Detect and Identify Topology Change in Power Distribution Systems Using Graph Signal Processing

Osten Anderson
Department of Electrical and Computer Engineering
University of California, Riverside
Riverside, CA

Nanpeng Yu
Department of Electrical and Computer Engineering
University of California, Riverside
Riverside, CA

***Abstract*—The proliferation of smart meters has led to the development of data-driven control algorithms in power distribution systems. However, many of these algorithms rely on accurate topological data. Network reconfiguration can change the topology of the system. It is critical for the system operators to quickly identify the updated topology and take appropriate control actions accordingly. This paper develops an algorithm based on graph signal processing to detect changes in the topology and to identify the new topology after reconfiguration using smart meter voltage magnitude measurements by comparing the smoothness of the signal on the possible topologies. The algorithm is fast, online, and requires only voltage magnitude measurements. Numerical results are presented on 16, 33, and 70-bus test cases. The algorithm achieves good performance for both detecting changes and identifying topology across all test cases.**

***Index Terms*—Topology identification, power distribution system, graph signal processing, smart meter.**

I. Introduction

As the costs of advanced metering infrastructure have dropped in recent years, the penetration of smart meters for commercial and residential customers has increased substantially. The phenomenon has created a surge in data on power distribution systems, enabling the development of advanced data-driven applications [1]. Many such applications, however, require knowledge of the physical status of the network, referred to as the network topology. At the same time, the development of new data-driven applications, such as network reconfiguration for loss mitigation [2], [3], has increased the motivation to make network topology increasingly mutable. Thus, there is a need to develop an algorithm which can provide insight to the topology of the network, when the remote monitoring function for switches fails.

We treat the problem in two parts. The first is topology change detection. A power distribution grid often operates in a default topology. A fault and subsequent closing of a tie switch to restore power forces the network to transition to a different topology. It is key to detect when these changes occur. The second is topology identification. When such a change occurs, it is necessary to determine what is the new topology. To solve the first sub-problem, we propose a graph signal processing approach to quantify the relationship between a signal of nodal voltages and the set of possible topologies on which that signal could have been measured. To solve the second sub-problem, we propose a robust algorithm that uses these relationships to detect topology changes and identify reconfigured topologies. The proposed method serves as a data-driven verification of the physical operating status of the network.

The remainder of this paper is organized as follows. Section II reviews works related to the problem of topology identification. Section III introduces the foundations of graph signal processing and the proposed algorithm for topology change detection and identification. Section IV describes the experimental setup and evaluates the performance of the proposed algorithm. Section V presents the conclusions.

978-1-6654-4875-8/21/$31.00 ©2021 IEEE

II. Related Work

The problem of topology change detection and identification in distribution systems is well-studied, and can be grouped into three main classes. The first class, fault detection, typically seeks to detect when a fault occurs, thus cutting off service to some part of the grid. Many such algorithms leverage knowledge of the physical parameters along with a limited measurements such as from micro phasor measurement units (μ-PMUs) [4]–[6] or power-flow meters [7]. Broadly speaking, these algorithms either rely on the assumption that μ-PMU penetration is at a sufficient level in distribution systems, or assume that additional metering infrastructure could be installed in order to enable the algorithm.

The second class focuses on reconstructing the topology as well as line parameters from measurement data. Algorithms have been proposed which take data-driven approaches to estimating topology and parameters using power injection data [8] or estimating topology using sensors on distributed energy resources [9]. Principle component analysis and graph theory have also been used to estimate the topology of the system from smart meter measurements alone [10]. However, these methods estimate topology on windows of data, with window lengths on the scale of days, or much longer. As topologies become increasingly mutable, it is important to detect the change in a topology as soon as it occurs.

