IOWA STATE UNIVERSITY

ECE Department

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

Zhaoyu Wang
Northrop Grumman Associate Professor
Iowa State University

Motivation of Data-Driven Power Outage Detection Method

- Based on EIA data, each customer lost power for around **4** hours on average in 2016.
- In August, more than **a million** customers across the Midwest are without power due to a powerful windstorm.
- Use of intelligent communication-capable devices in distribution systems has not become prevalent.
- Conventional expert-experience-based methods that use customer calls are laborious, costly, and time-consuming.

Ames, Iowa, 8/10/2020

Outage Detection in Partially Observable Distribution Systems

- **Problem Statement**: Developing a data-driven method for outage detection using smart meter data in partially observable distribution systems.

Reference	Data Source	Solution	Shortcoming
[1]	Smart meter-based	Multi-label support vector machine	System is fully observable.
[2]		Fuzzy Petri nets-based approach	
[3]		Probabilistic model-based method	
[4]	Non-smart meter-based (i.e., real-time power flow measurement, weather, social network data)	Hypothesis testing-based framework	Limiting data availability
[5]		Social network-based method	
[6]		Boosting algorithm	

[1] 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.
[2] 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.
[3] K. Sridharan and N. N. Schulz, "Outage management through AMR systems using an intelligent data filter," IEEE Trans. Power Deli., vol. 16, no. 4, pp. 669–675, Oct. 2001.
[4] 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.
[5] 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.
[6] 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.

Outage Detection in Partially Observable Distribution Systems

Challenges:

- Most distribution systems are **partially observable** (i.e., not every customer has smart meter).

- Most of the previous works handle the partially observable problem by involving extra data sources, such as real-time power-flow measurements and social network data.

- Outage detection can be considered as a **classification** problem (separating the data samples of normal and outage). However, the size of the outage data is far smaller compared to the data in normal conditions, which leads to a **data imbalanced problem**.

Outage Detection in Partially Observable Distribution Systems

Our Solution [7]:

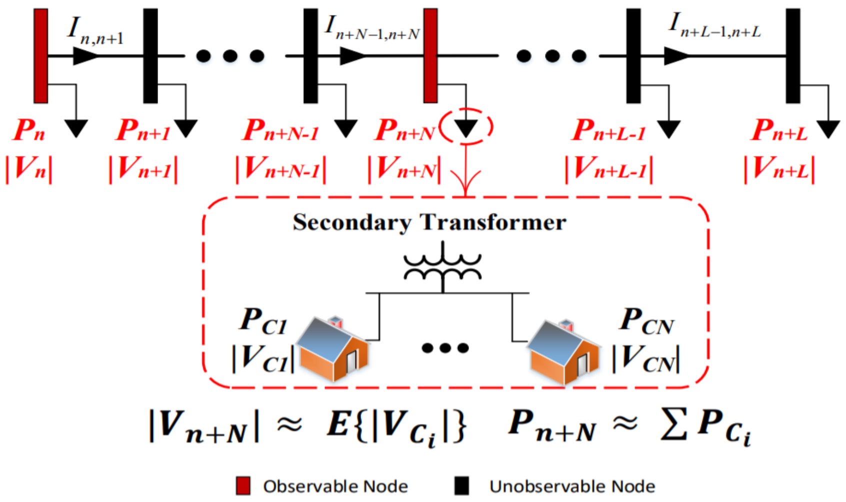
- ✓ Decomposing large-scale distribution networks into a set of intersecting **outage detection zones** and performing zone-based outage detection rather than branch-based outage detection.
- ✓ Granularity of zone-based outage detection depends on the system observability (i.e., when system is fully observable, our method provides branch-based results).
- ✓ Developing an **unsupervised-based** model for outage detection (only utilize the data in normal conditions for model training).
- ✓ Optimizing the zone selection by exploiting the tree-like structure of distribution systems.
- ✓ Providing an anomaly score coordination process to simplify outage location in large-scale networks.

[7] Y. Yuan, K. Dehghanpour, F. Bu, and Z. Wang, "Outage detection in partially observable distribution systems using smart meters and generative adversarial networks, " IEEE Trans. on Smart Grid, accepted for publication.

Outage Detection Zone Definition

Definition: 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}.

- ✓ Give that an outage event anywhere in the zone will lead to deviations from the (voltage-power) data distribution obtained from two observable nodes under normal operations.



Outage Detection Zone Definition

- Give the radial structure of the feeder, the voltage drop between two nodes can be expressed as [2]:

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

- The above equation can be rewritten in terms of nodal power measurements:

$$\Delta\mathbf{V} = |\mathbf{V}_n| - |\mathbf{V}_{n+N}| \approx \sum_{i=n+1}^{n+N} \sum_{j=i}^{n+L} \mathbf{K}_{i-1,i} \otimes \mathbf{I}_{i-1,i} \otimes \frac{\mathbf{P}_j}{cos\phi_j}$$

- When outage happens at a node s downstream of node n, $n+1 \leq s \leq n+L$, the post-outage voltage drop value can be determined as follows:

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

The difference between $\Delta\mathbf{V}$ and $\Delta\mathbf{V}_\mathrm{o}$ are almost proportional to the outage magnitude $\Delta\mathbf{P}_s$.

Step I: Breath-First Search-Based Zone Selection

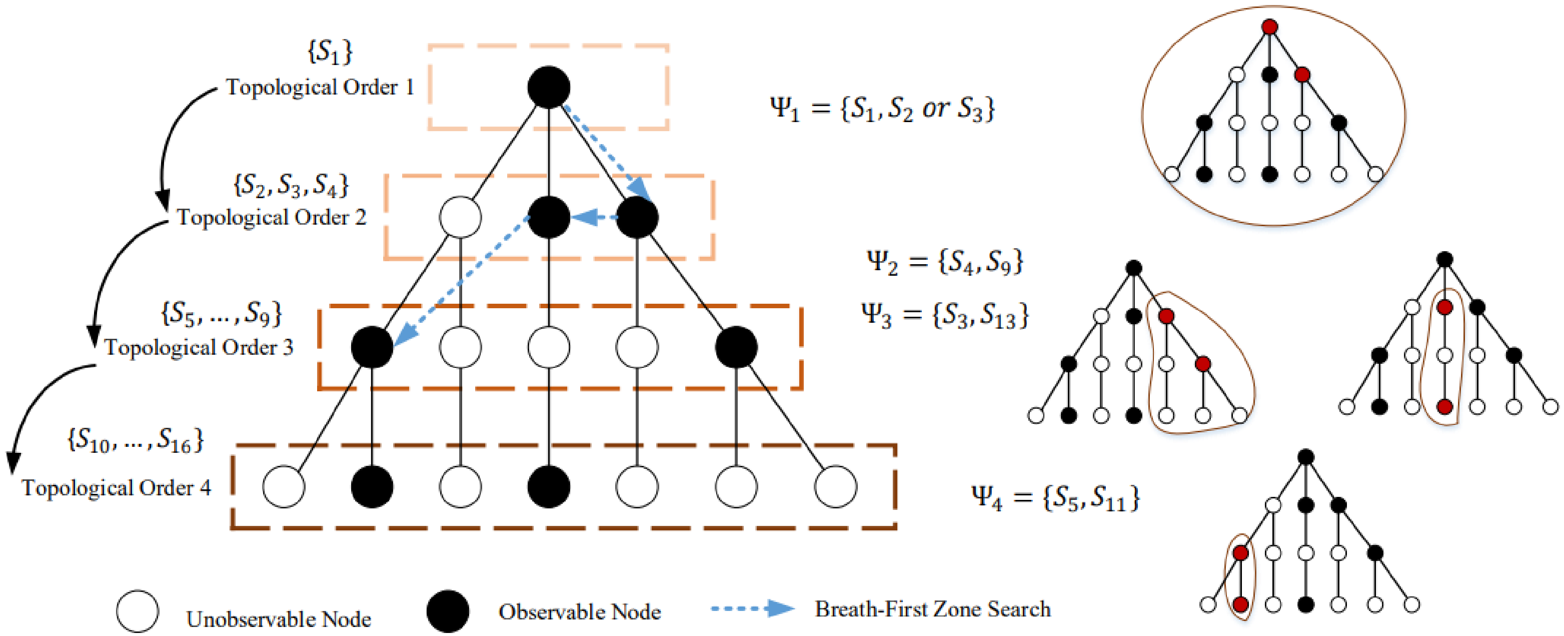
- **Problem**: Sectionalizing networks into multiple zones can be done in more than one way. How to find the optimal set of zones?

