IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 5 July 2025, accepted 21 July 2025, date of publication 4 August 2025, date of current version 15 August 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3595431

RESEARCH ARTICLE

Advanced Anomaly Detection in Smart Grids Using Graph Convolutional Networks With Integrated Node and Line Sensor Data

MAHDI ZARIF[1], (Member, IEEE), AND RAMIN MOGHADDASS[2,3], (Member, IEEE)
[1]Department of Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL 33431, USA
[2]Department of Industrial and Systems Engineering, University of Miami, Coral Gables, FL 33146, USA
[3]Miami Institute for Clean Energy, University of Miami, Coral Gables, FL 33146, USA

Corresponding author: Ramin Moghaddass (ramin@miami.edu)

This work was supported in part by the National Science Foundation under Grant 1846975, and in part by the Miami Dade County Office of Resilience through the BE305 Challenge.

ABSTRACT Identifying the location of faults, anomalies, and failures is a long-standing but critical challenge in power system networks. The implementation of smart meters and advanced sensor measurement technologies in recent years has allowed power systems operators to more accurately identify fault locations, quickly resolve problems, and improve the overall reliability of power networks. However, the presence of complex topology changes, penetration of renewable energy resources, stochastic propagation of anomalies in the network, and missing data require the use of new approaches that are sensitive to these issues. This article introduces an innovative method that uses a graph convolutional network (GCN) combined with a modified probability propagation matrix and dual graphs to identify and locate node/line anomalies using network sensors installed on both nodes and lines. Also, an optimization model was developed to find the most likely sources of anomalies across all nodes and edges of the network. The proposed method, which was evaluated on the IEEE 118-bus system and a set of simulated data, demonstrated outstanding performance in handling complex topologies and missing data. Although the proposed model is designed for power networks, its flexible characteristics make it applicable to many sensor-intensive networks or graph structures (e.g., transportation and social networks) where anomaly detection at nodes and/or edges is critical.

INDEX TERMS Anomaly detection, graph convolutional networks, node and line sensors, smart power networks.

I. INTRODUCTION

Power networks have undergone significant transformation in recent years, driven by advances in smart sensing technologies and communication infrastructures. Traditional power systems have evolved into smart grids through the integration of smart sensors, controllers, and real-time data transmission capabilities [1]. Despite these innovations, modern power networks remain vulnerable to a variety of anomalies that can cause temporary disruptions or extended outages. These anomalies generally fall into two categories: cyber-attacks targeting the digital infrastructure and physical anomalies caused by environmental factors or equipment failures, such as lightning strikes, storms, or hardware degradation. Accurate and timely fault identification is crucial to minimizing downtime and maintaining system reliability. However, locating faults in real time remains a major challenge due to the inherent complexity of power systems, the limited availability of high-resolution measurement devices, and the need for fast and low-latency communication [2]. The growing integration of distributed energy resources, particularly renewable sources, further complicates this task. Distributed generators (DG) introduce dynamic and bidirectional power flows that alter the operating state of the network, making anomaly detection and localization more difficult [3], [4].

Smart sensors such as Phasor Measurement Units (PMUs) and smart meters have significantly enhanced the accuracy of anomaly detection. However, the high cost of PMUs requires

The associate editor coordinating the review of this manuscript and approving it for publication was Fazel Mohammadi.

strategic and sparse deployment to balance coverage and economic feasibility [5]. In this context, the rise of deep learning provides promising tools to leverage the increasing volume of sensor data, particularly from affordable, widely deployed devices, while reducing dependence on expensive infrastructure. To fully exploit the potential of data-driven methods, it is crucial to integrate sensor observations with the structural and topological properties of the power grid. Power networks naturally form graph-structured systems, where measurements are distributed across nodes and edges. Therefore, models capable of jointly learning from multivariate sensor data and the underlying network topology are essential. As power systems continue to scale in complexity, the need for intelligent frameworks that combine sensor data with graph-based modeling becomes increasingly urgent for accurate, interpretable, and real-time fault detection.

Graph Convolutional Networks (GCNs) offer a compelling solution to the challenges of fault detection and localization in modern power networks. By naturally aligning with the graph-structured topology of these systems—where buses and substations are represented as nodes and transmission lines as edges—GCNs are well-suited for capturing the spatial and relational dynamics of fault propagation. Leveraging a propagation matrix that encodes how anomalies travel through the network, GCNs enhance the precision and interpretability of anomaly detection compared to traditional machine learning models, which often operate as non-interpretable. This paper proposes a novel GCN-based framework that utilizes both node- and edge-level sensor data to detect and isolate anomalies in power networks. Beyond identifying anomalous components, the framework integrates GCN outputs into an optimization model to refine the localization of faults and assess their broader impact. By combining structural information with data-driven learning, this approach addresses the critical need for accurate, scalable, and interpretable solutions in increasingly complex grid environments. To our knowledge, this is the first framework to simultaneously use edge sensor and node sensor data within a GCN architecture to locate anomalous components and trace the origin of faults in the network.

The remainder of this paper is organized as follows: Section II provides a review and discussion of previous research on network anomaly detection. Section III introduces the proposed anomaly detection model, covering the necessary background information, problem formulation, and identification of anomaly sources. Section IV offers a detailed analysis of numerical experiments and results. Finally, Section V concludes the paper and outlines directions for future work.

II. LITERATURE REVIEW

Several methods exist for anomaly detection and isolation in power system networks, ranging from traditional transmission line modeling techniques to modern machine learning approaches that enhance accuracy (e.g., [6], [7]). Chen et al. [7] provides a comprehensive review of fault detection, classification, and location techniques for transmission and distribution systems. In the following subsections, we review the key literature in both categories and discuss their limitations, which motivate our proposed approach.

A. TRADITIONAL MODELS FOR POWER SYSTEMS ANOMALY DETECTION AND ISOLATION

Traditional fault detection methods in power systems, such as overcurrent protection, distance protection, and differential protection, rely on deterministic models based on fixed thresholds, physical laws, and well-established equations. These techniques are simple, interpretable, and effective under predictable conditions, but struggle with the complexity and variability of modern large-scale networks. Among these, transmission line modeling stands out as a more flexible traditional approach. It uses impedance-based calculations, traveling wave analysis, distributed parameter models, and advanced numerical techniques to detect and localize faults. However, these methods require high-precision, synchronized data to maintain accuracy in dynamic network environments. Takagi et al. [8] introduced a method using only one-terminal voltage and current data, pioneering the use of microprocessor-based fault location by accounting for load flow, fault resistance, and mutual coupling. Signal processing techniques, particularly wavelet transforms, have also shown promise in transient-based fault detection, offering higher precision and robustness to fault resistance and system variations [9]. Other notable advances include Galijasevic and Abur's [10] use of voltage sags and fuzzy logic for low-cost, accurate fault location, and Brahma and Girgis's [11] technique employing synchronized voltage measurements to mitigate current transformer (CT) errors. Liu et al. [12] extended fault location methods to multi-terminal lines using PMUs, significantly reducing error rates. Further improvements were introduced by [13], who refined synchronized sampling techniques by accounting for series losses and solving the Telegrapher's Equations in the time domain. While accurate, this approach requires detailed line parameters, limiting its practicality in complex networks. The integration of PMUs represents a major step forward, enabling real-time, system-wide monitoring. Jiang et al. [14] demonstrated how PMU-based synchronization can overcome limitations of conventional methods. Similarly, Trindade et al. [15] proposed a low-cost approach for distribution networks using smart feeder meters.

In summary, traditional methods for anomaly detection and isolation in power systems, particularly those based on transmission line modeling, exhibit several critical shortcomings. These include their inability to effectively handle the complexity of modern large-scale network topologies, the challenge of acquiring precise and synchronized data across distributed systems, and the difficulty of accurately modeling dynamic system behavior under diverse and unpredictable

fault conditions. Furthermore, achieving a balance between accuracy and computational efficiency becomes increasingly problematic as the size, heterogeneity, and variability of the power system increases. Traditional approaches often lack the adaptability required to address evolving fault scenarios and are heavily reliant on detailed physical models, which can be impractical to develop and maintain. In contrast, machine learning techniques provide a more robust, flexible, and data-driven alternative, capable of learning complex patterns from historical and real-time data, thus reducing dependence on intricate physical modeling and improving adaptability to dynamic power system environments.

