} are two observable nodes, with S_{o1} being upstream of S_{o2}, and Z_{Ψ_i} is the set of all the branches downstream of S_{o1}.

C. Zone Selection

Based on Definition 1, for a specific distribution system, different zone selection strategies can result in different zone partitioning, which will impact the performance of outage detection and location. Hence, we propose a BFS-based zone selection method by exploiting the tree-like structure of distribution systems in this paper. Specifically, our method selects the zones using nodes at the present depth before moving on the nodes at the next depth level. As will be elaborated in Section IV, the proposed zone selection algorithm offers two advantages: (1) it is able to obtain the *optimal zone set* that maximizes the outage location information in arbitrary partially observable network. (2) The proposed BFS-based algorithm introduces a *valid topological ordering*, which significantly simplifies outage location identification process. Prior to discussing the zone selection algorithm, we provide the following useful definition [19].

Definition 2: In a radial network, node B is defined as an *immediate observable downstream node* for an arbitrary node A if two conditions are satisfied: 1) node B is located downstream of node A; 2) the path that connects A and B consists only of unobservable nodes.

The proposed algorithm involves the following steps:

- *Step I:* Consider a partially observable distribution system, g, with a total number of M branches, $B_g = \{b_1, \ldots, b_M\}$, and a set of $O+1$ observable nodes, $S_g = \{S_r, S_1, S_2, \ldots, S_O\}$, where S_r represents the network's root node (i.e., main substation).
- *Step II:* Define and initialize the zone set and the neighboring node set for g, as Ψ^g and $N(g) = \{\emptyset\}$. Note that the set Ψ^g is an *ordered set*, where new elements are added to the right side of the current elements in the set (i.e., order of elements matters). Initialize the set of candidate observable nodes as $S_B = \{S_r\}$, and the zone counter $k \leftarrow 1$.
- *Step III:* If $N(g) = \{\emptyset\}$, randomly select and then remove a node, S_{o1}, from S_B. Else if $N(g) \neq \{\emptyset\}$, randomly select and remove a node, S_{o1}, from $N(g)$.
- *Step IV:* Find all the immediate observable nodes downstream of S_{o1} (see Definition 2), and randomly select a node from this set, which is denoted as S_{o2}. If $N(g) = \{\emptyset\}$, add all the immediate observable nodes downstream of S_{o1} to $N(g)$; otherwise, add them to S_B.
- *Step V:* Select a new zone Ψ_k, with S_{o1} and S_{o2}, and include all the branches downstream of S_{o1} into Z_{Ψ_k} (see Definition 1). Add Ψ_k to the right side of the current zones in Ψ^g.
- *Step VI:* $k \leftarrow k+1$. Go back to Step III until $N(g)$ is empty for all the nodes in S_B.
- *Step VII:* Output the ordered set of all network zones, $\Psi^g = \{\Psi_1, \ldots, \Psi_w\}$, with w denoting the number of selected zones.

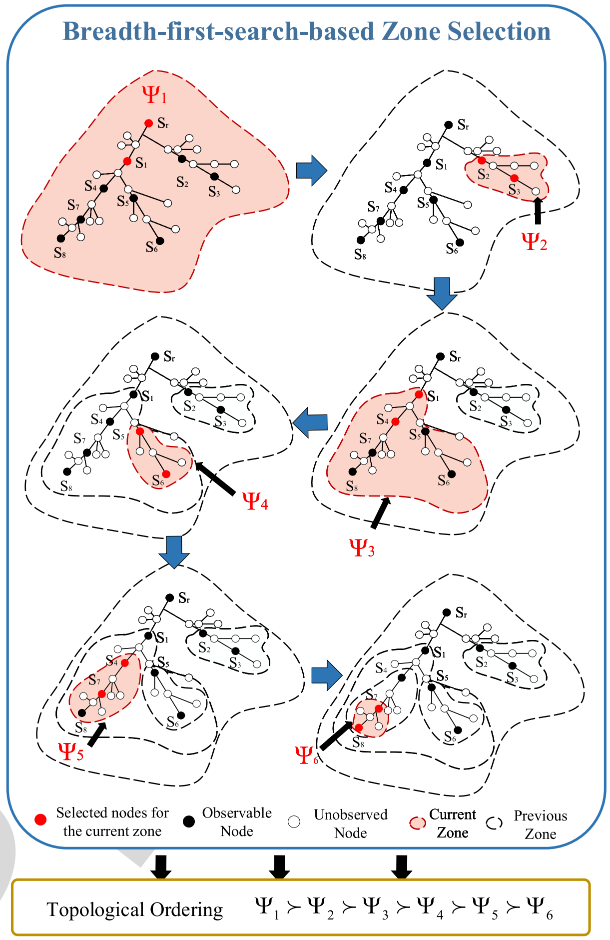

Fig. 3. Proposed BFS-based zone selection and ordering method.

To help the reader understand each step of the algorithm, an example of zone selection is shown in Fig. 3. In this exemplary system, $B_g = \{b_1, \ldots, b_{36}\}$ and $S_g = \{S_r, S_1, \ldots, S_8\}$. In the first iteration ($k = 1$), Ψ^g and $N(g)$ are both empty, $\{\emptyset\}$; In Step II, the root node is selected to be the first observable node, $S_B = \{S_r\}$. In Step III, since $N(g)$ is empty, S_{o1} is randomly selected and then removed from S_B; thus, $S_{o1} \leftarrow S_r$ and $S_B \leftarrow \{\emptyset\}$. In Step IV, S_1 and S_2 are identified as the immediate observable downstream nodes of S_r. Since $N(g)$ is empty, these two nodes are added to $N(g)$. Then, S_{o2} is selected randomly from $\{S_1, S_2\}$. In this example, $S_{o2} \leftarrow \{S_1\}$. In Step V, the first zone is defined based on the selected S_{o1} and S_{o2} and added to the set Ψ^g; $\Psi^g = \{\Psi_1\}$, where $\Psi_1 = \{S_r, S_1, Z_{\Psi_1}\}$. The algorithm will go back to Step III for the next iteration ($k \leftarrow k+1$).

Following the proposed zone selection method, the number of zones, ω, can be represented as a function of number of observable nodes: $\omega = O + 1 - O_{\text{end}}$, where O is the number of all observable nodes and O_{end} is the number of observable nodes that do not have any observable downstream nodes. This function indicates that the proposed method needs sensors installation at internal nodes to develop a meaningful zone partitioning. This requirement is consistent with the recent expansion of smart grid monitoring devices. In current distribution systems, metering devices are generally installed at some select locations, such as at the root node and other major utility equipment, which can be utilized to obtain a zone partitioning [20]. On the other hand, in many distribution systems monitoring devices are only installed at the terminal nodes, as claimed in [21]. To handle zone selection is such systems, we have provided an approximation method. Prior to discussing the method, we define passive and active internal nodes: active internal nodes are the subset of network internal nodes with non-zero current injection. In contrast, passive internal nodes do not have any current injection. The basic idea of this method is to utilize a part of measurement data of observed terminal nodes to represent their nearest unobserved passive internal nodes. The rationale behind this approximation is that the voltage drop between passive internal nodes internal nodes and the nearest terminal nodes can often be ignored. Using this approximation, the proposed approach can develop a reasonable zone partitioning when only terminal buses are metered. It should be noted that similar strategy has been utilized in previous works for learning the topology of distribution systems [22].