- **Our Solution**: Proposing a breadth-First Search-based Mechanism to use **all observable node pairs** to build the zones.

  - Each branch in the system belongs to at least one zone and each zone is unique.

  - Introducing a valid topological ordering, which simplifies outage location identification process.

Step II: Breath First Search-Based Zone Selection



- Each zone is determined by two neighboring observable nodes and contains all branches downstream of these two nodes.
- Selecting the zones using observable nodes at the present depth before moving on the observable nodes at the next topological order.
- The outcome of our zone selection algorithm follows a valid topological order, meaning that $\Psi_1 \succ \cdots \succ \Psi_w$.

Step III: Zone-Based Data Distribution Learning

Challenge: Learning the 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 (i.e., T = 3) for each zone (high-dimensional distribution).

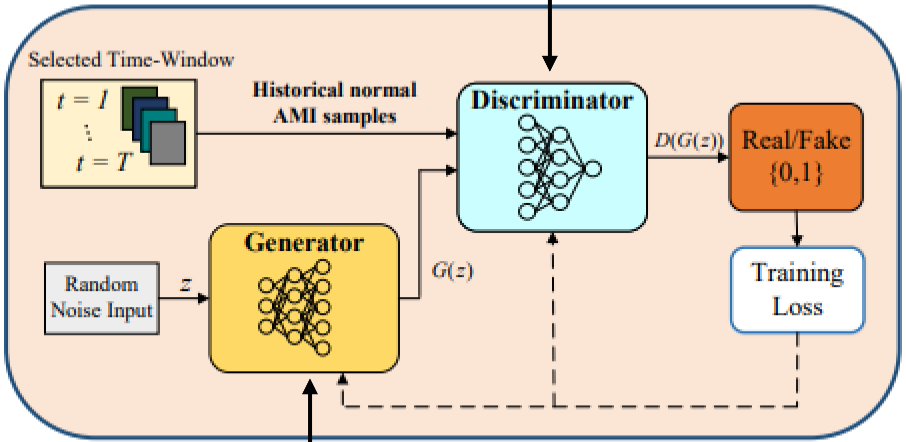
Existing methods:

- Parametric-based methods require distributional assumptions.
- Traditional nonparametric-based methods (i.e., KDE) lack of scalability for large dataset.

Our Solution: Using Generative Adversarial Network (GAN) to implicitly and efficiently represent complex distributions without any distributional assumptions.

- To address data imbalanced problem, we only use the data in normal conditions.

D: Distinguishing the generated data from real data



G: Generating data using random noise

Objective Function:

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

Probability of D assigning the correct label to real samples.

Probability of D assigning the incorrect label to artificial samples from G.

Step III: Zone-Based Data Distribution Learning

Why we use GAN to learn the data distribution?

Advantages:
1. GAN can learn complicated and high-dimensional data distributions without any dimensional assumptions.
2. The performance of GAN is superior (one of the state-of-the-art deep learning algorithms).
3. GAN requires few computation sources during online applications.
4. The discriminator network in GAN provides good guidance for outage detection.

Disadvantages:
1. GAN cannot provide an explicit representation of data distribution.
2. The training of GAN is often difficult (sensitive to hyperparameters).

[8] I. J. Goodfellow, P.-A. Jean, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville, and Y. Bengio, "Generative adversarial nets," NIPS, 2014.

Step IV: Zone-Based Outage Detection

- Zone-based outage detection is achieved by defining a **GAN-based anomaly score** that **quantifies deviations** between the learned normal data distribution and real-time measurements [9].

- The deviation is defined as follows:

$$\zeta_{\Psi_i}(x^t_{new}) = (1-\lambda)\cdot\delta_R(x^t_{new}) + \lambda\cdot\delta_D(x^t_{new})$$

δ_R is the **residual error** that describes the extent to which new measurement follows the learned distribution of the GAN:

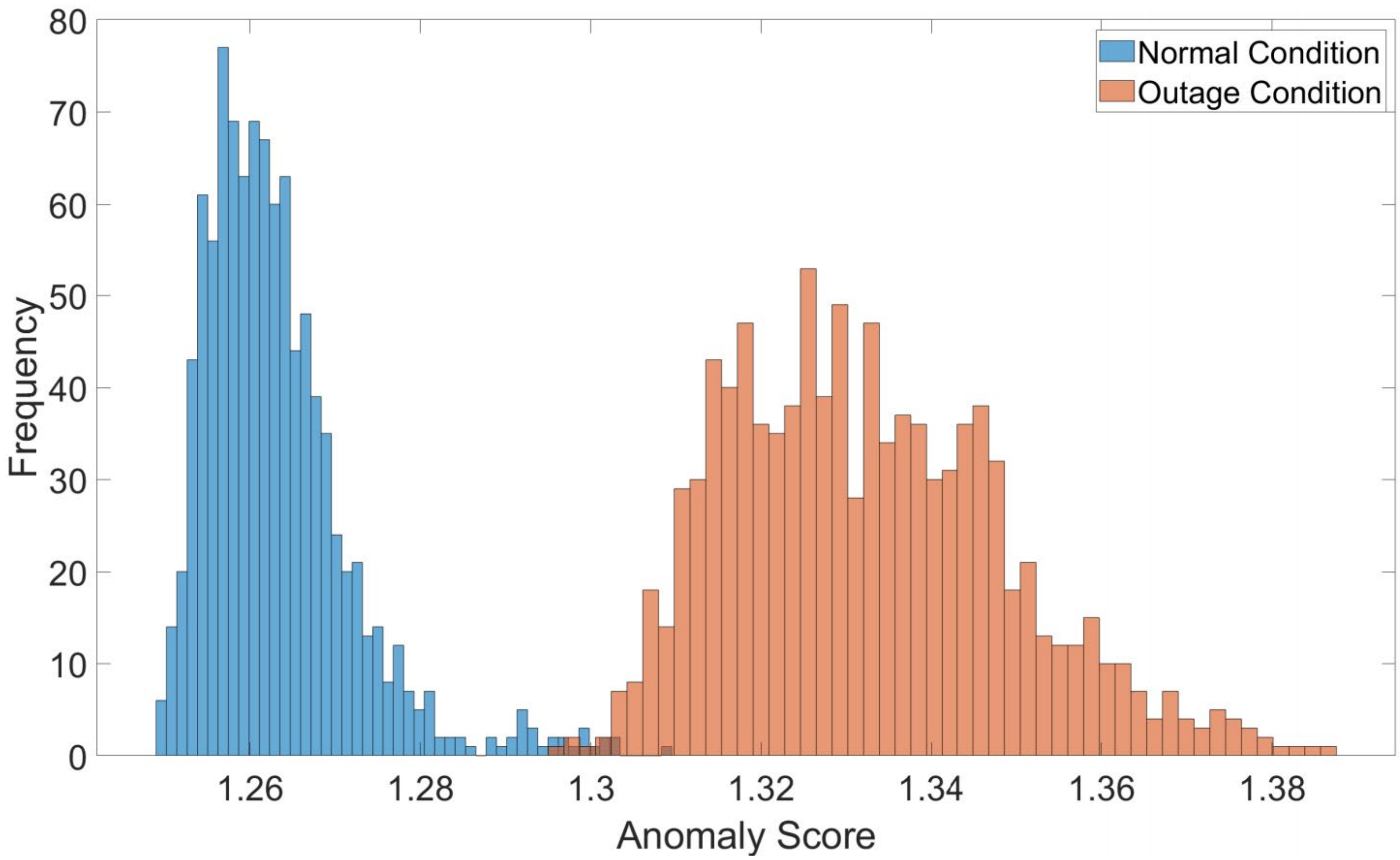
$$\delta_R(x^t_{new}) = \min_z |x^t_{new} - G(z)|$$

δ_D is the **discriminator error** that measures how well the optimal solution of the above optimization (z^*) follows the learned data distribution of the GAN.

$$\delta_D(x^t_{new}) = -\log D(x^t_{new}) - \log(1 - D(G(z^*)))$$

[9] P. S. S. Thomas, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs, "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," IPMI, 2017.

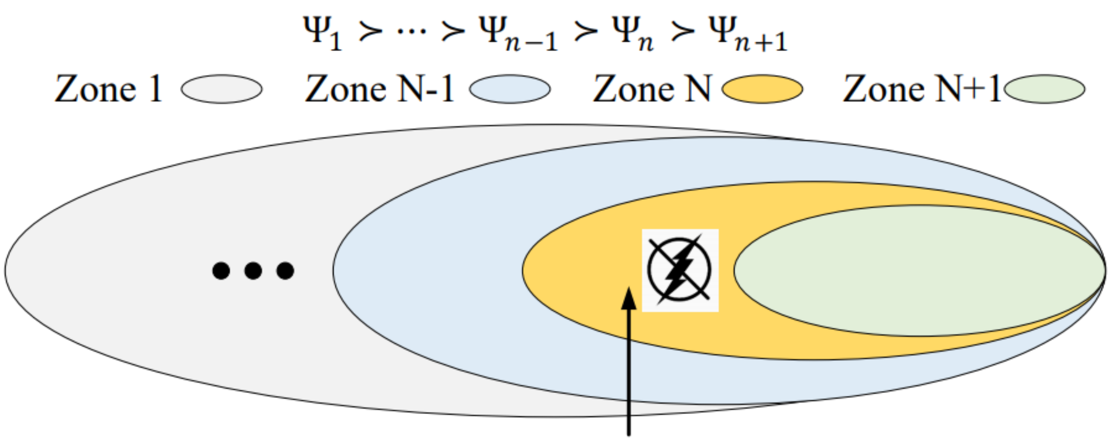
Step IV: GAN-Based Anomaly Score



✓ **A high anomaly score implies outage somewhere in the zone.**

Step V: GAN-Based Zone Coordination

- **Problem**: Multiple zones can contain the faulted branch. How to efficiently select the zone that contains the maximum information on the outage event?
- **Solution**: Using the topological ordering and multiple anomaly scores.
- Zone coordination follows a bottom-up fashion until no outage-related zone exits.



Multiple zones include the outage location (i.e., Zone 1, Zone N-1, etc).
Zone N contains the maximum information on the outage event.
The minimum branch candidates are $Z_{\Psi_N} \backslash Z_{\Psi_{N+1}}$

Theoretical Features of the Proposed Framework

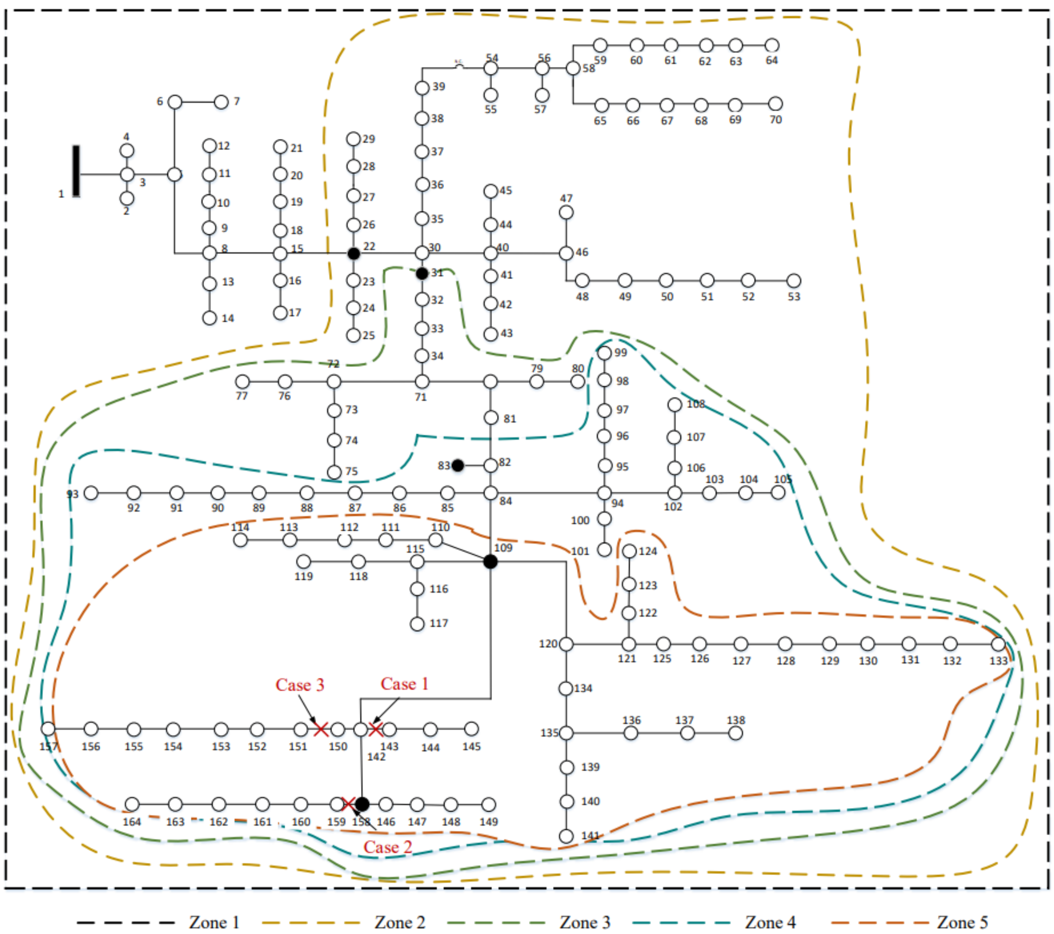
1) Maximum Outage Location Information Extraction