B. MACHINE LEARNING MODELS FOR POWER SYSTEMS ANOMALY DETECTION AND ISOLATION

In recent years, machine learning (ML) models—especially neural network-based approaches—have gained significant traction for fault detection and isolation in power systems. These models offer robust alternatives to traditional techniques by addressing challenges related to system complexity, measurement variability, and adaptability. Yadav and Dash [21] provided a comprehensive review of artificial neural networks (ANNs) for transmission line protection, highlighting their adaptability to changing system conditions and fast response times. Similarly, [20] emphasized the effectiveness of neural models in overcoming limitations in measurement collection at substations. Unsupervised learning has also emerged as a cost-effective solution. For example, [22] proposed a hierarchical anomaly detection and multimodal classification approach for photovoltaic (PV) systems, enabling more accurate fault diagnosis without requiring additional hardware. Their model addresses supervisory system limitations in handling complex abnormalities. In the context of smart grids, Chen et al. [23] introduced a GraphSAGE-based method for anomaly detection, which leverages temporal similarities in node features to identify hidden false data. This graph-based learning approach significantly improves key security metrics. A broader evaluation by Gholami and Srivastava [24] surveyed existing ML methods for anomaly detection, classification, and localization in distribution systems. The study identifies critical challenges, including the need for large labeled datasets, non-stationary operating conditions, and the demand for real-time processing. The authors advocate for integrating ML with domain knowledge and advanced data processing techniques to improve system reliability and efficiency. Furthermore, advanced deep learning architectures—such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), generative adversarial networks (GANs), and autoencoders—have enabled more sophisticated applications in time-series prediction, fault diagnosis, and control. These models excel at extracting latent features from complex data, driving significant progress in intelligent power system monitoring and management [25].

Although many machine learning models have been widely employed for anomaly detection and isolation in power systems, they face notable challenges. Their non-interpretable nature leads to limited interpretability and higher-than-expected false positive rates, making them difficult to apply in practice. Additionally, their performance often depends heavily on the network topology, reducing its robustness. To address these limitations, graph-based approaches—such as Graph Neural Networks (GNNs)—offer a more structured and topology-aware solution, better aligned with the complex and interconnected nature of power grids. Liao et al. [26] provide a concise overview of GNN applications in power systems, including fault analysis, time series prediction, and power flow estimation. Takiddin et al. [27] employ Chebyshev-based GNNs to identify system status and localize risks, showing improved detection rates over traditional methods. Vincent et al. [28] demonstrate the effectiveness of GCNs in detecting False Data Injection (FDI) attacks, significantly enhancing smart grid cybersecurity. Chen et al. [29] propose a GCN framework for fault localization in distribution networks, outperforming conventional ML methods in accuracy, noise resistance, and robustness on the IEEE 123-bus benchmark. An extension by Chen et al. [23] improves the GraphSAGE model by incorporating temporal correlations, further enhancing anomaly detection performance. Recent research also explores advanced GNN-based models for anomaly detection in attributed networks. These include ANOGAT-Sparse-TL for class imbalance handling [30], dual variational autoencoders with GANs for sparsity and nonlinearity [31], residual-based GCNs with hypersphere mapping for better anomaly separation [32], and sparse canonical correlation analysis for high-dimensional data alignment [33]. While promising, these methods are largely developed for social networks and do not yet incorporate physical constraints or the operational characteristics of power systems.

C. LITERATURE SUMMARY AND CONTRIBUTIONS OF THE PAPER

While numerous studies have addressed fault and anomaly detection in power systems, most fall short in three key areas: handling complex and dynamic network topologies, integrating both node and line sensor data, and providing flexible models that do not rely on predefined anomaly distributions. Existing methods often assume specific fault distributions and focus solely on node- or line-level anomalies, limiting their applicability in real-world scenarios. In contrast, our approach introduces a Graph Convolutional Network (GCN)-based framework that jointly utilizes bus and line sensor data, encoded as node and edge features via a dual-graph representation. We construct the network adjacency matrix using a transformation matrix, which is then updated using a learned fault propagation matrix. This enables accurate fault localization at both node and line levels, even under changing topologies or incomplete data. A key contribution of our work is the modeling of anomaly propagation as a stochastic process rather than a deterministic one, which better reflects real-world conditions in power systems. Factors such as fluctuating loads, varying line impedances,

and intermittent renewable generation cause anomalies to propagate in uncertain and non-uniform ways. Our model captures these dynamics by learning probabilistic influence patterns throughout the network. Additionally, we address a critical operational challenge: guiding system operators in dispatching maintenance teams across large, renewable-integrated grids. By improving fault localization accuracy and system interpretability, our method can enhance the resilience and reliability of modern power systems. We selected GCN over other graph-based learning models (e.g., GAT, GraphSAGE) due to its computational efficiency, simplicity, and scalability for fixed-topology networks, while still capturing attention-like behavior through a learned propagation matrix that models node influence.

III. THE PROPOSED FRAMEWORK

A. PRELIMINARIES

In this section, we introduce the foundational elements of power system equations and graph neural networks.

1) POWER SYSTEM MODELING

First, we review the mathematical models that govern electrical power systems, including the equations for electricity flow within the network. Power flow analysis examines power networks, focusing on currents, voltages, and power distribution at each bus in the system. The main objective is to identify voltage levels, power flows, and associated losses in the network during steady-state operations. Although elements of the power network, such as lines and transformers, have constant parameters, suggesting a linear system, the power flow problem is inherently nonlinear. This nonlinearity arises from the complex relationships between voltage, current, and power at each bus. Power flow calculations, which involve solving nonlinear equations, are essential to understand how an electrical transmission system responds to specific loads and generator output. These calculations are critical for power system operations and planning. For a network with n independent buses, we typically formulate a system of equations. This system defines the relationships between voltages and currents at each bus using the bus admittance matrix, as shown in Equation (1):

$$\begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1n} \\ Y_{21} & Y_{22} & \cdots & Y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{nn} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} = \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix}, \tag{1}$$

or, more concisely shows in Equation (2),

$$\mathbf{Y} \times \mathbf{V} = \mathbf{I}, \tag{2}$$

where $\mathbf{I}$ denotes the vector of current injections and $\mathbf{V}$ represents the vector of voltages at each bus. The term $\mathbf{Y}$, known as the bus admittance matrix, plays a crucial role. The diagonal entries, Y_{ii}, represent the self-admittance at bus i. This is calculated as the sum of all admittances for branches connected to bus i. Additionally, the bus current can be expressed in terms of bus voltage and power as follows:

$$I_i = \frac{S_i}{V_i} = \frac{S_{G_i} - S_{D_i}}{V_i} = \frac{(P_{G_i} - P_{D_i}) - j(Q_{G_i} - Q_{D_i})}{V_i}, \tag{3}$$

Here, S represents the complex power injection vector. P_{G_i} and Q_{G_i} are the real and reactive power outputs of the generator at bus i, while P_{D_i} and Q_{D_i} represent the real and reactive power loads at bus i. Equation (3) simplifies the relationship between voltage, power injections, and current at each bus. By integrating this into the bus admittance equation, we obtain the following:

$$(P_{G_i} - P_{D_i}) - j(Q_{G_i} - Q_{D_i})V_i = \sum_{j=1}^{n} Y_{ij}V_j, \quad i = 1, \ldots, n. \tag{4}$$

In power flow analysis, when load demands are known, we define the following bus power injections:

$$P_i = P_{G_i} - P_{D_i} \tag{5}$$

$$Q_i = Q_{G_i} - Q_{D_i} \tag{6}$$

Substituting these values into the power flow equation gives the general form:

$$P_i - jQ_iV_i = \sum_{j=1}^{n} Y_{ij}V_j, \quad i = 1, 2, \ldots, n. \tag{7}$$

Equations (4)- (7) highlight the core components of the network's power flow equations. By decomposing them into real and imaginary parts, we can formulate two equations per bus, involving four variables: real power P, reactive power Q, voltage V, and angle θ. To solve the power flow equations, two of these variables must be known for each bus. Several constraints must be taken into account when solving power flow equations, such as generator limitations, voltage limits, and transmission line capacities. These constraints introduce non-linearity into the equations, making them more difficult to solve.

In this paper, we use the Pandapower library in Python to test the system under various fault conditions. Its robust capabilities help manage the complexities of non-linear power system optimization. This enables us to calculate the optimal power flow and obtain comprehensive line data. This includes active and reactive power flowing through the lines, as well as detailed bus data, such as voltage angle, amplitude, and active and reactive power injection at each bus. In anomaly detection frameworks, it is essential to thoroughly understand the normal operating behavior of the power system. Anomalies often appear as unexpected deviations in system parameters. These deviations can be caused by faults, equipment failures, cyber-attacks, or other disruptions. These deviations can lead to irregularities in voltage levels, current flows, and power injections at various buses, potentially compromising system stability. Power flow

analysis plays a critical role in establishing baseline conditions for voltages, currents, and power flows during normal operations.