When the zone set is obtained, each branch in the system will belong to at least one zone, while at the same time, no two zones have the exact same set of branches. For example, branches of the zone Ψ_6 in Fig. 3, are also covered by zones $\Psi_1, \ldots, \Psi_5$. As will be shown in Section IV, these inter-zonal intersections introduce a *redundancy*, which will be leveraged for enhancing the robustness of the outage detection process by blocking bad data samples and outliers. Furthermore, to specify the outage location considering the zonal intersections, a zone coordination method is proposed in Section III.

III. GAN-Based Zone Monitoring

In this paper, to quantify deviations from the measurement data distribution in normal conditions caused by outage events, we have utilized a recently-invented non-parametric unsupervised learning approach, GAN [23]. One unique advantage of GAN is its ability to implicitly represent complex data distributions without constructing high-dimensional likelihood functions, thus addressing the challenge of dimensionality. Moreover, GAN does not assume a prior parametric structure over the data distribution. This ensures the performance of GAN for outage detection problem, since the utilities generally do not have a prior knowledge of the exact structure of data distribution in normal conditions. Meanwhile, since model training is done using only the data from normal condition, our method is not vulnerable to the outage data scarcity problem. When training is completed, a GAN-based anomaly score is assigned to real-time measurements to detect outage events inside the zone [14].

A. GAN Fundamentals and Training Process

For each zone, a GAN is trained to learn the joint distribution of measured variables $X = \{\Delta V^t, P_n^t, P_{n+N}^t\}_{t=1}^{T}$ within a time-window with length T (see Fig. 1), where P_n^t and P_{n+N}^t are the nodal power consumption for the two observable nodes in the zone, and ΔV^t is the voltage difference between the two nodes at time t. The purpose of defining a time-window over the observable variables is to exploit temporal relations between consecutive data samples in power distribution systems for more effective anomaly detection. In this paper, T is selected to be 3 hours based on calibration results from the grid search method [24]. It should be noted that the training procedure of GANs is an offline process; as a result, the high computational cost of the grid search approach does not impact the real-time performance of the proposed method. The training set consists of the SM data history of the variables defined in each zone, and is denoted as X_{Ψ_i} for zone Ψ_i. To account for the strong seasonal changes in customers' behavior that might mislead detecting the boundary between normal and outage behavior [25], the dataset has been decomposed into separate seasons to train different GAN models for each zone. Each dataset is randomly divided into three separate subsets for training (70% of the total data), validation (15% of the total data), and testing (15% of the total data).

GAN relies on two interconnected DNNs, which are simultaneously trained via an adversarial process: a *generator*, G, and a *discriminator*, D [26], as shown in Fig. 4 (part A). The interaction between the two DNNs can be modeled as a game-theoretic two-player nested minmax optimization [13]:

$$\min_{\theta_G} \max_{\theta_D} V(D, G) = \mathbb{E}_{x_{\Psi_i} \sim p_{X_{\Psi_i}}(x_{\Psi_i})}[\log(D(x_{\Psi_i}))] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z)))] \quad (4)$$

where, θ_G and θ_D are the learning parameters of G and D, respectively. $p_{X_{\Psi_i}}$ is the underlying probability density function of historical data obtained from the two observable nodes of the zone. In each iteration, D is trained to maximize the probability of assigning the correct label to both training examples and artificially generated samples from G. Thus, the output of D, $0 \leq D(x_{\Psi_i}) \leq 1$, represents the probability that x_{Ψ_i} is from the training dataset rather than generated artificially by G [13]. On the other hand, G is trained to generate artificial samples that maximize the probability of the discriminator D mislabeling. The input of G is defined as $z \in \mathbb{R}^{d \times 1}$, which is a noise signal with uniform distribution $p_z(z)$. In this case, $d = 4$ showed the best performance on the validation set. After a number of training iterations, G and D will reach a unique global optima at which both cannot improve. This means the generator can recover the underlying distribution of the training data and the discriminator cannot distinguish the true samples from the artificially generated samples [27]. The training process takes place offline and the detailed procedure is presented in Algorithm 1. In this paper, the hyperparameter set of GAN is calibrated by using the random search algorithm [28]; as a result, G consists of three components: an input layer of 4 neurons, two hidden layers of 8 neurons, and an output layer of 9 neurons. D also has three parts: an input layer of 9 neurons, two hidden layers of 8 neurons, and an output layer of 1 neuron. Moreover, $\{\alpha, m, n_D\}$ are selected as 0.01, 100, 1, respectively. To update θ_G and θ_D, a minibatch stochastic gradient descent method is utilized [13].

B. GAN-Based Anomaly Score Assignment

To detect potential outage events in each zone, a GAN-based anomaly score is utilized to evaluate sequential measurements

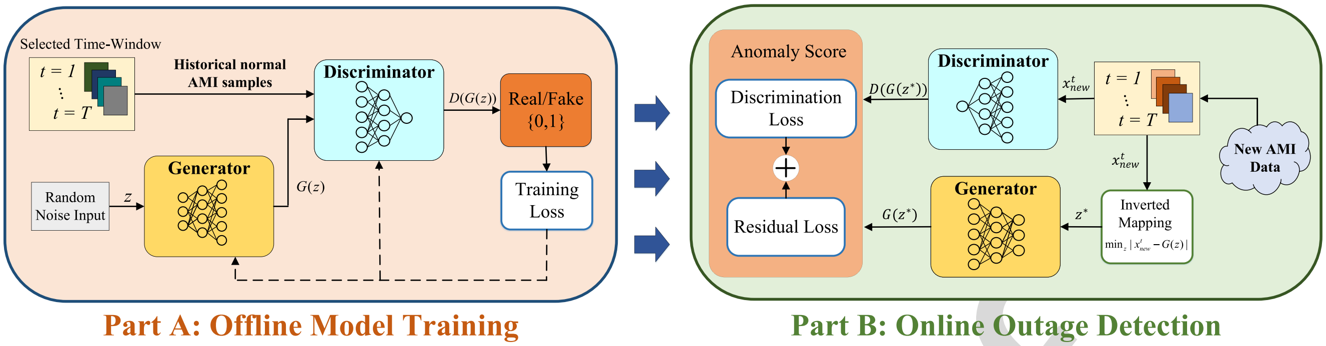

Fig. 4. GAN-based learning and testing structure.

Algorithm 1 GAN Training for Zone Ψ_i

Require: : Seasonal normal behavior data for zone Ψ_i
Require: : Learning rate α, batch size m, number of iterations for D per G iteration n_D, initial learning parameters for G and D, θ_D and θ_G

while Nash equilibrium has not been achieved **do**
  for $t = 0, \dots, n_D$ **do**
    Generate sample batch from the latent space z
    $p_z \rightarrow \{(z_j)\}_{j=1}^{m}$
    Obtain sample batch from the historical data
    $p_{X_{\Psi_i}} \rightarrow \{x_{\Psi_i}(j)\}_{j=1}^{m}$
    Update discriminator parameters using gradient descent with α based on the discriminator loss
    $\delta_D = \frac{1}{m}\sum_{j=1}^{m}[-\log D(x_{\Psi_i}(j)) - \log(1 - D(G(z_i)))]$
    $\theta_D := \theta_D - \alpha * \nabla_{\theta_D}\delta_D$
  end for
  Update generator parameters using gradient descent with α
  $\delta_G = \frac{1}{m}\sum_{j=1}^{m}[-\log D(G(z_j))]$
  $\theta_G := \theta_G - \alpha * \nabla_{\theta_G}\delta_G$
end while

of SMs online [14], as shown in Fig. 4 (part B). The anomaly score consists of two error metrics: the residual error, $\delta_R(\cdot)$, and the discriminator error, $\delta_D(\cdot)$. When a new data inquiry $x_{new}^t \in \mathbb{R}^{3T\times 1}$ is obtained (at the T time slots), the residual error describes the extent to which x_{new}^t follows the learned distribution of the G model, in the best case [14]:

$$\delta_R(x_{new}^t) = \min_z |x_{new}^t - G(z)| \tag{5}$$

After training, the generator, G, has learned an almost perfect mapping from the latent space z to the zonal measurement data distribution in normal conditions. Hence, if x_{new}^t is obtained from normal conditions, its residual error value is zero, $\delta_R(x_{new}^t) = 0$, since x_{new}^t and $G(z^*)$ are identical, where z^* is the optimal solution to (5). To obtain z^* during test time, a commercial nonlinear programming solver, "fmincon", is used in this work. Thus, higher $\delta_R(x_{new}^t)$ values represent deviations from normal operation conditions, suggesting occurrence of outage event within the zone.