- The proposed algorithm is able to obtain the optimal zone set as it maximizes the amount of information on the location of outage events in partially observable systems.

2) Robustness Against Bad Data Samples

- The proposed algorithm introduces robustness against bad data samples by taking advantage of existing redundancy of the zones (It is highly unlikely to have bad data problem for all zones simultaneously.)

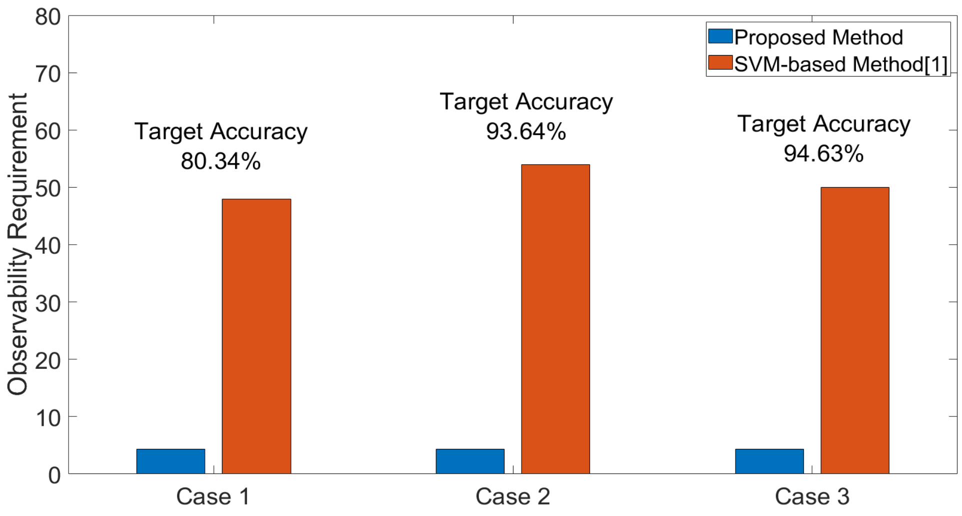
Numerical Results: 164-node Feeder Topology



- Six observable nodes are assumed in this feeder (Node 1, 22, 31, 83, 109, 158).
- Five zones are defined based on these nodes $\Psi_1 \succ \Psi_2 \succ \Psi_3 \succ \Psi_4 \succ \Psi_5$.
- Three outage events are simulated with different outage magnitudes (**case 1**: 20 customers are disconnected; **case 2**: 50 customers are disconnected; **case 3**: 80 customers are disconnected.)

Numerical Results: Accuracy Analysis

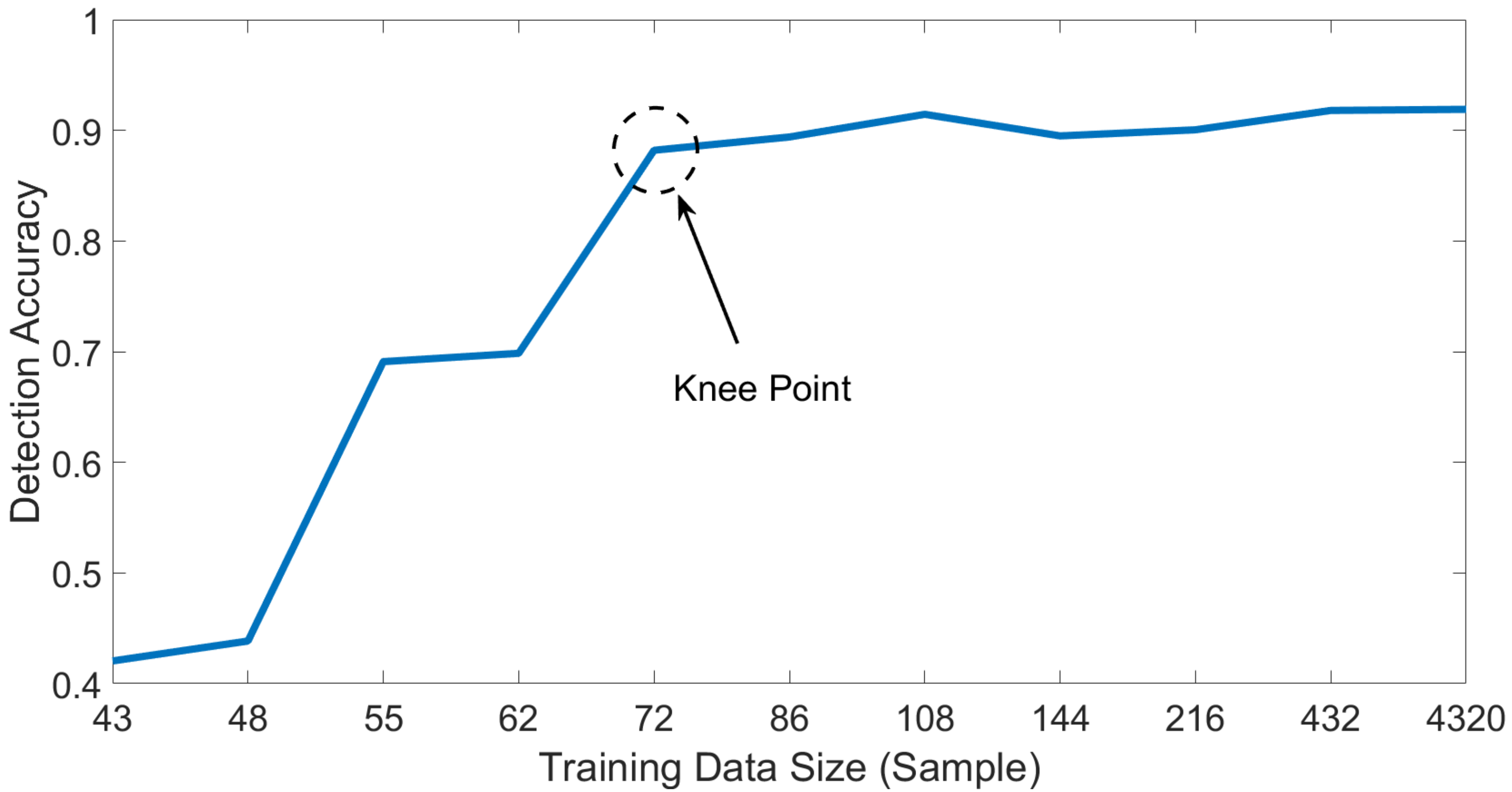
	Outage Detection Accuracy
Case 1	80.34%
Case 2	93.64%
Case 3	94.63%