The third class, and the one that the present work most relates, focuses on detecting changes to topology or identifying the current topology in real-time operations. In other words, the algorithms work on single time-step measurements or very short windows, so that a change can be detected almost as soon as it occurs. A kernel-node-map technique is used to identify topology, but not line parameters, from a subset of feasible topologies [11]. A time-series signature verification method is proposed to identify topology changes using µ-PMU data along with knowledge of physical parameters of the system [12]. However, these methods rely on voltage and current phasor measurements at high frequencies from µ-PMUs, which have very limited penetration in distribution grids and thus the assumption that these measurements are available may not be realistic.

This paper overcomes these shortcomings by providing an algorithm which is fast, reliable, and requires no additional sensor infrastructure other than the increasingly-prevalent smart meters. The method also requires relatively little training data. Finally, the algorithm can be used online on single time steps of data as soon as those data points are available.

III. Technical Method

A. Overall Framework

The proposed algorithm is based on graph signal processing (GSP), and in particular, the notion that voltage signals will be smoothest on the graph that represents the physical topology of the power distribution network. The overall framework of the proposed algorithm consists of a training phase and an online testing phase as shown in Fig. 1. In the training phase, the smoothness of the voltage signals on topologies representing normal operations is studied. In the online testing phase, the algorithm consists of two steps. The first step detects topology change in the distribution network. If a topology change is detected, in the second step the topology of the reconfigured network is identified.

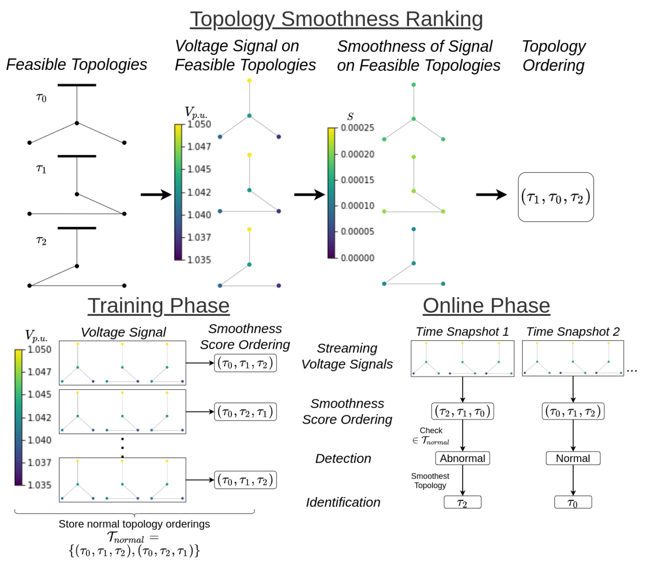

Fig. 1. Overall framework of the proposed topology change detection and identification algorithm for power distribution systems.

B. Review of Graph Signal Processing

A graph is defined by a set of N nodes or vertices $V \in [v_1, ..., v_N]$, and a set of M edges $E \in [e_1, \ldots, e_M]$ between nodes. These connections are modelled by the $N \times N$ adjacency matrix, a symmetric matrix where the ijth element $a_{ij} = 1$ if e_{ij} exists and 0 otherwise. The weight matrix W is an extension of the adjacency matrix, in which the w_{ij} may assume some value relating to the strength of the connection between nodes i and j. The degree matrix D is a diagonal matrix in which the entries d_{ii} are calculated as the sum of all weights of edges connected to node i: $d_{ii} = \sum_{j=1}^{N} w_{ij}$. Finally, the graph Laplacian, the basic building block of graph signal processing, can then be defined as $L = D - W$.

Graph signal processing extends traditional graph theory by defining a signal as a set of N measurements $\mathbf{x} = \{x_1, ..., x_N\}$ on the nodes of the graph. The graph Laplacian may be decomposed by eigenvalue decomposition as $L = U \Lambda U$, where the columns of U are the eigenvectors of L and Λ is a matrix comprised of the eigenvalues on its diagonal. Typically, the eigenvalues and corresponding eigenvectors are sorted in ascending order [13]. The magnitude of the eigenvalues corresponds to the frequency component of the signal, with eigenvectors corresponding to small eigenvalues being part of the low-frequency component.