The discriminator error, $\delta_D(x_{new}^t)$, is defined using the trained discriminator, D, to measure how well $G(z^*)$ follows the learned data distribution by the G model. The discriminator error can be written as [13]:

$$\delta_D(x_{new}^t) = -\log D(x_{new}^t) - \log(1 - D(G(z^*))) \tag{6}$$

The GAN-based anomaly score for zone Ψ_i is defined as the weighted sum of both error metrics [14]:

$$\zeta_{\Psi_i}(x_{new}^t) = (1-\lambda)\cdot\delta_R(x_{new}^t) + \lambda\cdot\delta_D(x_{new}^t) \tag{7}$$

where, $0 \le \lambda \le 1$ is a user-defined weight factor, the value of which is set at 0.1 in this paper, based on calibration results over the validation set. To determine the critical threshold for the anomaly score, above which new data points are identified as outage events, the GAN-based anomaly score, ζ_{Ψ_i}, is obtained for all training data samples of zone Ψ_i. The sample mean, μ_{Ψ_i}, and the sample variance, σ_{Ψ_i}, of the anomaly scores for the training data samples are calculated to determine the range of anomaly score in normal operations. When outage occurs, the real-time measurement data samples are expected to have anomaly scores above this range. We have used a rolling window approach in this work. Hence, the test point could use $T-1$ measurements before an outage, thus, we can detect an outage within one-time interval. The length of the time interval depends on the resolution of the smart meter data. The details of anomaly identification process are elaborated in the next section.

C. GAN-Based Zone Coordination

Using the trained GANs and GAN-based anomaly score method, outage events can be detected in each zone by comparing the anomaly scores between the new inquiry samples and the critical threshold. Considering that a GAN is trained for each zone, a high anomaly score only gives a rough estimation of event location by simply implying outage somewhere in the zone. In other words, all branches in the zone are the candidate event locations. Specifically, if we treat the whole grid as a single zone (i.e., if only a single GAN is trained for the whole grid), then a high anomaly score will only indicate that an outage has occurred somewhere in the system without any detailed location information. Since the granularity of location information depends on the number of candidate branches, it is necessary to reduce this number as much as possible. To achieve this, we have presented a GAN-based zone coordination method by integrating anomaly scores from multiple zones, which includes the following steps:

- *Stage I:* Assign a GAN to each zone, $\Psi_i \in \Psi^g$ and use Algorithm 1 over the historical seasonal data of the two observable nodes of each zone to learn the joint distribution of the measurement data.
- *Stage II:* After training for each zone, Ψ_i, obtain the anomaly score for training samples in the zone; determine the anomaly score sample mean and sample variance, denoted as μ_{Ψ_i} and σ_{Ψ_i}, respectively.
- *Stage III:* At time T, observe the anomaly scores of all the zones in the set Ψ^g based on the latest real-time measurements.
- *Stage IV:* Select the first zone from the right side of the set Ψ^g that has an abnormal anomaly score value and denote it as Ψ_a. We will show that this zone contains the maximum information on the outage event in Section IV. In other words, $a = \arg\max_{\xi} \xi,\ s.t.\ \zeta_{\Psi_\xi} > \mu_{\Psi_\xi} + h \cdot \sigma_{\Psi_\xi}$, where, h is a user-defined threshold factor.
- *Stage V:* Output the set of candidate branches that are potential locations of outage event as $B_c = Z_{\Psi_a} \setminus \{Z_{\Psi_{a+1}} \cup Z_{\Psi_{a+2}} \cup \cdots \cup Z_{\Psi_\omega}\}$, where $A \setminus B$ represents the elements of set A that are not in set B. Further, $\{\Psi_{a+1}, \Psi_{a+2}, \ldots, \Psi_\omega\}$ are the zones that have lower topology ordering than Ψ_a.

Based on the outcome of zone coordination, the DSO can obtain the minimum branch candidates that are potentially impacted by the outage, thus maximizing the outage information. This process will help the repair crew to rapidly find the outage location. Note that given the unbalanced nature of distribution networks, the proposed algorithm is applied to each phase separately. Hence, the zone set needs to be obtained for three phases. For the sake of conciseness we will continue our discussions for one phase, keeping in mind that the same logic applies to the other phases as well.

In practice, the distribution system often undergoes reconfiguration, on-load tap changing, and capacitor switching, which can strongly affect the actual data distribution. Thus, the proposed outage detection method needs to be customized to account for the effects of these events as well, as shown in Fig. 5. The basic idea is to integrate pre-trained GANs and fine-tuning strategy. Considering that the zone selection process and the training procedure of GANs are offline processes, the utility can obtain the zone sets and the corresponding GAN library in advance using historical data. When a capacitor switching occurs and raises an anomaly score flag, the existing GANs are treated as the pre-trained models which still maintain useful information. The new measurements from the observable nodes are utilized to fine-tune these pre-trained GANs for adapting to the changes of the underlying data distribution. The fine-tuning strategy can counter the overfitting problem on small datasets, thus, reducing the data size requirement [22].

IV. Theoretical Properties of the Proposed Framework

In this section, we discuss the theoretical properties of the proposed outage-detection framework. We will show that this approach has three fundamental properties.

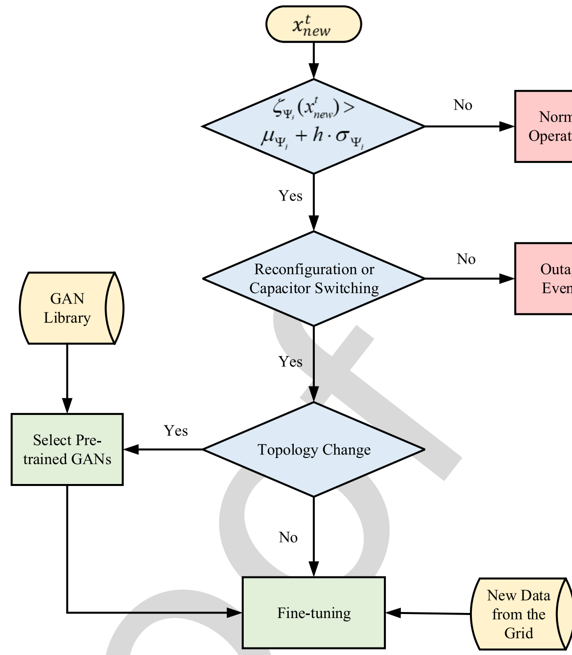

Fig. 5. Flowchart of the proposed method considering possibility of reconfiguration, on-load tap changing, and capacitor switching.