- ✓ **For three cases, we have tested if our method can detect outages in zone 5. The table shows the results for three cases.**
- ✓ **We have conducted numerical comparisons with a previous method.**
- ✓ **The previous method uses the last gasp signal from smart meters as the input of SVM to identify event location.**
- ✓ **The previous method requires a much higher level of observability (i.e., around 10 times) to achieve similar accuracy with our method.**

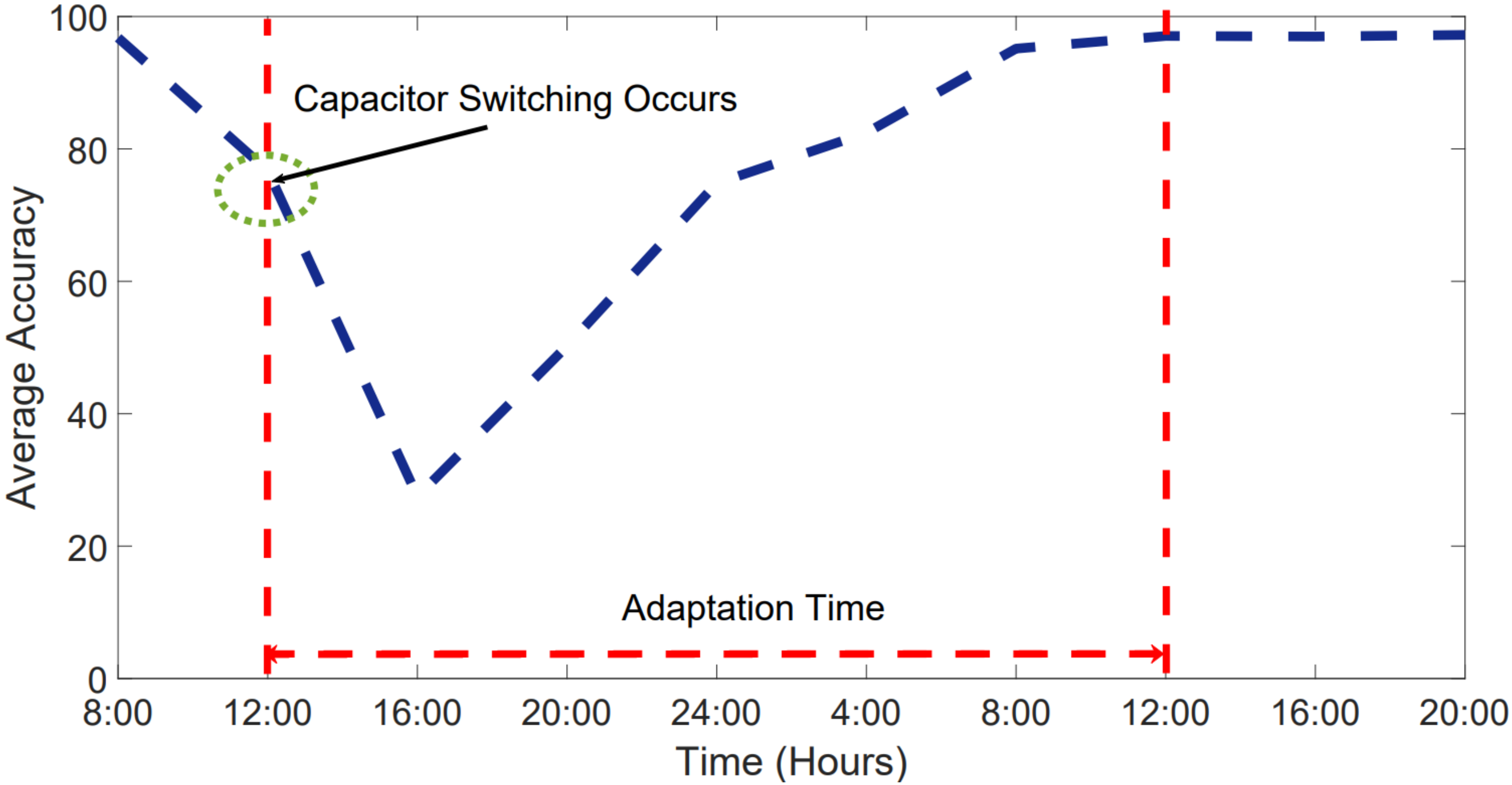
[1] Z. S. Hosseini, M. Mahoor, and A. Khodaei, "AMI-enabled distribution network line outage identification via multi-label SVM," IEEE Trans. Smart Grid, vol. 9, no. 5, pp. 5470–5472, Sep. 2018.

Numerical Results: Sensitivity of Outage Detection Accuracy



✓ **The performance of our model can reach acceptable detection accuracy with a small training set (around 3 days of data, hourly smart meter data).**

Numerical Results: Method Adaption



- ✓ **Our method can adapt to changes in system conditions (i.e., capacitor switching) with a relatively short time (around 1 day).**

Conclusion

- We have presented a new data-driven method to detect and locate outage events in partially observable distribution systems using only smart meter data.
- Our method performs zone-based outage detection rather than branch-based outage detection to handle the poor observability of systems.
- Our method follows an unsupervised learning fashion, thus solving the data imbalanced problem caused by outage data scarcity.
- Out method has been tested using a real distribution feeder and the corresponding smart meter data.

Recent Works Using Smart Meter Data

Distribution System Decision Making:

- Q. Zhang, K. Dehghanpour, Z. Wang, and Q. Huang, "A Learning-based Power Management Method for Networked Microgrids Under Incomplete Information," IEEE Transactions on Smart Grid, vol. 11, no. 2, pp. 1193-1204, March 2020.
- T. Lu, Z. Wang, J. Wang, Q. Ai, and C. Wang, "A Data-Driven Stackelberg Market Strategy for Demand Response-Enabled Distribution Systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2345-2357, May 2019.