Using the eigenvalue decomposition of the Laplacian, the graph Fourier transform (GFT) $\mathbf{X}$ of signal $\mathbf{x}$, or the spectral signal, is defined as:

$$\mathbf{X} = U^{-1}\mathbf{x} \tag{1}$$

Similar to the Fourier transform of a classical signal, the GFT takes the signal from the vertex domain to the frequency domain. Arising from this transformation is a notion of 'smoothness', a measure of the overall variation across the graph. A signal is smooth if it varies slowly in the frequency domain, and thus has mostly low-frequency components. A so-called 'smoothness score' may be defined [13] as:

$$s = \mathbf{x}^T L \mathbf{x} \tag{2}$$

This score quantifies the smoothness of a signal on a graph. A smaller score corresponds to a signal $\mathbf{x}$ that is more smooth on the graph with Laplacian L.

C. GSP for Power Distribution System Measurements

Power distribution systems can be readily represented by the graph theoretical framework. The nodes of the graph may be represented by the buses in the system. The edges of the graph are represented by the distribution lines connecting the nodes. We define the signal on the graph as the set of nodal voltage magnitude measurements recorded by smart meters.

1) Weight Matrix Definition: The graph representation of physical lines has a natural choice for the weights of edges: the line admittance. The admittance Y is the reciprocal of the line impedance Z: $Y = \frac{1}{Z}$. The intuition behind this choice is that the voltage signal should be more similar at nodes connected by lines with small impedance values than those connected by lines with large impedances. In essence, the line admittance functions as a proxy for the line voltage drop. However, impedance is complex, consisting of resistance and reactance. The ordering of eigenvalues for complex Laplacians is not well-defined in graph signal processing. To avoid such issues, we use the magnitude of the complex impedance in the weight matrix. The weight between nodes i and j, if they are connected by a line with impedance Z_{ij}, is then defined as $w_{ij} = \frac{1}{|Z_{ij}|}$.

D. Topology Change Detection and Identification Algorithm

A power distribution network consists of a set of normally-closed line switches and normally-open tie switches. The distribution network may have more than one feeder. First, we assume that there is a default operating topology for the distribution system. Under normal conditions, the distribution system operates in this default topology. The network may transition to a different topology by the opening of a normally-closed switch and the closing of a tie line. This transition may occur due to a fault or the operation of a network reconfiguration algorithm. To provide an example, in Fig. 1, τ_0 represents the default topology, and τ_1, τ_2 represent reconfigured topologies. It is also assumed that the operator has access to the set of all possible topologies that the distribution network may operate in. Let n_{tops} be the number of all possible topologies, and the set $T = \{\tau_0, \tau_1, ..., \tau_{n_{tops-1}}\}$ be the set of all possible topologies, where τ_0 denotes the default operating topology. Also, let brackets $\{\}$ denote an unordered set, and parentheses $()$ denote an ordered set.

The proposed algorithm is based on the smoothness score, and in particular, the idea that the voltage signal at any given time step will be smoothest on the graph that corresponds to the physical status of the power distribution network at that time. Thus, the ordering of the smoothness scores can be used to not only detect when the network changes from normal operating status, but also the reconfigured topology.

The algorithm consists of a training phase, and an online detection phase. In the online phase, the algorithm comprises two steps: one step for detecting a change in topology from the default operating topology, and a step for identifying the new topology. The pseudo code of the algorithm is described in Algorithm 1.

1) Training phase: The algorithm first requires a training phase in order to learn the behavior of smoothness scores for all topologies under default operation. Intuitively, it is expected that under normal operation, the smoothest score corresponds to the default topology τ_0. However, the physical parameters of the network as well as fluctuations in voltage signals can result in one or more altered topologies having comparable or better smoothness during certain periods of normal operation. As a result, requiring that τ_0 be the smoothest during normal operation may lead to many normal time steps being mislabeled as reconfigured. Much more stable than the smoothest topologies is the ordering of topologies, ranked by ascending smoothness score. In particular, for normal operation, a few topologies will consistently be amongst the smoothest.