2) GRAPH CONVOLUTIONAL NETWORKS (GCNs)

Graph theory is a powerful tool for detecting anomalies in power systems. This is because power systems naturally exhibit graph-like structures. In this representation, buses in the power system are modeled as nodes, while power lines are represented as edges connecting the nodes. This framework allows us to apply graph-theoretical methods to identify and analyze anomalies in the network. To handle the complexities of these structures, we use Graph Convolutional Networks (GCNs), which are specifically designed for graph-structured data. Traditional neural networks face limitations when dealing with data characterized by complex relationships, such as those found in social, biological, and communication networks. Their reliance on fixed-size input structures inhibits the effective processing of relational information embedded within graph topologies. GCNs overcome this limitation by performing a specialized convolution operation on the graph. Each node's features are updated based on its own characteristics and those of its neighboring nodes. This enables the GCN to capture and analyze the underlying topology of the power network, making it effective for anomaly detection. The convolution operation in a GCN is represented mathematically as follows:

$$X^{(l+1)} = \sigma(BX^{(l)}W^{(l)}). \tag{8}$$

The variable $X^{(l)}$ represents the feature matrix for all nodes at layer l, and $X^{(0)}$ is the input feature matrix. The weight matrix $W^{(l)}$ contains the trainable parameters for layer l. The symbol σ represents a non-linear activation function, such as ReLU (Rectified Linear Unit). Equation (8) shows the fundamental operation of a GCN. The features of each node are updated using both its own features and those of its neighbors. The normalization of the adjacency matrix, as shown below, is critical for handling varying node degrees and maintaining numerical stability:

$$B = D^{-\frac{1}{2}}AD^{-\frac{1}{2}}. \tag{9}$$

The adjacency matrix A represents the graph, where A_{ij} indicates the presence of an edge from node i to node j. Typically, A_{ii} is zero unless self-loops are included. The degree matrix D is diagonal, with each element D_{ii} representing the sum of the weights of all edges connected to node i, known as the degree of node i. Equation (9) shows how the adjacency matrix is normalized to prevent nodes with high degrees from disproportionately influencing the feature aggregation process.

As shown in Equation (8), the propagation rule is the mechanism by which the GCN updates the features of each node in each layer. The features are transformed by multiplying with the weight matrix and then processed through a non-linear activation function. As a result, the GCN leverages both feature information and the graph structure. This allows it to learn comprehensive representations of nodes based on their attributes and interconnections within the graph. GCNs are versatile and extend beyond anomaly detection. They are suitable for tasks like node binary classification, where the goal is to assign each node to one of two categories. In node binary classification, the GCN learns to distinguish between two classes by analyzing node features and the relationships between nodes in the graph. GCNs can also be used for tasks like multi-class classification, link prediction, and clustering. They take advantage of the graph's structure to enhance predictive accuracy and provide valuable insights. This capability makes GCNs particularly useful for complex systems, like power systems, where the interconnections between buses significantly affect the system's overall behavior. By using GCNs for these tasks, the model leverages the graph topology, ensuring that classifications and predictions are informed by the network' connections.

The output layer for node binary classification, using a sigmoid activation function, is mathematically represented as follows:

$$\hat{\mathbf{y}} = [\hat{y}_1, \ldots, \hat{y}_{|N|}] = \sigma\left(X^{(L)}W^{(L)} + b^{(L)}\right), \tag{10}$$

Here, $\hat{y}$ is the $|N| \times 1$ vector containing the predicted probabilities for each of the n nodes. $X^{(L)}$ is the $|N| \times d_L$ matrix of node features from the final layer L, where d_L is the number of features in layer L. $W^{(L)}$ is the $d_L \times 1$ weight matrix, which, when multiplied by $X^{(L)}$, results in a $|N| \times 1$ vector. $b^{(L)}$ is the $|N| \times 1$ bias vector, typically broadcast to match the number of nodes n. This ensures that the output $\hat{Y}$ is a $|N| \times 1$ vector, where each entry $\hat{y}_i$ represents the predicted probability that node i is anomalous.

B. PROPOSED METHOD

1) PROBLEM SETUP

Power networks are naturally represented as attributed graphs. In this representation, nodes correspond to buses, switches, or substations, while edges represent the transmission lines connecting them. This graph-based modeling allows us to apply GCNs, which are specifically designed to handle data with complex topological structures. The input to our GCN model consists of a modified adjacency matrix that captures the connectivity of the network. It includes both node and line connections, along with sensor data. These sensors collect data on various electrical parameters, providing rich features for both nodes (e.g., voltage, current) and edges (e.g., power flow). By constructing a dual graph, we integrate both the node and edge features within the GCN model. This provides a comprehensive view of the state of the network.

The adjacency matrix is used to create the propagation matrix in our model. This matrix governs how information is distributed across the network during convolution operations. The convolution layers aggregate information from a node's

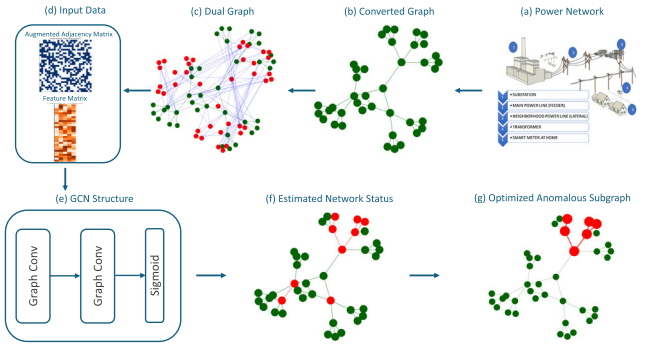

FIGURE 1. An overview of the proposed model. In (c), the edges are converted to nodes (shown in red) and all new edges are shown with dashed lines. The red nodes and edges in (f) and (g) represent the detected anomalous nodes and edges. The optimization model results in the removal of unconnected detected anomalous nodes and edges.

neighbors and connected edges. This effectively captures the complex relationships and dependencies inherent in the power network. This process enables the GCN to model how local anomalies can influence distant parts of the network due to the interconnected nature of power systems. A key enhancement in our approach is the dynamic update of the adjacency matrix. This updated matrix incorporates probabilistic weights that represent the likelihood of nonphysical connections between nodes and lines. This effectively captures the potential impact of faults at specific locations across the network. This probabilistic adjustment allows the model to adapt to network abnormalities and structural changes caused by failures, improving its response to anomalies. After the convolution and activation processes, the resulting feature maps are passed through fully connected layers. This leads to the final output layer. This layer provides the probabilities of faults occurring at all nodes and lines in the network. In addition, we propose a mathematical formulation that uses the output layer and the propagation matrix to accurately determine the sources of anomalies. This formulation enhances the model's ability to pinpoint the exact location of faults within the network by considering both the learned features and the network's topology.

The model detects and locates faults in a smart power grid by leveraging the structure and data of the power system in the form of a graph. The smart power system continuously provides data that is used to build a graph representation of the network. A Graph Convolutional Network (GCN) processes this graph, considering both the node/edge features and connectivity, to predict where faults are likely to occur. The final step uses these predictions to optimize fault localization, making the deployment of maintenance resources more efficient. This approach is particularly powerful because it not only uses sensor data (e.g., voltage, current), but also considers the relationships between different parts of the network. This makes it more effective in identifying complex fault patterns that might be missed by traditional methods. An overview of the proposed structure is shown in Fig. 1, including the workflow of a Graph Convolutional Network (GCN) model designed to predict fault locations in an electrical grid using dual graph representation. In the remainder of this section, we discuss each element of this model in detail.

2) CONSTRUCTION OF THE DUAL GRAPH AND AUGMENTED ADJACENCY MATRIX

The power network is represented as a graph $G(N, E)$, where N denotes the set of nodes (buses) and E denotes the set of lines (edges). Each node and line is equipped with sensors that provide crucial measurements for network monitoring and fault detection. Sensor data is organized into two feature matrices:

a: - NODE ATTRIBUTE/FEATURE MATRIX

$X_N \in \mathbb{R}^{|N|\times S_N}$, where S_N is the number of sensors per node. This matrix contains sensor readings for each node in the network.

b: - LINE ATTRIBUTE/FEATURE MATRIX

$X_E \in \mathbb{R}^{|E|\times S_E}$, where S_E is the number of sensors per line. This matrix contains sensor readings for each line.

To model the physical connectivity of the network, we use the original adjacency matrix $A \in \mathbb{R}^{|N|\times|N|}$, which captures the direct connections between nodes based on the physical layout of the power grid. To incorporate both node and line information into a unified framework, we construct a **dual graph** $G'(L, E')$. In this dual graph, the lines E of the original graph G are treated as nodes, and their connections are established based on shared nodes in G. This transformation is facilitated by the **incidence matrix** $B \in \mathbb{R}^{|N|\times|E|}$, which maps lines to their corresponding nodes in G. The adjacency matrix $A' \in \mathbb{R}^{|E|\times|E|}$ of the dual graph captures the connectivity between lines. Now, to fully integrate node-node, line-line, and node-line connections, we construct an **augmented adjacency matrix** A'' as follows:

$$A'' = \begin{pmatrix} A & B \\ B^\top & A' \end{pmatrix}$$

where

- A represents the adjacency between nodes (node-node connections) in the original graph.
- A' represents the adjacency between lines (line-line connections) in the dual graph.
- B and $B^\top$ represent the incidence relationships between nodes and lines (node-line connections).