Distribution System Situational Awareness:

- 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 Transactions on Smart Grid, vol. 10, no. 6, pp. 5942-5951, November 2019.
- H. Sun, Z. Wang, J.Wang, Z.Huang, N. Carrington, and J. Liao, "Data-Driven Power Outage Detection by Social Sensors," IEEE Transactions on Smart Grid, vol. 7, no. 5, pp. 2516-2524, September 2016.
- Y. Yuan, K. Dehghanpour, F. Bu, and Z. Wang, "Outage Detection in Partially Observable Distribution Systems using Smart Meters and Generative Adversarial Networks, " IEEE Transactions on Smart Grid, accepted for publication.
- Y. Yuan, K. Dehghanpour, F. Bu, and Z. Wang, "A Data-Driven Customer Segmentation Strategy Based on Contribution to System Peak Demand," IEEE Transactions on Power Systems, accepted for publication.

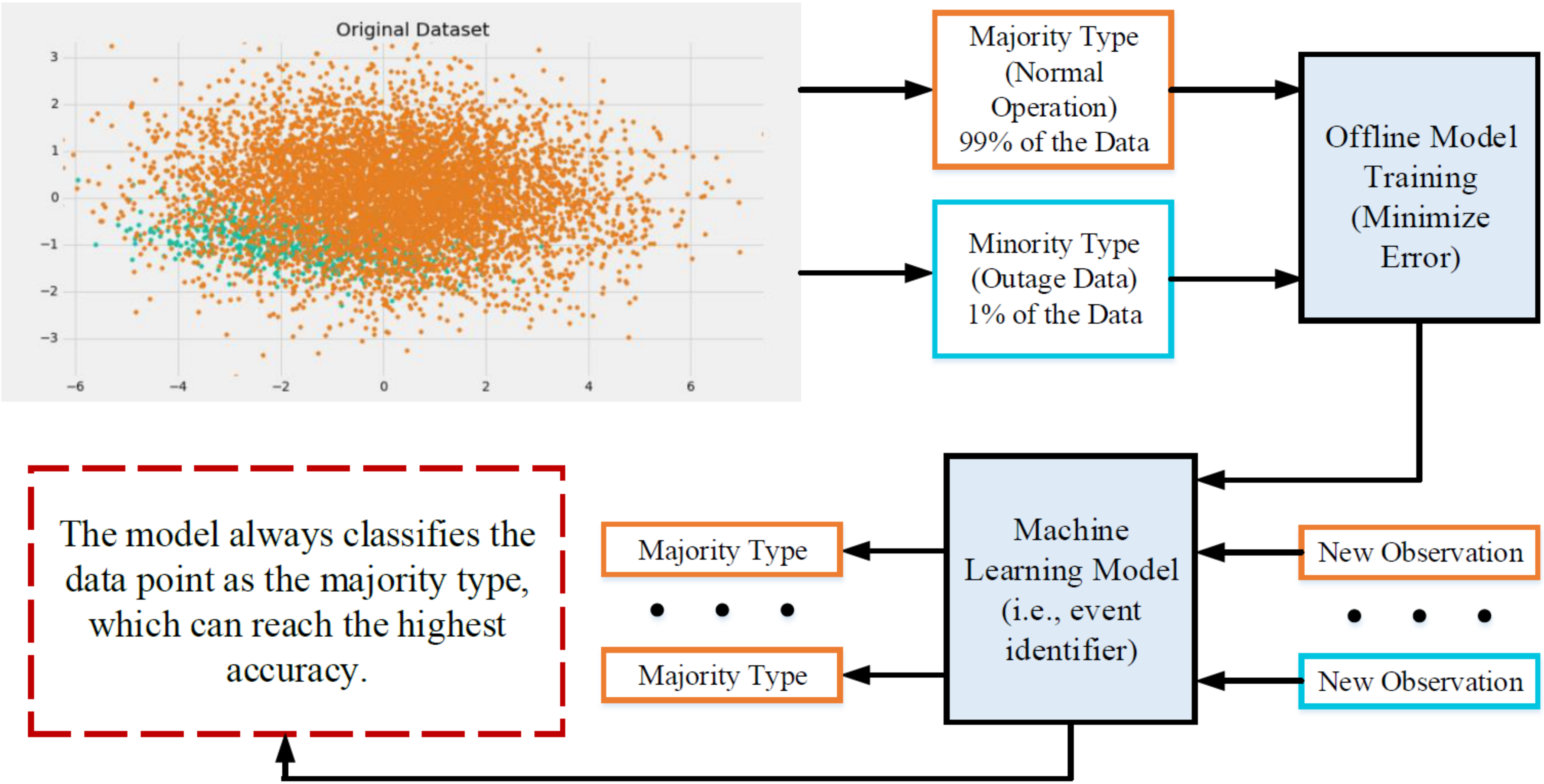
Distribution System Load Modeling:

- F. Bu, K. Dehghanpour, Z. Wang, and Y. Yuan, "A Data-Driven Framework for Assessing Cold Load Pick-up Demand in Service Restoration," IEEE Transactions on Power Systems, vol. 34, no. 6, pp. 4739-4750, November 2019.
- C. Wang, Z. Wang, J. Wang, and D. Zhao, "Robust Time-Varying Parameter Identification for Composite Load Modeling," IEEE Transactions on Smart Grid, vol. 10, no. 1, pp. 967-979, January 2019.
- C. Wang, Z. Wang, J. Wang, and D. Zhao, "SVM-Based Parameter Identification for Composite ZIP and Electronic Load Modeling," IEEE Transactions on Power Systems, vol. 34, no. 1, pp. 182-193, January 2019.

Thank You!
Q & A

Zhaoyu Wang
http://wzy.ece.iastate.edu

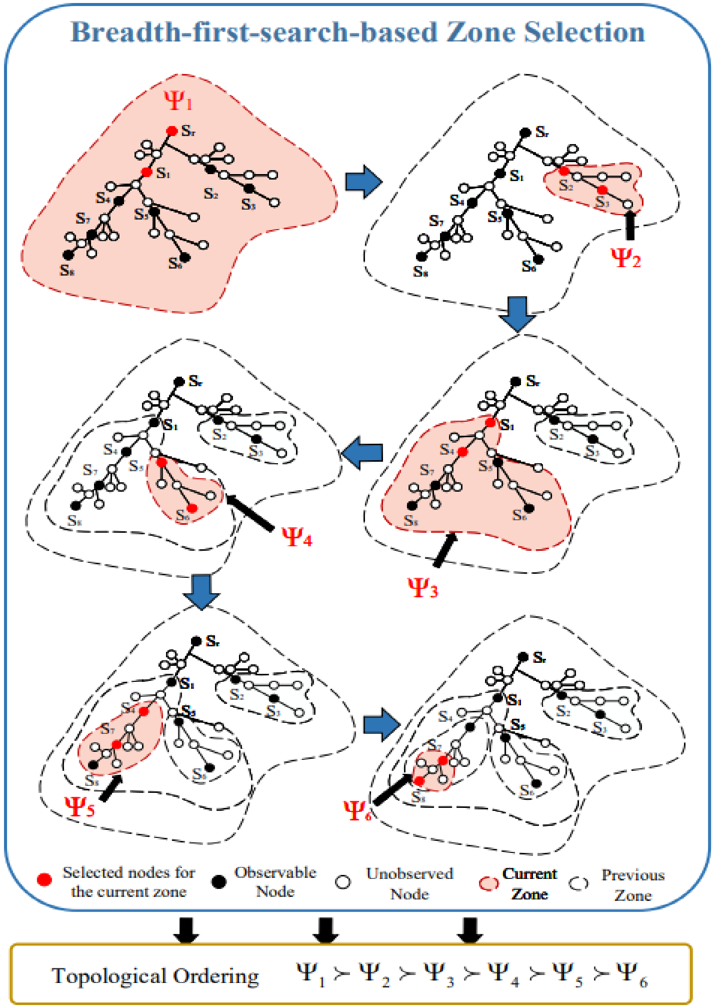
What is Data Imbalanced Problem?