Algorithm 1: Topology Change Detection and Identification

define n_{best};
define $T = \{\tau_0, \tau_1, ..., \tau_{n_{tops}}\}$;
initialize $\mathcal{T}_{normal} = \{\}$;
$normal \in \{True, False\}$
while *training* **do**
  $S \longleftarrow \{s_{\tau_0}, s_{\tau_1}, ..., s_{\tau_{n_{tops}-1}}\}$;
  $Ts \longleftarrow (T, \prec)$;
  $T^{n_{best}} \longleftarrow (Ts_i)_{1 \le i \le n_{best}}$;
  if $T^{n_{best}} \notin \mathcal{T}_{normal}$ **then**
    $\mathcal{T}_{normal} \longleftarrow \{\mathcal{T}_{normal}, T^{n_{best}}\}$;
  end
end
while *online* **do**
  $S \longleftarrow \{s_{\tau_0}, s_{\tau_1}, ..., s_{\tau_{n_{tops}}}\}$;
  $Ts \longleftarrow (T, \prec)$;
  $T^{(n_{best})} \longleftarrow (Ts_i)_{1 \le i \le n_{best}}$;
  if $T^{n_{best}} \in \mathcal{T}_{normal}$ **then**
    normal=True;
  else
    normal=False;
  end
  if *normal* **then**
    $\tau_{current} = \tau_0$
  else
    $\tau_{current} = \arg\min_{\tau_n} s_{\tau_n}$
  end
end

We assume the training phase data is exclusively from periods where the default topology is used. For each time step comprising the most recent nodal voltage snapshot, compute the smoothness score s_τ for each $\tau \in T = \{\tau_0, \tau_1, ...\}$. With a slight abuse of notation, let $Ts = (T, \prec)$ be the set T sorted by ascending s_τ. Then, $T^{n_{best}}$ is the truncation of this ordered set, keeping only the first n_{best} topologies: $T^{n_{best}} = (Ts_i)_{1 \le i \le n_{best}}$. These steps can be visualized by the topology smoothness ranking subsection of Fig. 1. Let $\mathcal{T}_{normal}$ be the set of $T^{n_{best}}$ corresponding to normal operation. $\mathcal{T}_{normal}$ is initialized empty. At each time step, $T^{n_{best}}$ is calculated and added to $\mathcal{T}_{normal}$ if it is not already in the set. $\mathcal{T}_{normal}$ is then the set of unique, ordered $T^{n_{best}}$ which corresponds to normal operating conditions. These steps are shown in the training phase subsection of Fig. 1. This set will then be used during the online phase to determine whether the network is operating in an altered topology.

2) Online Phase: In the online detection phase, $\mathcal{T}_{normal}$ is used to determine if the observed $T^{n_{best}}$ corresponds to normal operation. Streaming data is assumed, meaning data is made available to the algorithm at the same rate as it is collected. For each new voltage signal, the smoothness score for each topology is calculated, and $T^{n_{best}}$ is formed. Then $T^{n_{best}}$ is compared to the $\mathcal{T}_{normal}$ discovered during the training phase. If the current $T^{n_{best}}$ is found, then the current time step has normal topology and the topology at the current step $\tau_{current} = \tau_0$. The two sets must have both the same members and ordering to be considered equal. If the current $T^{n_{best}}$ was not observed during the training phase, then the time step is flagged as having abnormal topology. If abnormal topology is flagged, then the topology on which the signal is the smoothest is labelled as the current topology, $\tau_{current} = \arg\min_{\tau} s_{\tau}$.