This augmented matrix $A'' \in \mathbb{R}^{(|N+E|)\times(|N+E|)}$ effectively combines adjacency information from both the original graph and the dual graph, allowing us to treat edge features as node features within a comprehensive graph structure. In addition to the physical connections captured by A'', we introduce an **attention (propagation) matrix** $P = [p_{ij}] \in \mathbb{R}^{(|N+E|)\times(|N+E|)}$, which captures the probabilistic impact of anomalies propagating from one element (node or line) to another. Here, p_{ij} denotes the probability that an anomaly

in element i will affect element j. To model both the physical topology and the probabilistic propagation of anomalies, we update the augmented adjacency matrix A'' by integrating the attention matrix P:

$$A''_{\text{updated}} = \max(A'', P).$$

This updated matrix A''_{updated} incorporates both direct physical connections and indirect probabilistic influences between nodes and lines, increasing our ability to analyze the behavior of the network under fault conditions. We use the **max** function to calculate A'' in order to capture the *strongest influence*, whether it arises from physical adjacency or probabilistic propagation, between any two elements. This ensures that the final matrix reflects the most important connection, which is critical in scenarios where a probabilistic anomaly may have a stronger effect than a direct physical link. Moreover, using the max function avoids *diluting strong propagation paths* that could occur with averaging, especially when the propagation matrix P reveals latent but impactful interactions not captured in the physical topology.

c: - FEATURE MATRIX AUGMENTATION

We also augment the feature matrices to align with the structure of A''_{updated}:

$$X = \begin{pmatrix} X_N \\ X_E \end{pmatrix},$$

where $X \in \mathbb{R}^{(|N+E|)\times S}$ combines the node and line features into a single matrix, with $S = \max(S_N, S_E)$.

3) GRAPH CONVOLUTIONAL NETWORK IMPLEMENTATION

Using the updated adjacency matrix A''_{updated} and the augmented feature matrix X, we employ a graph convolutional network (GCN) to process the graph. The GCN leverages the enhanced graph structure to learn representations that capture both physical and probabilistic interactions within the network. This approach improves the accuracy of the location and prediction of the fault by considering the full spectrum of available data. The GCN operates by iteratively updating the feature representations of nodes and lines through layers that aggregate information from their neighbors, as defined by A''_{updated}:

$$X^{(l+1)} = \sigma\left(\hat{D}^{-\frac{1}{2}}\hat{A''}\hat{D}^{-\frac{1}{2}}X^{(l)}W^{(l)}\right), \tag{11}$$

Eq. (11) describes how the features of the nodes and lines are updated and aggregated from the neighboring nodes and edges in each layer of the GCN. The convolution utilizes the normalized adjacency matrix $\hat{A''}$, the degree matrix $\hat{D}$, and the weight matrix $W^{(l)}$ to apply the nonlinear activation function σ. With $X^{(0)} = X$, the augmented feature matrix that includes the node and edge features, where $\hat{D}_{ii} = \sum_j \hat{A''}_{ij}$. The weight vector $W^{(l)}$ is learned during the training process. The final output layer of the GCN uses a *sigmoid* activation function to map the learned features to the anomaly probabilities:

$$\boldsymbol{q} = \sigma\left(X^{(L)}W^{(L)}\right),$$

where $\boldsymbol{q} = [q_1, q_2, \ldots, q_{|N+E|}]$ is the vector of estimated anomaly probabilities for all nodes and lines. Each element q_i represents the estimated probability that an anomaly occurs at node or line i. By analyzing $\boldsymbol{q}$, we can identify and localize faults within the network using two common approaches:

1) **Thresholding**: Set a predefined threshold τ (e.g., $\tau = 0.5$). If $q_i \geq \tau$, node or line i is considered anomalous; otherwise, it is considered normal.
2) **Ranking**: Rank all nodes and lines based on their anomaly probabilities q_i. Nodes or lines with higher q_i values are deemed more likely to be anomalous. This method is particularly useful when the number of anomalies is unknown or when prioritizing resources for further investigation.

Using these methods, $\boldsymbol{q}$ effectively facilitates the detection and localization of anomalies within the network.

C. LOCATING THE SOURCES OF ANOMALIES: AN OPTIMIZATION APPROACH

Using the anomaly probability vector $\boldsymbol{q}$ through thresholding and ranking is a foundational approach for anomaly detection. However, these methods face significant challenges. Selecting appropriate thresholds is non-trivial and can greatly affect detection accuracy. An improperly chosen threshold may result in numerous false positives or negatives. The potential for detecting widespread anomalies can overwhelm resources and obscure meaningful insights. The model might predict a high probability of anomalies across many nodes and lines without a clear justification. In addition, the lack of contextual and relational information hampers the ability to understand and justify detected anomalies. The model does not connect anomalies based on network topology or the relationships between nodes and lines. Scalability and adaptability issues also arise in large or dynamic networks. Computational demands and evolving conditions can affect performance. Addressing these challenges may require integrating additional information, such as network topology, using adaptive thresholds, enhancing the interpretability of the model, and designing algorithms that consider interconnections between nodes and lines. This would provide a more cohesive and justifiable anomaly detection strategy. In this section, we propose an optimization model that addresses some of these challenges in an interpretable manner by incorporating topological information about the network.

Given the probability p_{ij} of the propagation of the anomaly from node i to node j, a system of equations can be formulated to identify the sources of the anomalies. This system incorporates a balancing parameter α and is expressed as follows:

$$q_i = \alpha x_i + (1-\alpha)\sum_{j=1}^{|N+E|} x_j p_{ji},\ i \in \{1, \ldots, |N+E|\}. \tag{12}$$

The equation for balancing, as depicted in Eq. (12), demonstrates the adjustment of probability by considering both direct effects and propagation influences within the network. The variables x_i represent the likelihood that node i is the cause of anomalies. The parameter α is a weighting factor that changes the model's emphasis on the inherent anomaly generation capabilities of each node and the impacts exerted on each node by other nodes. When the value of α is set to 1, the model only takes into account the inherent probabilities x and fully disregards the influence of propagation from other nodes. This scenario operates on the assumption that the nodes produce anomalies in a manner that is not influenced by each other. When the value of α is set to 0, the model assigns the observed anomalies entirely to the effects of other nodes, as defined by the propagation matrix P, without considering any inherent characteristics of the nodes themselves. The values of α ranging from 0 to 1 enable the model to consider both inherent and transmitted influences, effectively balancing these aspects based on the magnitude of α. A higher value of α increases the importance of the node's own likelihood of generating anomalies, whereas a lower value of α highlights the relevance of interactions between nodes. In basic network layouts, such as radial distribution networks, nodes generally function with significant independence. In this context, positioning α around 1 is beneficial, highlighting the intrinsic anomaly production capacities of individual nodes. In contrast, in intricate or interconnected network architectures, where node interactions substantially affect operational dynamics, an α value between 0 and 1 is ideal. This enables the model to detect both anomalies produced by the nodes and those transmitted through network interactions.

Selecting the appropriate value of α is essential to accurately represent the behavior of the network. To optimize this parameter, it is necessary to have an extensive understanding of the network dynamics and potentially perform an iterative calibration using real-world data. This equilibrium aids in discerning the most probable origins of anomalies by taking into account both the potential independent generation of anomalies by nodes and their potential impact on, or sensitivity to, the network's structure. Given that x_is are probability values for the dual network with both nodes and edges as nodes, we need to ensure the following is satisfied:

$$\sum_{i=1}^{|N+E|} x_i = 1 \quad \text{and} \quad 0 \le x_i \le 1 \tag{13}$$

To rewrite Eqs. (13)-(14) in matrix form, we define the vectors and matrices as follows:

$$\boldsymbol{q} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}, P = \begin{bmatrix} p_{11} & \cdot & p_{1|N+E|} \\ \cdot & \cdot & \cdot \\ p_{N1} & \cdot & p_{N|N+E|} \end{bmatrix} \tag{14}$$

The system of equations in Eq. (10) can be rewritten in vector-matrix form as:

$$\boldsymbol{q} = \alpha \boldsymbol{x} + (1-\alpha)\mathbf{P}^T \boldsymbol{x},$$

where $\mathbf{P}^T$ denotes the transpose of the matrix $\mathbf{P}$. Rearranging the equation, we get:

$$\boldsymbol{q} = (I + (1-\alpha)\mathbf{P}^T)\boldsymbol{x}.$$

To solve for $\boldsymbol{x}$, we need to isolate it as

$$(I + (1-\alpha)\mathbf{P}^T)\boldsymbol{x} = \boldsymbol{q}.$$

Given that $I + (1-\alpha)\mathbf{P}^T$ is invertible (because all its eigenvalues are positive, ensuring that its determinant is nonzero), we can multiply both sides by the inverse of $I + (1-\alpha)\mathbf{P}^T$:

$$\boldsymbol{x} = (I + (1-\alpha)\mathbf{P}^T)^{-1}\boldsymbol{q}.$$

In light of the limitations outlined in Eq. (13)), it is imperative to verify that the solution vector $\boldsymbol{x}$ complies with these requirements. These constraints can be effectively addressed by employing numerical optimization approaches. An example of a method is the constrained optimization strategy in which we solve the following linear problem:

$$\begin{aligned} &\text{minimize } \|(I + (1-\alpha)\mathbf{P}^T)\boldsymbol{x} - \boldsymbol{q}\| \\ &\text{subject to } \sum_{i=1}^{|N+E|} x_i = 1, \\ &\qquad 0 \le x_i \le 1, \quad \forall i = \{1, 2, \ldots, |N+E|\}. \end{aligned}$$