Step II: Breath First Search-Based Zone Selection - Algorithm

- **Step I**: Consider a partially observable distribution system, g, with a total number of M branches, $B_g = \{b_1, \dots, b_M\}$, and a set of O + 1 observable nodes, $S_g = \{S_r, S_1, \dots, S_O\}$, 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 $N(g) = \{\phi\}$. 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. Initialize the set of candidate observable nodes as $S_B = \{S_r\}$, and the zone counter $k \leftarrow 1$.
- **Step III**: If $N(g) = \{\phi\}$, randomly select and then remove a node,S_{o1}, from S_B. Else if $N(g) \neq \{\phi\}$, randomly select and remove a node, S_{o1}, from $N(g)$.
- **Step IV**: Find all the immediate observable nodes downstream of S_{o1}, and randomly select a node from this set, which is denoted as S_{o2}. If $N(g) = \{\phi\}$, 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}. Add Ψ_k to the right side of the current zones in Ψ^g.
- **Step VI**: k ← 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, $\Psi^g = \{\Psi_1, \dots, \Psi_w\}$with w denoting the number of selected zones.

Step II: Breath First Search-Based Zone Selection – An Example



simple

1) In this exemplary system, $B_g = \{b_1, \dots, b_{36}\}$ and $S_g = \{S_r, S_1, \dots, S_8\}$.

2) $k = 1$, Ψ^g and $N(g)$ are both empty. S_r is selected to be the first observable node, $S_B = \{S_r\}$.

3) $S_{o1} \leftarrow S_r, S_B \leftarrow \{\emptyset\}$

4) S_{o2} is selected randomly from the immediate observable downstream node of S_r, $S_{o2} \leftarrow S_1$.

5) $\Psi_1 = \{S_r, S_1, Z_{\Psi_1}\}, \Psi^g = \{\Psi_1\}$

6) $k \leftarrow k + 1$

Step III: Zone-Based Data Distribution Learning

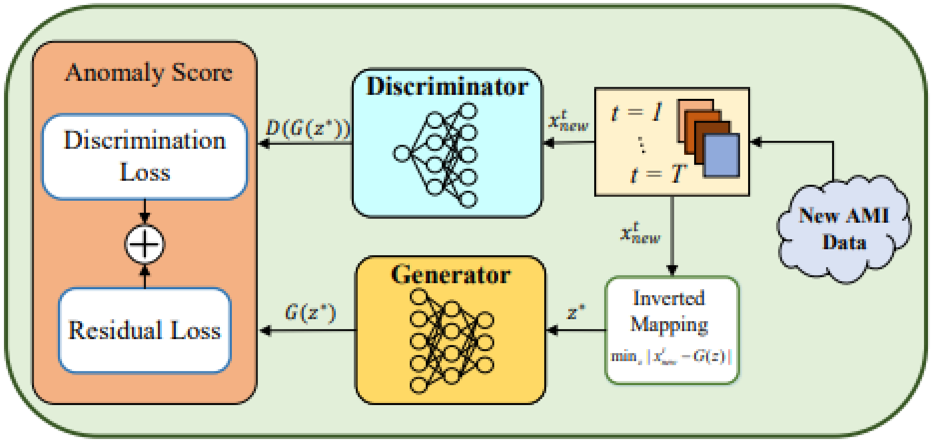
- GAN relies on two interconnected DNNs, which are simultaneously trained via an adversarial process [8]:
  - ✓ **Discriminator D**: maximizing the probability of assigning the correct label to both training examples and generated samples from G.
  - ✓ **Generator G**: generating artificial samples that maximize the probability of the discriminator D mislabeling.
- After training, G can recover the underlying distribution of the training data and the D cannot distinguish the true samples from the artificially generated samples

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 $t = 0, ..., n_D$ **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

Step IV: Zone-Based Outage Detection



GAN-based Anomaly Score [9]:

$$\varsigma_{\Psi_i}(x^t_{new}) = (1-\lambda)\cdot\delta_R(x^t_{new}) \\ +\lambda\cdot\delta_D(x^t_{new})$$

- **Basic Idea**: Detecting the outage by utilizing the learned data distribution in normal operations.
- **Methodology**: Quantifying deviations between the learned distribution and real-time measurements by combining two error metrics: the residual δ_R, and the discriminator error δ_D.
- **Outcome:** GAN-based anomaly score, $\varsigma_{\Psi_i}(\cdot)$, for each zone.

Step V: GAN-Based Zone Coordination - Algorithm

- **Stage I**: Assign a GAN to each zone, $\Psi_i \in \Psi^g$ and use 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.

- **Stage V**: Output the set of candidate branches that are potential locations of outage event as $B_c = Z_{\Psi_a} \setminus \{Z_{\Psi_{a+1}} \cup Z_{\Psi_{a+2}} \cup \cdots \cup Z_{\Psi_w}\}$, where $A \setminus B$ represents the elements of set A that are not in set B.

Step V: GAN-Based Zone Coordination

Valid Topological Ordering of the Zones

- $\Psi_i \succ \Psi_j$: Ψ_i has a higher topological order than Ψ_j, thus $Z_{\Psi_i} \not\subset Z_{\Psi_j}$ (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)
- The outcome of our zone selection algorithm follows a valid topological order, meaning that $\Psi_1 \succ \cdots \succ \Psi_w$.
- This property eliminates the need for a burdensome comprehensive search process in zone coordination process (only need to check one zone).

Theoretical Properties of the Proposed Framework

Property – Maximum Outage Location Information Extraction

- To mathematically prove this property, we first define a set of undetectable branch sets $U(\Psi^g) = \{u_1, \dots, u_V\}$, where $u_k = \left\{b_{k_1}, \dots, b_{k_n}: \forall b_{k_i}, b_{k_j}, \gamma^g\left(b_{k_i}\right) = \gamma^g\left(b_{k_j}\right)\right\}, \gamma^g\left(b_{k_i}\right) = \{\forall \Psi_i: b_{k_i} \in Z_{\Psi_i}, \Psi_i \in \Psi^g\}$.

- Leverage the concept of entropy to quantify the amount of outage location information in Ψ^g [10]:

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

Where, $|u_i|$ is the cardinality of the set u_i and M is the total number of branches.