The algorithm requires very little parameter tuning. The only true hyperparameter, n_{best}, is not extremely sensitive. The parameter can be tuned by observing the number of sets discovered in the training phase. If a very small number of sets (e.g., less than 3) are found, the parameter is likely too low, and there is a likelihood that in the online phase, some abnormal time steps would have the same $T^{n_{best}}$. Essentially, the sets are too general. Conversely, if there are a very large number of sets (e.g., more than 10), the sets are likely too specific and the chosen n_{best} is too large. In this case, the training phase sets will not generalize well to the online detection phase.

The choice of training set size also functions as a pseudo-hyperparameter. If the training set is too small, the model is more likely to incorrectly flag normal topologies as topology changes. If the training set is too large, the model may label altered topologies as normal. In essence, too small and too large training sets may cause underfitting and overfitting, respectively. There is a notable drop in the true positive rate when the training set size is increased from 100 to 300. Essentially, when the training set is this small, the full set of normal topologies can not be presented in the training set.

This behavior invites the suggestion that, rather than including every set in $\mathcal{T}$, there should be a minimum frequency of $T^{n_{best}}$ to be included. By doing so, it would be expected that the problems too-large training sets would be avoided. In practice, this has several drawbacks. The first is that selecting a value for this threshold is more sensitive than selecting a training set size. The second is that if several unique $T^{n_{best}}$ share the same frequency in the training phase, they will either be all included or all excluded from $\mathcal{T}$. Conversely, increasing or decreasing the set size by some amount may only affect the inclusion or exclusion of a single $T^{n_{best}}$. For these reasons, no such threshold is used.

IV. Experimental Validation

A. Experimental Setup

To validate the performance of the proposed method, we implement the algorithm on three distribution networks. The 16-bus network has 3 feeders and 3 tie switches [14]. The 33-bus network has 1 feeder and 5 tie switches [15]. The 70-bus network has 2 feeders and 11 tie switches [16].

We model load data with Commission for Energy Regulation smart meter data from the Irish Social Science Data Archive [17]. Time-series data is generated by representing each load in the networks by 200 randomly-sampled customers in the data set. The aggregated loads are then scaled such that the average load at each node is equal to the load specified in the original test cases, yielding approximately 8 months of data in 30-minute intervals. Nodal voltage data is obtained by the power flow solutions, where the topology at each time step is defined in the following way.

To generate the chosen topology for each timestep, a window-based approach is taken. Windows of static topology are assigned with lengths uniformly distributed between 12 and 16 time steps, corresponding to 6 to 8 hours of data. In each window, normal topology is assigned with probability 0.7. The probability 0.7 is chosen to provide a reasonable balance between normal and reconfigured topologies to evaluate the algorithm. In the altered topology windows, a normally-closed line is chosen at random to open from the set of all normally-closed lines and a tie switch to close to restore service to every node. In the cases where more than one tie switch could close to restore service for the selected open line, a simple reconfiguration algorithm is implemented by comparing the power flow solution for each topology, and choosing the one that best preserves voltage quality. In these networks, under normal loading, a given open line will generally have a single tie corresponding to the best preserved voltage quality across all time steps. Some topologies have poor power flow convergence characteristics, so the corresponding open-close switch pairs are not allowed. We also do not allow configurations which isolate a feeder completely and shift all loads to another feeder, as this would cause voltage magnitude to dip below acceptable levels. In total, the 16-, 33-, and 70-bus feeders each have 9, 26 and 52 possible reconfigured topologies, not including the default topology.