Now, we can use the mode of the optimized vector $\boldsymbol{x}$ to identify the most probable origins of faults in the network. The objective function is quadratic because it relies on the metric norm, which is common in least squares problems. Quadratic objective functions are convex when the Hessian matrix is positive semi-definite. Since the objective function is quadratic and the constraints are linear, we can conclude that the optimization problem is convex. Convex optimization problems have the advantage that any local minimum is also a global minimum, making it easier to find the solution. Overall, the optimization problem, with its quadratic objective function and linear constraints, is well-suited for convex optimization methods, which can solve it effectively. The complexity of the problem is typically expressed as a polynomial function of the number of variables. This means that it can be effectively managed for moderately large problems. This allows network analysis and anomaly detection to be applied in practical settings, even with a large number of nodes and variables. The vector $\boldsymbol{x} = [x_1, \ldots, x_{|N+E|}]$, obtained from the optimization model in Section III-C, represents the likelihood that each node or line in the network is the source of an anomaly. The most likely fault source is identified as the element with the highest value in $\boldsymbol{x}$, which is labeled as faulty in the network status. The probability vector $\boldsymbol{x}$ also influences the status of other nodes and lines. If a node or line is connected to others with a high probability of fault, it is more likely to be marked as faulty. This

Algorithm 1 Update Network Status Based on Optimal Anomaly Source Probabilities

Input:
- Initial labels: $\hat{y} = \{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_{|N+E|}\}$
- Source probabilities: $x = \{x_1, x_2, \ldots, x_{|N+E|}\}$
- Augmented propagation matrix: $P_{\mathbb{R}^{(|N+E|)\times(|N+E|)}}$

Output:
Updated labels $\hat{y} = \{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_{|N+E|}\}$

Step 1: Identify Most Likely Anomaly Source
$i^* \leftarrow \arg\max_j x_j$
$\hat{y}_{i^*} \leftarrow 1$
$e^{i^*} \leftarrow$ one-hot vector for i^*

Step 2: Compute Influence Scores
for $i = 1$ to $|N + E|$ **do**
  $f_i \leftarrow \sum_{j=1}^{|N+E|} e_j^{i^*} \cdot p_{ji}$
end for

Step 3: Update Labels
for $i = 1$ to $|N + E|$ **do**
  $\hat{y}_i \leftarrow \hat{y}_i \cdot \text{sign}(f_i)$
end for

Return: Updated label vector $\hat{y}$

approach ensures that the network status reflects both the direct identification of faults and their ripple effects throughout the network. Any nodes or lines previously estimated to be anomalous but not stochastically connected to the detected source should be labeled as normal. This approach aims to remove false positives—elements that were mistakenly identified as anomalous. In Algorithm 1, we present all the steps in a matrix form, providing a scalable method to update the estimated status of the network once the results of the optimization model are obtained.

IV. NUMERICAL EXPERIMENTS

In this article, we used the IEEE 118-bus test system [34]. This system is commonly used in power system research to evaluate new methodologies for fault analysis, stability evaluation, and grid optimization techniques. This system presents a challenging environment that closely resembles the operational characteristics and constraints of real large-scale power networks. The complexity of fault analysis in this system arises from its sophisticated network topology. The IEEE 118-bus test system is complex due to its interconnected loops and paths. These complexities make it difficult to accurately determine the location of faults. This is because it is challenging to determine the effect of faults based solely on voltage and current readings. Furthermore, the presence of several loops and pathways adds complexity in calculating changes in network characteristics caused by faults. The network configuration includes 12 interconnection lines across

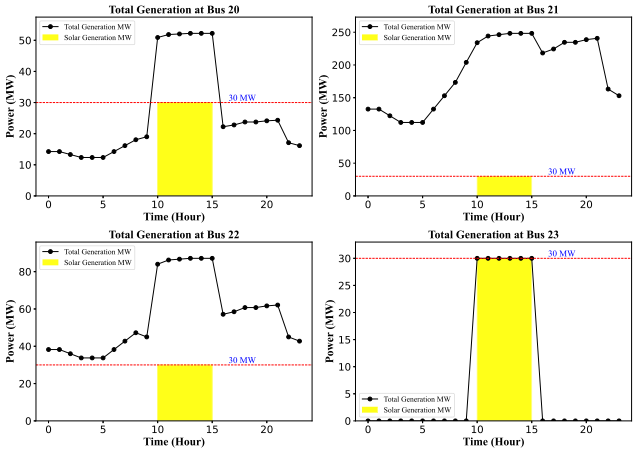

FIGURE 2. Total generation (MW) at selected buses.

three zones: 7 lines connect Zone 2 to Zone 1, and 4 lines link Zone 2 to Zone 3. Each of these lines is equipped with a switch, allowing the operator to disconnect them as needed or according to changes in the operational plan. We thoroughly evaluated the performance of the proposed model and compared it with benchmark models in three distinct sections. The power system in our study operates continuously for 24 hours and consists of 22 solar generators distributed in three zones. Zone 1 has 10 solar generators, each capable of producing 30 MW. Zone 2 has 4 generators, each generating 25 MW. Zone 3 consists of 8 generators, each with a capacity of 2 MW. Fig. 2 shows the total electricity generation from buses 20, 21, 22, and 23, including the contribution from solar power. We analyze various scenarios, including line and bus outages, over a 24-hour period. We simulated 118 anomaly scenarios, one for each bus outage, and 173 anomaly scenarios, one for each line outage, on the IEEE 118-bus system.

The network uses a configuration of six sensors: four for nodes (buses) and two for lines. The sensors deployed at the nodes measure active power, reactive power, voltage amplitude, and the node's angle. These measurements are crucial for monitoring and controlling the power system's state. Meanwhile, the sensors on the power lines monitor the active and reactive power flow. These smart meters help achieve highly reliable load forecasts. In addition, we used these predictions as input for Monte Carlo simulations, treating them as mean load values with a variance of 0.5. The system simulation was conducted within the Jupyter environment, using the Pandapower library for power system modeling. We assume that the available sensors at the nodes and buses can be used for fault detection. In addition, we studied the impact of partial sensor availability in the network to make the analysis more realistic (see Section 4.3.3). -*Summary of Power Flow Simulation*: We used Pandapower to simulate our network, running the simulation for each time step (one hour). Pandapower can simulate the load profile on the IEEE 118-bus system by leveraging its time series simulation

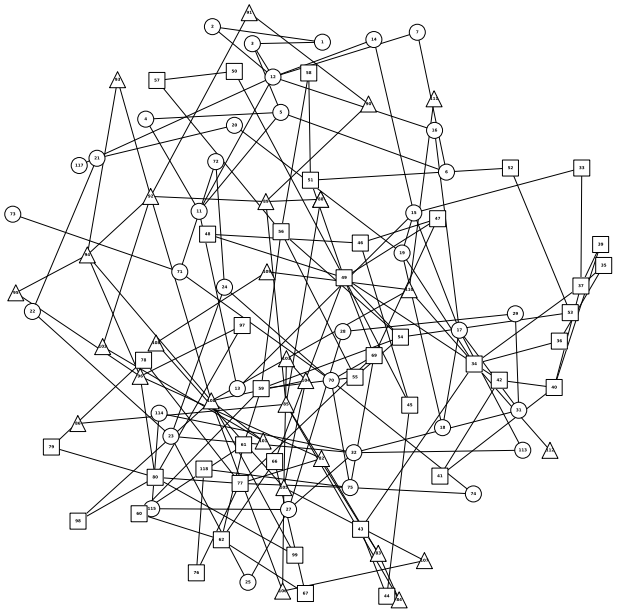

FIGURE 3. Graph of IEEE 118 - Buses on the same zone have the same shape.

TABLE 1. Model hyperparameters.

Hyperparameter	Value / Description	Hyperparameter	Value / Description
Number of GCN layers	2	Activation functions	ReLU (hidden), Sigmoid (output)
Input feature size	6 (4 nodes; 2 edges)	Epochs	100
Hidden layer sizes	16, 32	Loss function	Binary Cross-Entropy
Optimizer	Adam	Learning rate	0.001

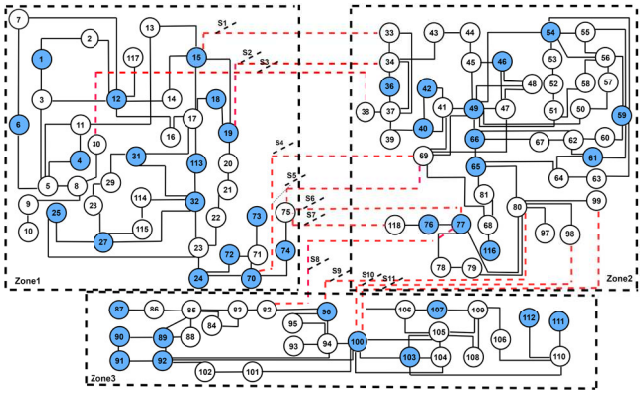

FIGURE 4. Single diagram of IEEE 118-bus test system (blue-colored buses indicate the presence of power generation sources).

capabilities. Initially, the standard IEEE 118-bus network is loaded, providing the system topology and parameters. A load profile representing the variation of loads over 24 hours is defined according to the IEEE 118-bus system, typically using scaling factors to reflect peak and off-peak demand periods throughout the day. To achieve more accurate estimations, we use Monte Carlo simulations for the load modeling of each bus. These scaling factors are applied to adjust the loads on the network at each time step. Subsequently, for each bus containing renewable energy, we account for specific time periods during which electricity output is available. We assumed a fixed value for the solar panel output rather than using a probabilistic approach. Pandapower runs power flow calculations for each hour, updating the loads according to the defined profile. This process enables the analysis of the performance of the system under varying load conditions throughout the day. It allows the assessment of voltage profiles, line loads, and other operational parameters without having to refer to specific code implementations.