Framework Property 1—Valid Topological Ordering of the Zones: The framework introduces a *valid topological order* among the zones. This order can be leverage to simplify outage location in large-scale networks. A valid topological order for any pair of zones is a relationship denoted as $\Psi_i \succ \Psi_j$, indicating that Ψ_i has a higher topological order than Ψ_j. This means that $Z_{\Psi_i} \not\subset Z_{\Psi_j}$; i.e., either all branches in Ψ_j are located downstream of the branches of Ψ_i or the branches of Ψ_i and Ψ_j do not share any common path starting from the network's root node. Note that $\Psi^g = \{\Psi_1, \ldots, \Psi_w\}$ obtained from the proposed BFS-based zone selection algorithm follows a valid topological order, meaning that $\Psi_1 \succ \cdots \succ \Psi_w$. The reason for this is that the proposed zone selection algorithm explores all the immediate downstream nodes at each depth level without backtracking in Stage II (Section II), prior to moving to the next level.

To show this, note that when an outage event happens the anomaly scores for a subset of zones in Ψ^g, will increase above their normal range. Due to the radial structure of the networks these zones will follow a relationship of the form $Z_{\Psi_1} \supset Z_{\Psi_2} \supset \cdots \supset Z_{\Psi_{v_O}}$, with v_O denoting the number of the zones containing the faulted branch. Thus, the zones within Ψ^g that are impacted by outage also follow a valid topological order. At Stage IV (Section III), the proposed zone coordination algorithm selects $\Psi_{v_0} \leftarrow \Psi_a$ (i.e., the zone with the lowest topological order) as the zone that has the most specific information on the location of outage among all the impacted zones, since it contains the least number of candidate branches. Hence, higher order zones on the same path with abnormal anomaly scores, which are supersets of the selected zone and have less information on outage location, are automatically ignored. This eliminates the need for a burdensome comprehensive search process. Finally, to infer the candidate

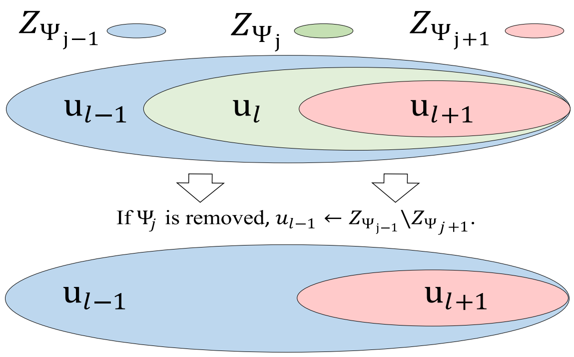

Fig. 6. Venn diagram for demonstrating proof of Theorem 1.

branches that are potentially the location of the outage event, all the branches in the healthy zones with lower topological orders than Ψ_{v_O} have to be removed, as shown in Step IV (Section II). This helps the operator to directly pick the smallest set of branches among thousands of candidate branches in a large-scale network. For example, when outage occurs in any branches within Ψ_6 in Fig. 3, the DSO can ignore the anomaly scores of zones that have a higher topological ordering (i.e., $\Psi_1, \dots, \Psi_5$) to directly infer outage location as $\Psi_a \leftarrow \Psi_6$.

Framework Property 2—Maximum Outage Location Information Extraction: The proposed algorithm is able to obtain the optimal zone set as it (locally) maximizes the amount of information on the location of outage events in partially observable systems. To show this, first, we leverage the concept of entropy to assess the amount of outage location information in Ψ^g. The set $\gamma^g(b_j)$ is defined as $\gamma^g(b_j) = \{\forall \Psi_i : \ b_j \in Z_{\Psi_i}, \Psi_i \in \Psi^g\}$. Hence, $\gamma^g(b_j)$ is the set of all zones in Ψ^g that include b_j. Based on this definition, for each Ψ^g, a set of *undetectable branch sets* is defined as $U(\Psi^g) = \{u_1, \dots, u_V\}$, where $u_k = \{b_{k_1}, \dots, b_{k_n}: \forall b_{k_i}, b_{k_j}, \gamma^g(b_{k_i}) = \gamma^g(b_{k_j})\}$. Thus, u_k defines a set of branches that are covered with the exact same set of zones and cannot be distinguished from each other in terms of outage event location. Given the set $U(\Psi^g)$ the outage location information can be measured using the concept of *entropy*, as follows [29]:

$$H(U(\Psi^g)) = -\sum_{i=1}^{V} \frac{|u_i|}{M} \log \frac{|u_i|}{M} \tag{8}$$

where $|u_i|$ is the cardinality of the set u_i. The higher entropy value implies a higher number of distinguishable branches, and consequently, more information on outage location. The theoretical upper boundary for the entropy is $log(M)$; this case only happens when each u_k only includes a single branch and $V = M$ (i.e., all branches are fully distinguishable and $|u_i| = 1$). This indicates any individual branch is distinguishable using two zones that intersect exactly at that branch. The theoretical lower boundary value for the entropy is zero, which implies that all the branches are covered by identical set of zones (i.e., the branches are not distinguishable and $|u_i| = M$). Based on this metric, the following theorem and proof are obtained.

Theorem 1: For any partially observable network, the proposed BFS-based zone selection algorithm maximizes the outage detection entropy.

Proof: We will prove the local optimality of the selected zone set, Ψ^g, by showing that any deviation from this set results in a decline in outage detection information entropy. Here, a deviation is defined as the addition or removal of any one zone. First, consider the case of removing an arbitrary zone $\Psi_j \in \Psi^g$, and without loss of generality assume that $\Psi_{j-1} \in \Psi^g$ and $\Psi_{j+1} \in \Psi^g$ are the smallest and largest zones, respectively, where $\Psi_{j-1} \supset \Psi_j \supset \Psi_{j+1}$ holds. As is demonstrated in Fig. 6, all the branches that are covered by Ψ_{j-1}, Ψ_j and Ψ_{j+1} are partitioned into three branch sets that belong to the set U : $u_{l-1} = Z_{\Psi_{j-1}} \setminus Z_{\Psi_j}$, $u_l = Z_{\Psi_j} \setminus Z_{\Psi_{j+1}}$, and $u_{l+1} = Z_{\Psi_{j+1}}$. Based on the proposed GAN-based zone coordination algorithm, the status of u_{l-1} can be determined by comparing the anomaly scores of Ψ_{j-1} and Ψ_j. The status of u_l and u_{l+1} are determined by the anomaly scores of Ψ_j and Ψ_{j+1}. Note that all these three sets are distinguishable from each other in outage detection. When Ψ_j is removed, u_l will be eliminated from $U(\Psi^g)$. The new branch partition is reduced to two sets $u_{l-1} \leftarrow Z_{\Psi_{j-1}} \setminus Z_{\Psi_{j+1}}$ and $u_{l+1} = Z_{\Psi_{j+1}}$. This means that the status of u_l cannot be determined anymore (i.e., u_l is merged into u_{l-1}). In other words, Ψ_j is the one zone that enables discrimination between branches u_l and u_{l-1}. Mathematically, this leads to a decrease in entropy, $H(U(\Psi^g))$; the decline in entropy equals $\frac{1}{M} \log \frac{(|u_{l-1}|+|u_l|)^{|u_{l-1}|+|u_l|}}{|u_{l-1}|^{|u_{l-1}|}|u_l|^{|u_l|}}$. This decrease shows that removal of any zone in Ψ^g will reduce the amount of outage location information. Now consider the case of adding a zone to Ψ^g: assume that the newly added zone, Ψ_j, is defined by two observable nodes $S_{o1} \in S_g$ and $S_{o2} \in S_g$; however, the proposed algorithm has already utilized all the observable nodes in S_g as S_{o1}, shown in Step II (Section II); this means that there is at least one zone in Ψ^g that is identical to Ψ_j. Hence, adding a zone to the set Ψ^g will not change $U(\Psi^g)$ and the entropy remains unchanged. ■