Theoretical Properties of the Proposed Framework

Property – Maximum Outage Location Information Extraction

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

Proof. We prove the local optimality of the selected zone set, Ψ^g, by showing that any deviation from Ψ^g in a decline in outage detection information entropy.

The case of removing an arbitrary zone $\boldsymbol{\Psi_j \in \Psi^g}$: mathematically, this leads to a decrease in $H(U(\Psi^g))$; the decline equals $\frac{1}{M} log \frac{(|u_{l-1}|+|u_l|)^{|u_{l-1}|+|u_l|}}{|u_{l-1}|^{|u_{l-1}|} \cdot |u_l|^{|u_l|}}$.

The case of adding a zone to $\boldsymbol{\Psi^g}$: since the proposed algorithm has already utilized all the observable nodes to build zone, any additional zone is duplicated. $U(\Psi^g)$ will not change and the entropy remains unchanged.

Theoretical Properties of the Proposed Framework

Property – Robustness Against Bad Data Samples

- Bad AMI data samples could generate high anomaly scores, thus leading to misclassification.
- The proposed method has integrated a bad data mechanism by taking advantage of existing redundancy of the zones in Ψ^g.
- In the zone coordination process, a set of redundant zones Ψ^R is selected that consists of the zones with lower topological order than Ψ_a.
- 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 $\eta^{|\Psi^R|}$.

Numerical Results: Accuracy Analysis

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)}$$

$$Recall = \frac{(TP)}{(TP + FN)}$$

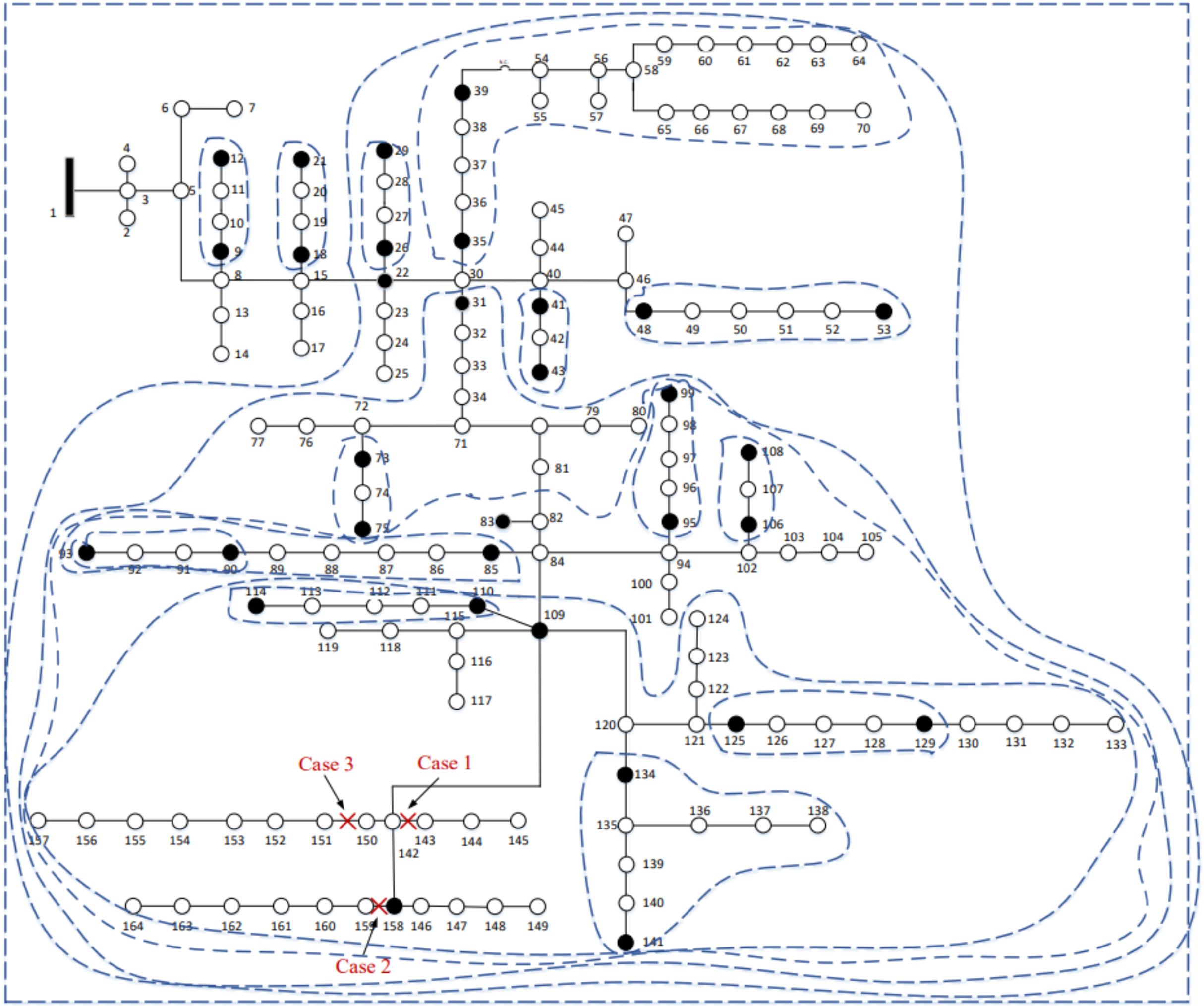
$$Precision = \frac{(TP)}{(TP + FP)}$$

$$F_1 = \frac{(\beta^2 + 1) * Prec * Recall}{(\beta^2 * Prec + Recall)}$$

TP: True positive (correctly predict the outage class)
TN: True negative (correctly predict the normal class)
FP: False positive (incorrectly predict the outage class)
FN: False negative (incorrectly predict the normal class)

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

Numerical Results: Accuracy Analysis for 19-Zone Case



- We have conducted a test case with more observable nodes, and hence finer zones.
- 33 observable nodes are assumed in the feeder (node 1, 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.

Numerical Results: Accuracy 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

Reference

[1] 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.
[2] 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.
[3] K. Sridharan and N. N. Schulz, “Outage management through AMR systems using an intelligent data filter,” IEEE Trans. Power Deli., vol. 16, no. 4, pp. 669–675, Oct. 2001.
[4] 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.
[5] 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.
[6] 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.
[7] Y. Yuan, K. Dehghanpour, F. Bu, and Z. Wang, "Outage detection in partially observable distribution systems using smart meters and generative adversarial networks, " IEEE Trans. on Smart Grid, accepted for publication.
[8] I. J. Goodfellow, P.-A. Jean, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville, and Y. Bengio, “Generative adversarial nets,” NIPS, 2014.
[9] P. S. S. Thomas, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs, “Unsupervised anomaly detection with generative adversarial networks to guide marker discovery,” IPMI, 2017.
[10] L. Jost, “Entropy and diversity,” Oikos, vol. 113, no. 2, pp. 363–375, 2006.