-*Graph Topology*: The graph topology of the IEEE 118-bus system is shown in Figs. 3 and 4. Fig. 3 presents the graph of the IEEE 118-bus system, illustrating its complexity. Fig. 4 illustrates the IEEE 118-bus test system, where blue-colored buses indicate power generation sources, such as generators and solar panels. The system is divided into three zones, with various connections and switches (S1, S2, etc.), which are typically closed to manage the flow of power throughout the network. As observed, the network is complex, which leads some references to divide it into three distinct zones. These zones are interconnected, which means that a fault or interruption in one zone can affect the power flow and stability of other zones. Managing the power flow between these zones requires sophisticated control mechanisms to maintain voltage levels, frequency, and system stability. In particular, the power system exhibits numerous rings or cycles within the graph topology of each zone. These cycles add complexity to the analysis, particularly in the context of anomaly detection. The proposed framework includes several hyperparameters and model setup variables, which we tuned using a limited grid search over a small, representative subset of data. The final selection reflects empirically optimal values within this scope. A summary of the key hyperparameters and model configuration is provided in Table 1.

A. SIMULATING ANOMALIES AND THEIR PROPAGATION

Fault detection in the IEEE 118-bus system is particularly challenging due to its size and complexity. The interconnected nature of the system means that a fault in one part of the network can propagate and impact other areas, making it difficult to isolate the fault quickly. Furthermore, the presence of multiple zones adds to the complexity, as each zone could have different operating conditions and sensitivities to faults. To conduct a comprehensive analysis, we assumed numerous scenarios for outages, considering anomalies at all nodes and lines.

In this study, we simulated faults in both nodes and lines of the power system network, assuming a complete outage for each affected line or node. Based on this assumption, we modeled the power system's response to these faults. After each simulation, all sensor data was collected and analyzed to assess the impact of the faults on the entire network, providing insights into fault detection and system stability under different fault conditions. Within the power network simulation, we established fault indicators by setting precise thresholds for active and reactive power, phase angle, and voltage levels. These indicators serve to detect abnormal

conditions and expected problems. We assumed that a 30% increase or decrease in active or reactive power on the transmission lines implies a significant departure from typical operational circumstances, potentially indicating faults such as short circuits, uneven loads, or equipment breakdowns. In addition, we considered a power angle deviation exceeding 45 degrees or falling below 30 degrees as an additional indication of instability. These angles imply a deviation from stable power flow, potentially caused by synchronization problems or disturbances in the system. Regarding voltage levels, we employed a threshold of surpassing 1.05 per unit or falling below 0.95 per unit. This threshold represents the boundaries at which voltage conditions could compromise both the reliability of the system and the quality of power, potentially prompting the activation of preventative measures. These criteria collectively function as fault indicators to aid in the detection and diagnosis of faults, hence assuring the stability and safety of the network. For line outages, there are 173×24 instances. For bus outages, there are 118×24 instances. After each outage simulation, our predefined indicators allow us to determine the impact of each fault on specific lines and nodes, depending on the timing and location of the fault occurrence. Using this data, we can calculate the probability of a fault impacting each particular line or node over the 24-hour period. This result helps us to update the adjacency matrix of the network. Given this realistic preparation of scenarios, anomalies are propagated according to the current load of the system, the topology of the network, and other relevant factors, making the model highly dynamic and realistically simulated for anomalies. To construct the anomaly propagation matrix, we simulated anomalies 50 times for each node and line, recording the frequency with which anomalies propagated from node i to node j. This resulted in an empirical probability matrix capturing the likelihood of anomaly spread across the network. The choice of 50 simulations was based on preliminary observations indicating that the propagation probabilities stabilized around this point, offering a practical balance between statistical reliability and computational efficiency. It should be pointed out that although generating the propagation matrix P requires a large number of simulations (e.g., 15,000 for the IEEE 118-bus system), this process is performed entirely offline and does not impact the model's real-time performance. Additionally, because each simulation is independent, the process is highly parallelizable and well-suited for distributed computing, making it scalable to larger networks.

It is important to note that if an anomaly occurs at a node or line in the IEEE 118 system, it does not always impact the same downstream nodes. The impact depends on the system topology, the type of anomaly, and the effectiveness of the protection and control mechanisms. In meshed systems such as IEEE 118, power can be rerouted and the system can adjust to anomalies, potentially minimizing or even preventing impacts on downstream nodes. In summary, anomalies were introduced for each line and bus within the power system, with the assumption that these anomalies directly result in an outage of the affected line or bus. A 24-hour simulation was performed during which a node/line outage was assumed to occur at each time step. Sensor data was collected from 173 lines and 118 buses throughout this period. Similarly, a bus failure was simulated at each time interval over a 24-hour duration, with sensor data collected again from 173 lines and 118 buses.

B. GRAPH CONSTRUCTION AND INITIAL FEATURE TRANSFORMATION

In the proposed model, we employ a Graph Convolutional Network (GCN) to predict fault locations within an electrical grid comprising 118 buses and 173 lines. The architecture is uniquely designed to handle dual graph representations — one representing the buses and another for the lines. Each bus node is characterized by data from four sensors, resulting in a node feature matrix with dimensions 118×4. Similarly, each line, represented as an edge in the graph, is characterized by data from two sensors, forming an edge feature matrix of size 173×2. For the initial graph construction, we create two adjacency matrices: A_{node} for the nodes (buses), and A_{edge} for the edges (lines), sized 118×118 and 173×173 respectively. These matrices are instrumental in defining the connectivity and relationship between the various components of the grid. The input features are normalized to standardize the data range between different sensors. To effectively utilize edge features within our GCN model, we employ a dual graph representation. This approach involves treating lines as nodes in a secondary graph, allowing us to directly incorporate line-specific sensor data into the graph-based learning framework. Using a transfer function, we construct a comprehensive adjacency matrix that encapsulates both node and edge features. This matrix serves as the foundational structure for the GCN, enabling the propagation of features across the graph according to the defined relationships.

- *Output Layer and Model Training*: The final output layer of GCN applies a sigmoid activation function to map the features to probabilities, resulting in a 291×1 matrix. Each element of this matrix represents the probability of a fault at the corresponding node or edge, which is vital to identify potential problems throughout the grid. The model uses a binary cross-entropy loss function, which is well-suited for the binary nature of fault prediction. We employ the Adam optimizer, known for its effective handling of sparse gradients and adaptive learning rate capabilities, making it particularly suitable for training deep graph-based neural networks. This architecture not only facilitates effective feature learning and transformation specific to the topology and dynamics of electrical grids, but also enhances the ability to accurately predict and localize faults by leveraging both node- and edge-level data.

C. ANOMALY DETECTION RESULTS

We evaluated the effectiveness of our model in detecting anomalies in a binary setting for all nodes and lines, where label 1 indicates an anomaly and label 0 indicates

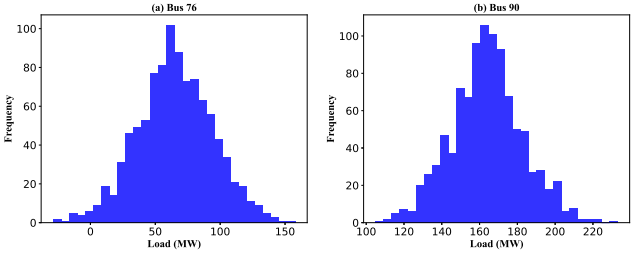

FIGURE 5. Histogram of load on buses 76 and 90.

normal. The performance metric used are accuracy, precision, F1-score, coverage, and recall, each providing unique insights into the model's ability to detect anomalies. Accuracy measures the overall proportion of correctly identified instances, both anomalies and normal cases, but can be misleading in imbalanced datasets where one class dominates. Precision focuses on the model's ability to correctly identify anomalies among all instances it predicts as anomalies, indicating a low false positive rate when high; however, overly emphasizing precision may cause the model to miss actual anomalies if it becomes too conservative. Recall, also known as sensitivity, assesses the model's capacity to detect all actual anomalies, minimizing false negatives, but optimizing for recall alone might increase false positives. The F1-Score balances precision and recall by calculating their harmonic mean, offering a more comprehensive metric that considers both false positives and false negatives, which is particularly useful in cases with uneven class distribution or differing costs of errors. Coverage evaluates the proportion of true anomalies detected out of all instances, reflecting the model's efficiency in detecting anomalies throughout the dataset. Collectively, these metrics provide a thorough assessment of the model's strengths and weaknesses in anomaly identification, allowing for a nuanced understanding of its performance. Our benchmark models include GCN (Graph Convolutional Network), GNN (Graph Neural Network), GAT (Graph Attention Network), NN (Neural Network), and SVM (Support Vector Machine), all of which are used to detect anomalous nodes in a supervised manner.