Framework Property 3—Robustness Against Bad Data Samples: Bad AMI data samples could generate high anomaly scores, which can lead to misclassification of bad data as outage event. Hence, it is essential to block these data samples from the outage detection algorithm. To do this, we have integrated a bad data detection mechanism into the algorithm by taking advantage of existing redundancy of the zones in Ψ^g. The basic idea is that since bad measurement data are not actually generated by outage events, it is highly unlikely to cause deviations in anomaly scores assigned to several intersecting zones at the same time, given that intersecting zones do not share the data from the same measurement devices. To introduce robustness against bad data, a set of redundant zones is selected for Ψ_a, Stage IV (Section III). This set consists of the zones with lower topological order than Ψ_a, and is denoted as $\Psi^R = \{\Psi_{r_1}, \dots, \Psi_{r_n}\}$, where $\Psi_a \subset \Psi_{r_i}, \ \forall \Psi_{r_i} \in \Psi^R$. If $\exists \Psi_{r_i}$ such that $\zeta_{\Psi_{r_i}} \leq \mu_{\Psi_{r_i}} + h \cdot \sigma_{\Psi_{r_i}}$ then the outage in Ψ_a is dismissed as bad data. The number of redundant zones $|\Psi^R|$ depends on the desired reliability of the algorithm against bad data. If the probability of receiving an anomaly due to bad data for each zone is η, then the probability of misclassifying a case of bad data as outage decreases with $\eta^{|\Psi^R|}$.

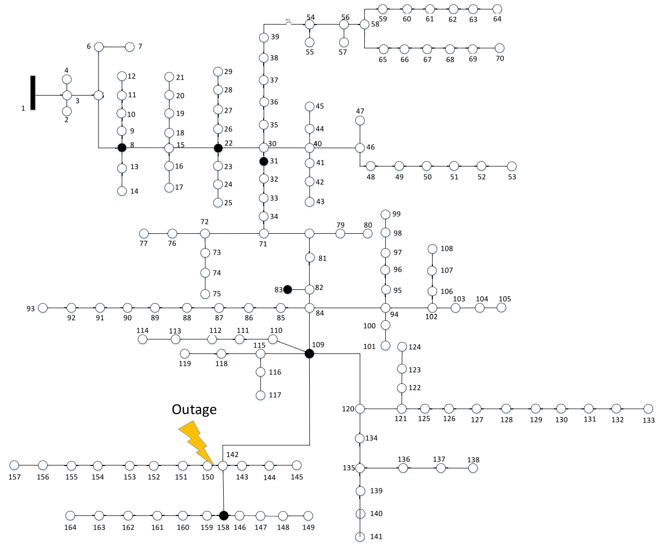

Fig. 7. 164-node feeder topology.

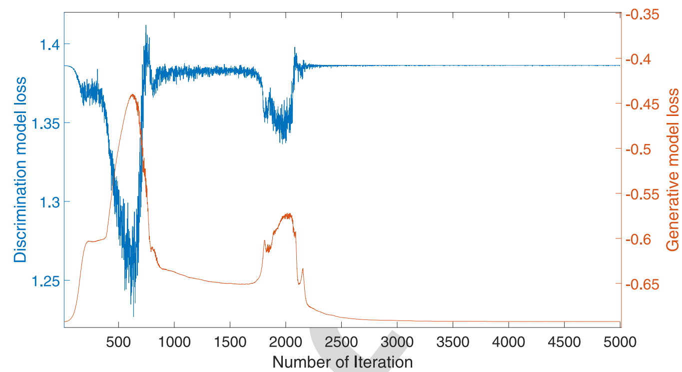

Fig. 8. Training result for a GAN model.

V. Numerical Results

The proposed outage detection method is tested on a real distribution feeder with corresponding 3-year hourly SM data. To provide convincing results, the most complex real distribution feeders is selected from our dataset. The topology of this network is shown in Fig. 7. This feeder consists of 164 nodes and around 800 customers [17]. Six observable nodes are assumed in this feeder (node 8, node 22, node 31, node 83, node 109, and node 158), where five zones are defined based on these nodes. These zones are denoted $\{\Psi_1, \ldots, \Psi_5\}$ and include branches downstream of node 8, node 22, node 31, node 83, and node 109, respectively. Note that $\Psi_1 \succ \Psi_2 \succ \cdots \succ \Psi_5$.

A. Performance of GAN Model

To validate the performance of GAN training process, we calculate the loss values of G and D that can be leveraged to verify if the model has the model has converged to the Nash equilibrium or not. The loss values are calculated based on the objective function of GAN. In the training process, G is trained to maximize $log(D(G(z)))$ and D is trained to maximize the probability of assigning the correct label to both training examples and samples from G, $-log(D(x_{\Psi_i})) - log(1 - D(G(z)))$. According to the theoretical analysis in [13], when the Jensen-Shannon divergence between the G model's distribution and the data distribution is zero, $D(G(z))$ and $D(x_{\Psi_i})$ should converge to $1/2$, which indicates that the loss values of G and D should converge to $2\log(2)$ and $\log(\frac{1}{2})$ at the equilibrium, respectively. This has been confirmed in Fig. 8. After a number of training iterations, both D and G losses converge to the desired values and these indicate that the GAN has been trained successfully and the underlying joint data distribution in normal condition has been learned.

The case study is conducted on a standard PC with an Intel Xeon CPU running at 3.70 GHz and with 32.0 GB of RAM. The average computational time for training each GAN over the available SM dataset is around 840 seconds. It should be noted that multiple GANs can be trained independently and in parallel with each other, which can reduce the adaptation time after system reconfiguration and capacitor switching. Since the training procedure is offline this parallel training method can be conveniently scaled to large distribution systems.

B. Performance of Outage Detection

The performance of the GAN-based outage detection method is tested for different outage cases. The outage event is located between node 142 and node 164, as shown in Fig. 7; three outage events are simulated with three different outage magnitudes to evaluate the performance of the proposed method. The first case is designed as a small-size event where around 20 customers are disconnected (with 40kW aggregate average hourly demand). The second case is designed to represent a middle-size event, where around 50 customers are impacted (with 100kW aggregate average hourly demand). The third case is a large-size event, with around 80 customers (with 150kW aggregate average hourly demand). For each case, GAN models are trained using the historical SM data of the five zones. These three outage cases where simulated in OpenDSS using our real datasets, in which voltage drop was calculated according to simulation outcomes. Meanwhile, to represent standard measurement deviations, error samples were generated from a normal distribution with zero mean and 1% variance and added to the voltage values obtained from the simulator [30]. Fig. 9 presents the histogram of anomaly score for one zone under normal and outage conditions. The mean values of ζ are 1.263 and 1.33 in the normal and outage conditions with variance values 7.7×10^{-5} and 2.7×10^{-4}, respectively. Based on Fig. 9, the difference between anomaly score under normal and outage conditions is large enough to enable DSOs to distinguish these conditions. Meanwhile, Fig. 10 presents the consistency of anomaly score for training and test sets when the system is in normal conditions. However, when the outage event takes place in the zone, the real-time anomaly score reaches considerably higher values.