B. Numerical Results

The performance of the method is evaluated using metrics both for detection and identification. For detection, the true positive rate and true negative rate are presented, where positives represent reconfigured topology and negatives represent normal topology. We present the F_β score with $\beta = 0.5$, a variation of the well-known F_1 score.

$$F_\beta = (1+\beta^2) \cdot \frac{precision \cdot recall}{\beta^2 \cdot precision + recall} \tag{3}$$

Precision and recall in the above equation are calculated as:

$$precision = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$recall = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

In practice, there is a high degree on unbalance between normal operation and reconfigured operation, with a network operating in the normal configuration for the vast majority of the time. In such a scenario, a small increase in the false positive rate would have a disproportionate effect on the ratio between true and false positives, which would quickly erode the ability of an operator to trust results of the algorithm. As such, greater importance is placed on precision than recall. For this reason, we present the F_β score with $\beta = 0.5$.

Each unique topology can be represented by the changes in switches with regard to the normal topology. In other

TABLE I
TOPOLOGY CHANGE DETECTION PERFORMANCE OF GSP-BASED ALGORITHM

Network	True Positive Rate	True Negative Rate	$F_{0.5}$ Score
16 bus	1.00	0.999	0.999
33 bus	0.925	0.997	0.982
70 bus	0.944	0.992	0.982

TABLE II
TOPOLOGY IDENTIFICATION PERFORMANCE OF GSP-BASED ALGORITHM

Network	Open Accuracy	Closed Accuracy	τ Accuracy
16 bus	1.000	1.000	1.000
33 bus	0.875	0.876	0.875
70 bus	0.848	0.940	0.848

words, the reconfigured topology can be described by the opened line and the closed tie. In some cases, the algorithm correctly identifies the tie that closed but not the opened line. For this reason, we present the accuracy for closed switch identification, open switch identification, and entire topology τ identification. Metrics for identification are calculated with respect to only reconfigured time steps. Thus these metrics include false negatives from the previous step but not true negatives nor false positives.

Detection metrics are shown in Table I. The algorithm achieves great performance with regards to detecting topology change across all test cases. Most importantly, high true positive rate is accompanied by a high true negative rate. The result is that there is a high degree of confidence in those events identified as altered topology, and the unbalanced nature of real-world operation would not result in an undue number of false positives. This is reflected in the high $F_{0.5}$ score. Identification metrics are shown in Table II. In all cases, the algorithm successfully identifies the reconfigured topology with high accuracy. In general, the closed switch accuracy is equal or greater than the open switch accuracy. This is due to the nature of distribution systems, where the number of tie switches is generally much fewer than the number of normally-closed switches, making it easier to select the correct tie switch amongst them.

Figures 2 and 3 show the performance of the algorithm with different values of the parameter n_{best} and the training set size pseudo-parameter. The figures show the algorithm is not particularly sensitive to the values of these parameters. The figures also show how choice of parameter involves a tradeoff between true positive and true negative rate, with a higher true positive achievable at the expense of lower true negative and vice versa.

V. CONCLUSION

In this paper, we proposed a novel algorithm for monitoring power distribution system topology. The algorithm exploits the behavior of the smoothness scores of a voltage signal on possible topologies to determine when a change occurs, and identify the new topology. The algorithm relies on very little parameter tuning, and can be used online. The performance of the algorithm is verified using real-world data modeled on 16-bus, 33-bus, and 70-bus test cases. In all cases, the algorithm has excellent performance with regards to both topology change detection and topology identification, with extremely low misdetection of normal topology time steps.

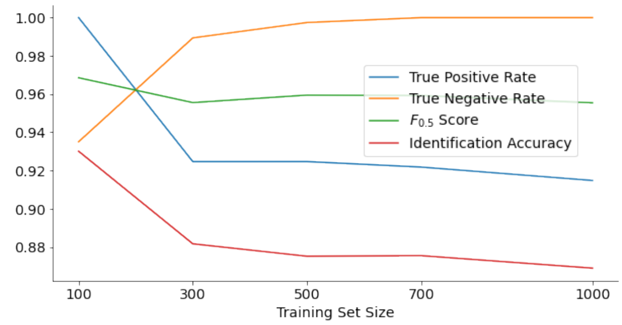

Fig. 2. Detection and identification metrics for training set sizes of 100, 300, 500, 700, and 1000 measurements in the 33-bus system.