1) POWER OF THE MODEL IN DETECTING ANOMALOUS NODES AND REGIONS

In this section, we assume that the switches on the communication lines are closed and that the data from the line and bus sensors have been collected completely by the operator. Under these conditions, we conducted simulations using both the standard GCN and the proposed method, which employs the propagation matrix with a threshold setting of 0.25, applied separately to lines and nodes. The performance results of these experiments are detailed in Table 2 and the accumulated results for the lines and nodes are summarized in Table 3. In both tables, the results based on benchmark models are also reported. Using a dataset of 50 samples $\times$ 118 buses $\times$ 24 hours for nodes and 50 samples $\times$ 173 lines $\times$ 24 hours for lines, this simulation assesses the effectiveness of several techniques. We evaluated the performance of our model by comparing it with an SVM with a sigmoid kernel $\gamma = 0.01$ and $C = 100$, a GNN with a 2-layer architecture (582 and 291 units), and a Neural Network (NN) with 291 input neurons and two hidden layers consisting of 150 and 75 neurons, respectively. The nodes and edge features and anomaly labels are utilized as inputs to train these models. The proposed approach consistently surpasses previous models, achieving the highest accuracy (97.83% for nodes and 98.95% for lines) and superior precision (96.12% for nodes and 98.48% for lines). The F1 score (97.95% for nodes and 99.19% for lines) and recall (99.85% for nodes and 99.91% for lines) substantially exceed the standards, demonstrating better detection of true positives. Furthermore, it has exceptional coverage (95.98% for nodes and 98.40% for lines), highlighting its efficiency in identifying anomalies throughout the network. Table 3 illustrates the overall performance of the proposed technique of combining nodes and lines. The proposed strategy significantly improves the alternative methods, achieving the highest scores in all parameters. It achieves the highest accuracy (98.50%), precision (97.63%), F1 score (98.75%), coverage (97.52%), and recall (99.89%). These results highlight the effectiveness and reliability of the proposed approach in accurately detecting faults, making it the most robust strategy compared to the alternative models evaluated.

In our experiments on the IEEE 118-bus system, the inference time of the trained model was on the order of seconds, suggesting a strong potential for real-time or near-real-time deployment. Moreover, the offline training phase—although more resource-intensive—is a one-time cost, while the forward pass and the optimization model used in real-time operation remains lightweight. For larger networks, scalability can be further enhanced through batching, model pruning, or deployment on parallel computing platforms. These characteristics position our approach as a promising candidate for practical integration into real-time grid control systems.

D. ADDITIONAL INSIGHTS FOR SELECTED NODES AND LINES

Tables 4 and 5 present the precision and accuracy of a selected set of important lines and nodes. Line 42 is a crucial transmission line that handles significant power transfers, but it has the lowest precision compared to all other lines. Line 42's lower precision and accuracy are due to its high and variable power flow. Although other methods might still face this challenge, the proposed method demonstrates higher precision in predicting anomalies for this line compared to alternative approaches. Node 76, located in the third zone, is essential for the integration of solar energy into the network. However, it shows the lowest accuracy compared to all other nodes. The lower precision and accuracy at this location are likely caused by variations in both load and generation. Our examination of the load and production patterns in this region shows significant fluctuations in both load and generation. The existence of these variations at node 76 is likely to be the fundamental

TABLE 2. Comparison with benchmark models - separated by node and line levels.

		Accuracy (%)	Precision (%)	F1 Score (%)	Coverage (%)	Recall (%)
Node Anomalies ($50 \times 118 \times 24$)	Proposed	97.83	96.12	97.95	95.98	99.85
	GCN	86.30	84.87	87.17	77.25	89.59
	GAT	87.00	89.54	89.95	95.26	90.01
	GNN	74.66	79.19	75.37	60.47	71.89
	NN	69.29	60.42	70.82	54.82	85.54
	SVM	76.72	77.27	78.10	64.07	78.94
Line Anomalies ($50 \times 173 \times 24$)	Proposed	98.95	98.48	99.19	98.40	99.91
	GCN	89.50	89.68	90.92	83.35	92.19
	GAT	87.15	89.00	89.68	95.16	90.25
	GNN	85.20	87.56	86.90	76.84	86.25
	NN	70.71	79.13	72.98	57.46	67.72
	SVM	78.11	91.13	79.17	65.52	69.98

TABLE 3. Comparison with benchmark models - combined for node and line levels.

Method	Accuracy (%)	Precision (%)	F1 Score (%)	Coverage (%)	Recall (%)
Proposed Method	98.50	97.63	98.75	97.52	99.89
GCN	88.20	87.81	89.47	80.94	91.19
GAT	87.07	89.27	89.81	95.21	90.13
GNN	80.93	84.44	82.48	70.19	80.62
NN	70.13	70.58	72.12	56.40	73.73
SVM	77.55	84.96	78.73	64.92	73.35

TABLE 4. Performance metrics for selected lines: proposed method and GCN.

	Line 64 (45-49)	Line 36 (17-31)	Line 132 (92-93)	Line 42 (33-18)	Line 107 (75-69)
Proposed Method					
Accuracy (%)	98.23	99.43	96.65	93.12	94.76
Precision (%)	97.12	96.34	95.19	92.20	94.45
F1 Score (%)	98.64	96.12	97.65	90.34	93.56
Coverage (%)	95.76	96.45	95.17	89.94	94.87
Recall (%)	96.23	97.43	96.76	92.67	95.76
GCN					
Accuracy (%)	85.65	89.10	74.26	83.69	86.59
Precision (%)	83.56	87.85	73.36	80.30	85.26
F1 Score (%)	84.25	86.21	74.15	81.20	84.25
Coverage (%)	83.12	87.12	74.87	80.12	85.13
Recall (%)	82.35	87.54	74.12	79.25	84.96

TABLE 5. Performance metrics for selected nodes: Proposed method and GCN.

Metric	Node 17	Node 69	Node 76	Node 90
Proposed Method				
Accuracy (%)	97.25	96.36	94.98	98.89
Precision (%)	95.46	94.36	92.50	95.63
F1 Score (%)	97.87	97.52	94.52	97.62
Coverage (%)	94.52	94.52	91.85	96.23
Recall (%)	94.12	93.89	93.65	95.89
GCN				
Accuracy (%)	90.89	86.98	87.69	87.79
Precision (%)	86.62	89.74	85.21	88.84
F1 Score (%)	91.25	87.24	86.57	87.96
Coverage (%)	85.96	89.32	84.99	88.54
Recall (%)	84.62	89.76	83.26	87.96

cause of the reduced accuracy and precision reported. With a 35 MW solar panel and a demand of 68 MW at node 76, the solar plant produces more than 50% of the total power, especially when considering the operational hours of the solar panels from 9 AM to 4 PM. The precision and accuracy of this node are affected by the significant variance in load and the oscillations in generation due to the uncertain nature of solar power. The histogram in Fig. 5 illustrates the load distribution at node 76, showing significant fluctuations. However, the load distribution of node 90, shown in Fig. 5, demonstrates lower fluctuations and a smaller variation. This comparison reveals that node 90 exhibits more stable load patterns, resulting in higher precision and accuracy for anomaly detection. In contrast, node 76 exhibits worse precision due to its significant load and generation fluctuations.

1) POWER OF THE PROPOSED MODEL TO LOCATE THE SOURCE OF ANOMALIES

As discussed earlier, one important aspect of the proposed model is its ability to detect the sources of anomalies. To show the effectiveness of our model in locating anomalies using the optimization approach described in Section III, we visualized our simulation results for Line 99 in Fig. 6. Panel A highlights the neighborhood affected when a fault occurs on this line. Panel B displays the estimated anomaly locations, while panels C and D compare the selected nodes/lines without and with the application of the optimization model, respectively. Similarly, Fig. 7 presents the simulation results for

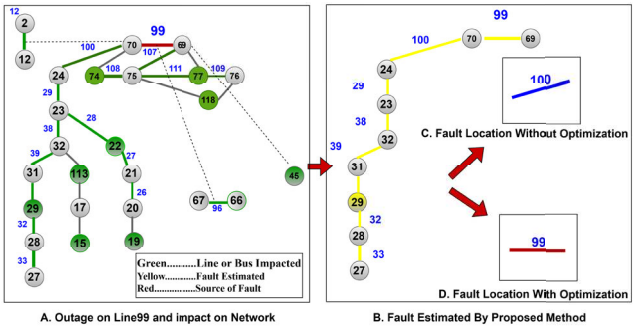

FIGURE 6. A. Anomaly on line 99 and its impact on the network; B. Estimated anomalous nodes and edges.