It is critical to show that an outage event *outside* a zone will not lead to abnormal anomaly scores for that zone. Fig. 7 shows the distribution of anomaly score changes for one zone, when the outages of different magnitudes happen outside the zone. Hence, this figure depicts the histogram of

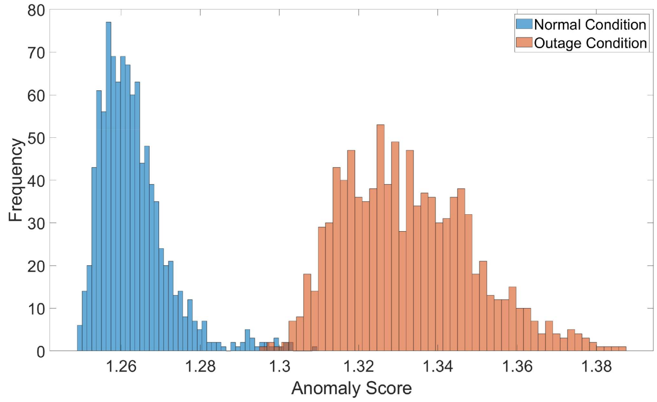

Fig. 9. Anomaly score histogram under the normal and outage conditions.

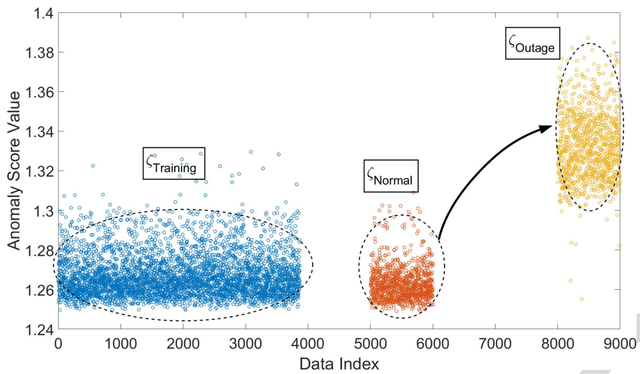

Fig. 10. Anomaly score of the training set, with respect to the normal/outage test set.

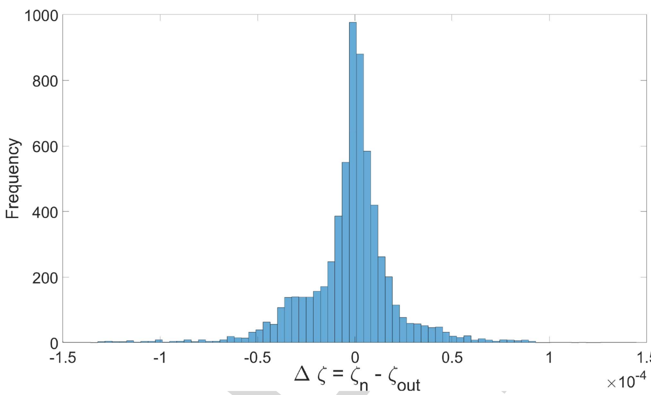

Fig. 11. The histogram of $\Delta\zeta$.

TABLE I
OUTAGE DETECTION QUALITY ANALYSIS

Zone	Case	Accu	Recall	Prec	F_1
Ψ_1	case 1	0.752	0.645	0.8206	0.7223
	case 2	0.913	0.967	0.8727	0.9175
	case 3	0.928	0.9970	0.8761	0.9326
Ψ_2	case 1	0.8355	0.784	0.874	0.8266
	case 2	0.9435	1	0.8985	0.9465
	case 3	0.9435	1	0.8985	0.9465
Ψ_3	case 1	0.673	0.506	0.7685	0.6074
	case 2	0.912	0.984	0.8601	0.9179
	case 3	0.914	0.988	0.8606	0.9199
Ψ_4	case 1	0.9225	0.884	0.964	0.9223
	case 2	0.953	0.939	0.966	0.9523
	case 3	0.981	0.995	0.968	0.9813
Ψ_5	case 1	0.834	0.738	0.9134	0.8164
	case 2	0.9605	0.991	0.934	0.9617
	case 3	0.965	1	0.9346	0.9662

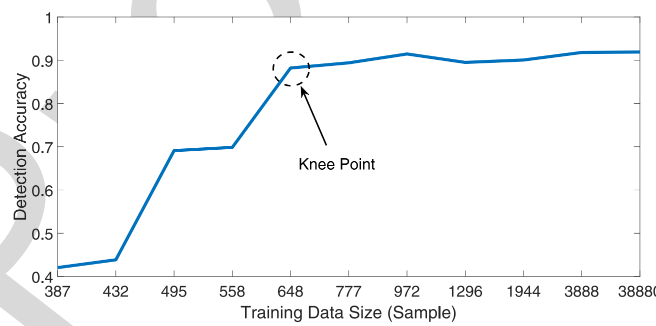

Fig. 12. Sensitivity of outage detection accuracy to the size of training set.

$\Delta\zeta = \zeta_n - \zeta_{out}$, where ζ_n is the anomaly score obtained in normal conditions and ζ_{out} is the anomaly score obtained when the outage happens outside the zone. As can be observed, the anomaly score assigned to the zone does not change and remains almost constant for these outside-zone outages, which indicates that the anomaly score can be relied upon to correctly distinguish the outages inside and outside the zone.

To evaluate the quality of outage detection performance of the proposed method for a multi-zone network, several statistical metrics are applied, such as accuracy (Accu), precision (Prec), recall, and F_1 score [31]. The values of these indexes are presented in Table I for the three outage cases and different zones. Based on the results, we can conclude that the performance of the proposed outage detection method improves as the event size increases, due to higher levels of deviation from normal joint measurement data distribution. For medium and large outage cases, all indexes reach values over 0.9. Moreover, to represent the sensitivity of the outage inference accuracy to the duration of training data, we have tested the average performance of the GAN under various sizes of training dataset as shown in Fig. 12. As is demonstrated in the figure, the performance of the GAN can reach acceptable detection accuracy with a small training set (around 700 data samples, which translates to around 3 days of data).

To prove the performance of our method, we have conducted one more test with more smart meters, and hence finer zones. In this case, 33 observable nodes are assumed in the feeder (node 8, 9, 12, 18, 21, 22, 26, 29, 31, 35, 39, 41, 43, 48, 53, 73, 75, 83, 85, 90, 93, 95, 99, 106, 108, 109, 110, 114, 125, 129, 134, 141, 158), where 19 zones are defined based on these nodes. These zones are denoted as $\{\Psi_1, \ldots, \Psi_{19}\}$ using the proposed zone selection method. The values of the statistical indexes are presented in Table II. Based on this table, most of the statistical indexes are above 0.9, which corroborates good detection performance. When the outage does not occur in the zones, the accuracy of these zones remains stable and high. In general, the proposed method can handle distribution systems