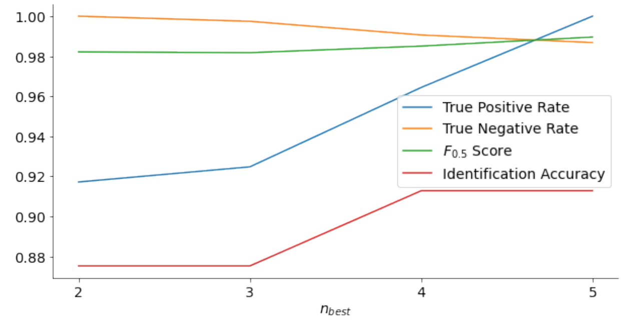

Fig. 3. Detection and identification metrics for n_{best} values of 2, 3, 4, and 5 measurements in the 33-bus system.

REFERENCES

[1] N. Yu, S. Shah, R. Johnson, R. Sherick, M. Hong, and K. Loparo, "Big data analytics in power distribution systems," in *2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pp. 1–5, 2015.

[2] Y. Gao, J. Shi, W. Wang, and N. Yu, "Dynamic distribution network reconfiguration using reinforcement learning," in *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pp. 1–7, IEEE, 2019.

[3] Y. Gao, W. Wang, J. Shi, and N. Yu, "Batch-constrained reinforcement learning for dynamic distribution network reconfiguration," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5357–5369, 2020.

[4] Y. Sharon, A. M. Annaswamy, A. L. Motto, and A. Chakraborty, "Topology identification in distribution network with limited measurements," in *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*, pp. 1–6, 2012.

[5] Hao Zhu and G. B. Giannakis, "Lassoing line outages in the smart power grid," in *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 570–575, 2011.

[6] J. E. Tate and T. J. Overbye, "Line outage detection using phasor angle measurements," *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1644–1652, 2008.

[7] Y. Zhao, R. Sevlian, R. Rajagopal, A. Goldsmith, and H. V. Poor, "Outage detection in power distribution networks with optimally-deployed power flow sensors," in *2013 IEEE Power & Energy Society General Meeting*, pp. 1–5, 2013.

[8] X. Li, H. V. Poor, and A. Scaglione, "Blind topology identification for power systems," in *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 91–96, 2013.
[9] Y. Weng, Y. Liao, and R. Rajagopal, "Distributed energy resources topology identification via graphical modeling," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2682–2694, 2017.
[10] S. J. Pappu, N. Bhatt, R. Pasumarthy, and A. Rajeswaran, "Identifying topology of low voltage distribution networks based on smart meter data," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5113–5122, 2018.
[11] M. Xiao, S. Wang, Z. Ullah, Y. Li, and R. Arghandeh, "Topology detection in power distribution system using kernel-node-map deep networks," *IET Generation Transmission & Distribution*, vol. 14, pp. 4033–4041, 2020.
[12] G. Cavraro and R. Arghandeh, "Power distribution network topology detection with time-series signature verification method," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3500–3509, 2018.
[13] S. Ljubisa and S. Er, *Vertex-Frequency Analysis of Graph Signals*. Springer International Publishing, 2019.
[14] S. Civanlar, J. J. Grainger, H. Yin, and S. S. H. Lee, "Distribution feeder reconfiguration for loss reduction," *IEEE Transactions on Power Delivery*, vol. 3, no. 3, pp. 1217–1223, 1988.
[15] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, 1989.
[16] D. Das, "Reconfiguration of distribution system using fuzzy multi-objective approach," *International Journal of Electrical Power & Energy Systems*, vol. 28, no. 5, pp. 331–338, 2006.
[17] "CER Smart Metering Project - Electricity Customer Behaviour Trial, 2009-2010 [dataset]." www.ucd.ie/issda/CER-electricity, 2012.