TABLE II
OUTAGE DETECTION QUALITY ANALYSIS FOR 19-ZONE CASE

Zone	Case	Accu	Recall	Prec	F_1	Zone	Case	Accu	Recall	Prec	F_1
	case 1	0.752	0.645	0.8206	0.7223		case 1	0.9225	0.884	0.964	0.9223
Ψ_1	case 2	0.913	0.967	0.8727	0.9175	Ψ_{11}	case 2	0.953	0.939	0.966	0.9523
	case 3	0.928	0.997	0.8761	0.9326		case 3	0.981	0.995	0.968	0.9813
	case 1	0.9495	0.955	0.9446	0.9498		case 1	0.94	0.94	0.94	0.94
Ψ_2	case 2	0.95	0.956	0.944	0.951	Ψ_{12}	case 2	0.94	0.94	0.94	0.94
	case 3	0.951	0.958	0.9447	0.951		case 3	0.9405	0.941	0.9401	0.9405
	case 1	0.922	0.929	0.916	0.923		case 1	0.96	0.96	0.96	0.96
Ψ_3	case 2	0.9225	0.93	0.9163	0.9231	Ψ_{13}	case 2	0.961	0.962	0.9601	0.961
	case 3	0.9175	0.92	0.9154	0.9177		case 3	0.958	0.956	0.9598	0.9579
	case 1	0.8355	0.784	0.874	0.8266		case 1	0.9625	0.962	0.963	0.9625
Ψ_4	case 2	0.9435	1	0.8985	0.9465	Ψ_{14}	case 2	0.962	0.961	0.9629	0.962
	case 3	0.9435	1	0.8985	0.9465		case 3	0.9635	0.964	0.963	0.9635
	case 1	0.9335	0.932	0.9348	0.9334		case 1	0.945	0.946	0.9441	0.9451
Ψ_5	case 2	0.9315	0.928	0.9345	0.931	Ψ_{15}	case 2	0.9455	0.947	0.9442	0.9456
	case 3	0.9365	0.938	0.9352	0.9366		case 3	0.946	0.948	0.9442	0.9461
	case 1	0.973	0.972	0.9739	0.973		case 1	0.834	0.738	0.9134	0.8164
Ψ_6	case 2	0.975	0.977	0.974	0.975	Ψ_{16}	case 2	0.9605	0.991	0.934	0.9617
	case 3	0.976	0.978	0.947	0.976		case 3	0.965	1	0.9346	0.9662
	case 1	0.9455	0.94	0.9505	0.9452		case 1	0.929	0.93	0.9281	0.9291
Ψ_7	case 2	0.945	0.94	0.95	0.945	Ψ_{17}	case 2	0.928	0.928	0.927	0.928
	case 3	0.9465	0.942	0.9506	0.9463		case 3	0.934	0.94	0.9289	0.9344
	case 1	0.902	0.908	0.8981	0.903		case 1	0.976	0.972	0.9798	0.9759
Ψ_8	case 2	0.9055	0.914	0.8987	0.9063	Ψ_{18}	case 2	0.977	0.974	0.979	0.9769
	case 3	0.9065	0.916	0.9	0.9074		case 3	0.9785	0.977	0.98	0.9785
	case 1	0.673	0.506	0.7685	0.6074		case 1	0.9115	0.908	0.9144	0.9112
Ψ_9	case 2	0.912	0.984	0.8601	0.9179	Ψ_{19}	case 2	0.9165	0.918	0.9153	0.9166
	case 3	0.914	0.988	0.8606	0.9199		case 3	0.9195	0.924	0.9158	0.92
	case 1	0.9295	0.929	0.93	0.929		case 1	0.9051	0.881	0.922	0.899
Ψ_{10}	case 2	0.9305	0.931	0.9301	0.9305	Mean	case 2	0.9406	0.952	0.932	0.941
	case 3	0.9296	0.93	0.93	0.9295		case 3	0.944	0.9575	0.931	0.945

with different number of smart meters distributed across the grid.

C. Method Adaption

To validate our fine-tuning strategy, we have conducted additional numerical experiments as shown in Fig. 13. As demonstrated in the figure, a capacitor switching is assumed to have occurred at 12:00 pm. Due to the change in the underlying data distribution, the performance of the proposed method decreases from around 97% to 76%. Here, instead of performing Monte Carlo simulation based on a single set of demand data, we have tested the model with one-month data (under the capacitor switching) and calculated the average accuracy. At the beginning of the fine-tuning process and immediately after the switching event, the model accuracy drops to a low level compared to the previous time-point (around 25%). This is due to the extremely small size of the newly-acquired training dataset and re-calculation of the critical threshold of the anomaly score. Then, the average accuracy of the proposed method clearly improves as the size of training data increases, which allows the model to be fine-tuned reliably. Around a day later, our method achieves similar accuracy levels as before capacitor switching, which means the proposed method has adapted to changes in system conditions. Compared to the results of Fig. 12, the data collection time can be reduced from 3 days to 1 day using our fine-tuning strategy.

D. Method Comparison

We have conducted numerical comparisons with a previous support vector machine-based approach [6] to show that our proposed method can achieve good outage detection accuracy with smaller number of smart meters. To ensure a fair comparison between the two methods, the accuracies of both are evaluated based on the same zone-level criteria. As is demonstrated in Fig. 14, for the three different outage cases, the previous method [6] requires a much higher level of observability (i.e., almost 10 times more) to achieve similar detection accuracy with our method. This indicates that our approach can accurately detect outage events and is a suitable method in most current distribution grids that have limited observability.

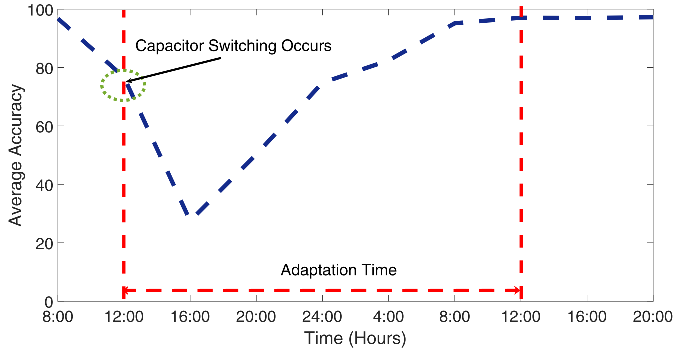

Fig. 13. The performance of the fine-tuning strategy under capacitor switching.

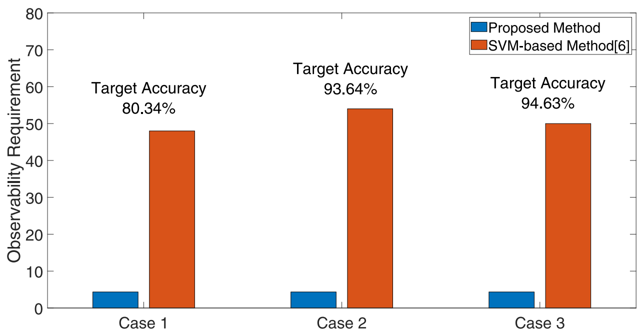

Fig. 14. Comparison results between [6] and the proposed method.

VI. Conclusion

In this paper, we have presented a new data-driven method to detect and locate outage events in partially observable grids using SM measurements. The proposed GAN-based approach is able to implicitly represent the distribution of data in normal conditions and determine potential outage events online. The developed multi-zone outage detection mechanism is based on an unsupervised learning approach, which can address several challenges in outage detection: 1) the poor observability of system caused by the limited number of SMs. 2) data imbalance problem caused by outage data scarcity. 3) the high-dimensionality of the data caused by the temporal-spatial relationship. Meanwhile, our proposed robust BFS-based zone selection and ordering mechanism is guaranteed to capture the maximum amount of information on outage location for any given partially observable system. This method is validated on a real utility feeder using real SM data.

References

[1] *Average Frequency and Duration of Electric Distribution Outages Vary by States*, Energy Information Administration, Washington, DC, USA, Apr. 2018. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=35652

[2] R. A. Sevlian, Y. Zhao, R. Rajagopal, A. Goldsmith, and H. V. Poor, "Outage detection using load and line flow measurements in power distribution systems," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 2053–2069, Mar. 2018.

[3] T. A. Short, *Electric Power Distribution Handbook*. Boca Raton, FL, USA: CRC Press, 2018.

[4] H. Sun, Z. Wang, J. Wang, Z. Huang, N. Carrington, and J. Liao, "Data-driven power outage detection by social sensors," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2516–2524, Sep. 2016.

[5] F. C. L. Trindade, W. Freitas, and J. C. M. Vieira, "Fault location in distribution systems based on smart feeder meters," *IEEE Trans. Power Del.*, vol. 29, no. 1, pp. 251–260, Feb. 2014.

[6] Z. S. Hosseini, M. Mahoor, and A. Khodaei, "AMI-enabled distribution network line outage identification via multi-label SVM," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5470–5472, Sep. 2018.

[7] R. Moghaddass and J. Wang, "A hierarchical framework for smart grid anomaly detection using large-scale smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5820–5830, Nov. 2018.

[8] S.-J. Chen, T.-S. Zhan, C.-H. Huang, J.-L. Chen, and C.-H. Lin, "Nontechnical loss and outage detection using fractional-order self-synchronization error-based fuzzy petri nets in micro-distribution systems," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 411–420, Jan. 2015.

[9] K. Sridharan and N. N. Schulz, "Outage management through AMR systems using an intelligent data filter," *IEEE Trans. Power Del.*, vol. 16, no. 4, pp. 669–675, Oct. 2001.

[10] R. A. Fischer, A. S. Laakonen, and N. N. Schulz, "A general polling algorithm using a wireless AMR system for restoration confirmation," *IEEE Power Eng. Rev.*, vol. 21, no. 4, p. 70, Apr. 2001.

[11] P. Kankanala, S. Das, and A. Pahwa, "Adaboost$^+$: An ensemble learning approach for estimating weather-related outages in distribution systems," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 359–367, Jan. 2014.

[12] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," Univ. Minnesota, Minneapolis, MN, USA, Rep., p. 1, 2007. AQ3

[13] I. J. Goodfellow *et al.*, "Generative adversarial nets," in *Advances in Neural Information Processing Systems (NIPS)*. Red Hook, NY, USA: Curran, 2014.

[14] P. S. S. Thomas, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs, "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," in *Information Processing in Medical Imaging (IPMI)*. Cham, Switzerland: Springer, 2017. AQ4

[15] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Process.*, vol. 35, no. 1, pp. 53–65, Jan. 2018.

[16] S. Beamer, K. Asanovic, and D. Patterson, "Direction-optimizing breadth-first search," *Sci. Program.*, vol. 21, nos. 3–4, pp. 137–148, 2013.

[17] F. Bu, Y. Yuan, Z. Wang, K. Dehghanpour, and A. Kimber, "A time-series distribution test system based on real utility data," in *Proc. North Amer. Power Symp. (NAPS)*, Wichita, KS, USA, 2019, pp. 1–6.

[18] L. Kersting and W. Grigsby, *Distribution System Modeling and Analysis*. Boca Raton, FL, USA: CRC Press, 2016.

[19] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri, "Distributed reactive power feedback control for voltage regulation and loss minimization," *IEEE Trans. Autom. Control*, vol. 60, no. 4, pp. 966–981, Apr. 2015.

[20] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning and statistical estimation in distribution networks—Part I," 2015. [Online]. Available: arXiv:1501.04131.

[21] D. Deka, S. Backhaus, and M. Chertkov, "Learning topology of distribution grids using only terminal node measurements," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Sydney, NSW, Australia, 2016, pp. 205–211.

[22] N. Tajbakhsh *et al.*, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imag.*, vol. 25, no. 5, pp. 1299–1312, May 2016. AQ5

[23] S. Nowozin, B. Cseke, and R. Tomioka, "f-GAN: Training generative neural samplers using variational divergence minimization," in *Advances in Neural Information Processing Systems*. Red Hook, NY, USA: Curran Assoc., Inc., 2016, pp. 271–279.

[24] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, and Y. Bengio, "An empirical evaluation of deep architectures on problems with many factors of variation," in *Proc. 24th Int. Conf. Mach. Learn.*, 2007, pp. 473–480.

[25] K. Dehghanpour, Y. Yuan, Z. Wang, and F. Bu, "A game-theoretic data-driven approach for pseudo-measurement generation in distribution system state estimation," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 5942–5951, Nov. 2019.

[26] E. Denton, S. Chintala, A. Szlam, and R. Fergus, "Deep generative image models using a Laplacian pyramid of adversarial networks," in *Advances in Neural Information Processing Systems (NIPS)*. Red Hook, NY, USA: Curran, 2015.

[27] Y. Chen, Y. Wang, D. Kirschen, and B. Zhang, "Model-free renewable scenario generation using generative adversarial networks," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 3265–3275, May 2018.

[28] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, pp. 281–305, Feb. 2012.

[29] L. Jost, "Entropy and diversity," *OIKOS*, vol. 113, no. 2, pp. 363–375, 2006.

[30] A. Abdel-Majeed and M. Braun, "Low voltage system state estimation using smart meters," in *Proc. 47th Int. Univ. Power Eng. Conf. (UPEC)*, London, U.K., Sep. 2012, pp. 1–6.

[31] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006.

Yuxuan Yuan (Graduate Student Member, IEEE) received the B.S. degree in electrical and computer engineering from Iowa State University, Ames, IA, USA, in 2017, where he is currently pursuing the Ph.D. degree. His research interests include distribution system state estimation, synthetic networks, data analytics, and machine learning.

Kaveh Dehghanpour (Member, IEEE) received the B.Sc. and M.S. degrees from the University of Tehran in electrical and computer engineering, in 2011 and 2013, respectively, and the Ph.D. degree in electrical engineering from Montana State University in 2017. He is currently a Postdoctoral Research Associate with Iowa State University. His research interests include machine learning and data mining for monitoring and control of smart grids, and market-driven management of distributed energy resources.

Fankun Bu (Graduate Student Member, IEEE) received the B.S. and M.S. degrees from North China Electric Power University, Baoding, China, in 2008 and 2013, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. From 2008 to 2010, he worked as a Commissioning Engineer for NARI Technology Company Ltd., Nanjing, China. From 2013 to 2017, he worked as an Electrical Engineer with the State Grid Corporation of China, Nanjing, China. His research interests include distribution system modeling, smart meter data analytics, renewable energy integration, and power system relaying.

Zhaoyu Wang (Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from Shanghai Jiao Tong University in 2009 and 2012, respectively, and the M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology in 2012 and 2015, respectively. He is the Harpole-Pentair Assistant Professor with Iowa State University. His research interests include power distribution systems and microgrids, particularly on their data analytics and optimization. He is the Principal Investigator for a multitude of projects focused on these topics and funded by the National Science Foundation, the Department of Energy, National Laboratories, PSERC, and Iowa Energy Center. He is the Secretary of IEEE Power and Energy Society (PES) Award Subcommittee, the Co-Vice Chair of PES Distribution System Operation and Planning Subcommittee, and the Vice Chair of PES Task Force on Advances in Natural Disaster Mitigation Methods. He is an Editor of the IEEE TRANSACTIONS ON POWER SYSTEMS, the IEEE TRANSACTIONS ON SMART GRID, IEEE PES LETTERS, and the IEEE OPEN ACCESS JOURNAL OF POWER AND ENERGY, and an Associate Editor of *IET Smart Grid*.