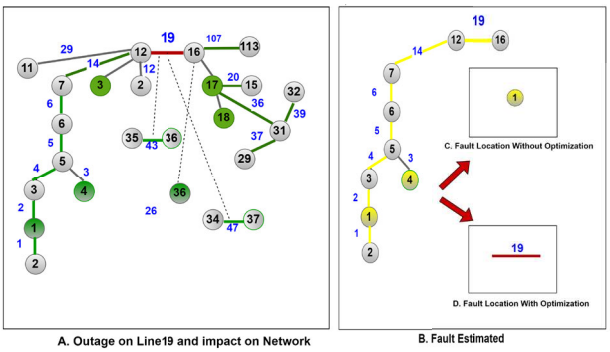

FIGURE 7. A. Anomaly on line 19 and its impact on the network; B. Estimated anomalous nodes and edges.

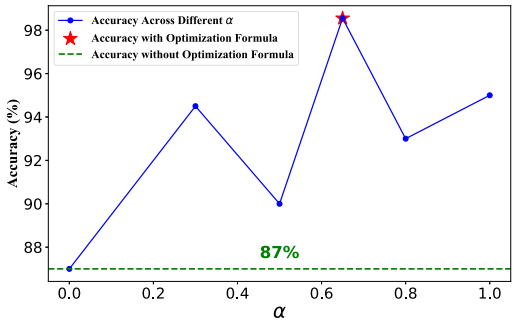

FIGURE 8. Accuracy versus different values of α in the optimization model.

Line 19, which show a comparable outcome to Line 99. In both examples, the sources of anomalies can be accurately identified using the optimization model. By applying the optimization model, we calculated the accuracy for different values of the parameter alpha. As shown in Fig. 8, the highest accuracy, 98.55%, is achieved at $\alpha = 0.65$. In contrast, the maximum accuracy of fault detection without using the optimization algorithm is only 87%. This clearly demonstrates that our proposed method significantly improves fault detection accuracy. In general, the optimization approach consistently produces more precise and reliable results for fault localization.

The sensitivity analysis of the propagation threshold, as depicted in the Fig. 9, indicates a distinct trend across multiple performance indicators. As the propagation threshold diminishes, the model's performance—assessed by accuracy, precision, F1 score, coverage, and recall—considerably enhances, especially within the threshold interval of around 0.25 to 0.5. This suggests that reduced threshold values improve the model's capacity to detect sophisticated propagation patterns of anomalies in the network.

The results in Fig. 10 further validate the effectiveness of the Proposed Method in detecting the true sources of anomalies. illustrates the performance of an anomaly detection model in identifying the exact sources of anomalies in a power network. It compares the total number of cases to the number of true matches for three categories: Line, Node, and Line and Node. The model successfully detected approximately 64.19% of Line anomalies, 51.63% of Node anomalies, and 59.10% of combined Line and Node anomalies. In addition to the 59.1% exact match rate (hop = 0), our proximity-based evaluation shows that the model identifies the true source within 1-hop in approximately 85% of cases, within 2-hops in about 91%, and within 3-hops in over 95%. These results underscore the model's strong operational value, even when strict exact matches are not achieved.

This indicates that the model is more effective at detecting Line-related anomalies compared to Node-related ones, with the combined category falling in between. Although the model does not always detect the exact source of an anomaly, it often identifies sources that are in close proximity, often adjacent to the true source. From an operational standpoint, this level of accuracy is highly beneficial for the maintenance team. However, in terms of our strict performance metrics, it is still considered a "no match." This distinction highlights that the practical performance of our model is stronger than the raw metrics suggest. Furthermore, MOST existing benchmark models are still unable to pinpoint anomaly sources as accurately as our model. Additionally, our model has the capability to filter out many false positives based on the propagation matrix once the sources are identified, providing more precise and actionable results for power system management.

2) MODEL PERFORMANCE ACROSS VARIOUS NETWORK TOPOLOGIES

The power system studied in this article consists of 11 interconnection lines, which we consider as switches used by the Independent System Operator (ISO) to change the system's configuration. As a result, there are a total of 2048 (2^{11}) possible states. We used these switches to assess the system's performance across 11 randomly selected topologies. Changing a switch in the power system results in alterations in topology, adjacency matrix, and power flow.

Initially, we conducted tests using the same training data as before, meaning that we did not modify the adjacency matrix or power flow data used to train the proposed Graph Convolutional Network (GCN). However, changing a switch generates a new adjacency matrix and affects the power flow. To improve performance, we retrained the system based on

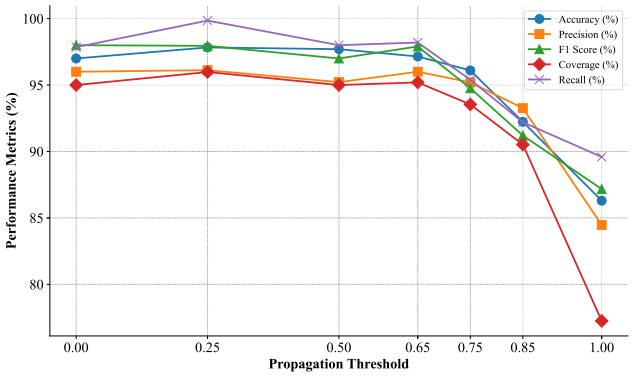

FIGURE 9. The propagation threshold performance.

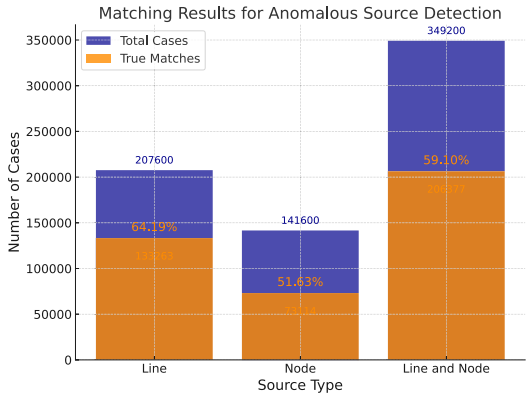

FIGURE 10. Performance of the model in source detection - There are 349,200 total cases (50 runs × 24 hours × 118 nodes + 50 runs × 24 hours × 173 lines).

each topology and conducted subsequent tests under modified circumstances. After retraining, we applied the proposed method to the new data, referring to this approach as the Proposed Method-TR. We then investigated the performance of the GCN and the proposed approach under different conditions by changing the switches. Performance measures are presented in Fig. 11. It should be pointed out that we used a fixed value of $\alpha = 0.65$ across all topologies without individual tuning. However, α is sensitive to the network topology, and topology-specific tuning is a potential avenue for further improvement.

As indicated in Fig. 11, the model exhibits resilience in the face of these modifications. The improved efficiency of the retrained model, which uses the propagation matrix to update the adjacency matrix, is clearly visible compared to the basic model. The fault detection accuracy is lowest when there is a topological change between areas 1 and 2, particularly on line 41. Analysis of this connection revealed that it provides the most effective transmission of power between these regions. Modifying the network topology in this manner requires increasing the production capacity in area 1 to meet consumer demands, which results in notable changes in power distribution between lines and buses in that region. Consequently, these significant adjustments lead to a decrease of approximately 20% in the probability of

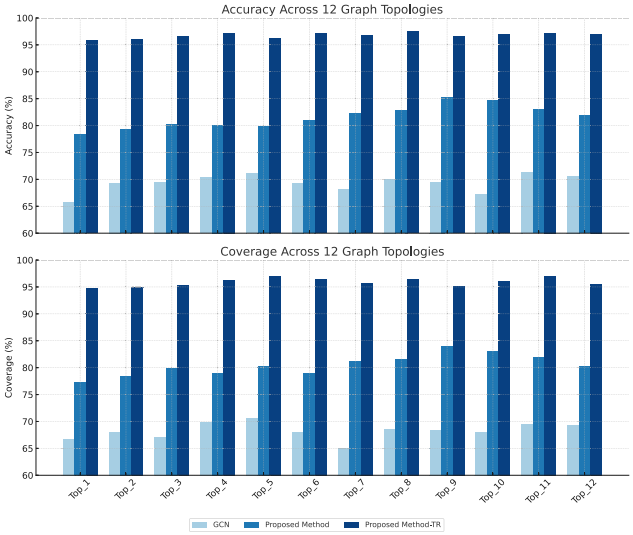

FIGURE 11. Performance comparison across graph topologies.

detecting faults in the improved model before retraining. To optimize system performance, as explained in the previous section, we reprocessed the required data for training and testing. This adjustment significantly improved the accuracy rate, increasing it from 75.34% to 95.76%. In addition, the proposed approach demonstrates reliability and robustness in the face of changing topology, allowing for accurate fault location prediction. This resilience is crucial for network operators, considering the high likelihood of operational strategy adjustments within the network for various reasons. Given the numerous configurations possible in complex networks, it is essential that the proposed model remains robust against changes. Since the status of the switches (whether they are on or off) is known to the operator in practice, the operator can determine the current network topology based on the active switches. This knowledge allows the operator to select and run the trained model corresponding to the specific switch configuration. By applying the appropriate model, the operator can effectively detect any anomalies in the network, ensuring accurate fault localization despite changes in the network topology. This practical approach enhances the robustness of the proposed method as it leverages the operator's knowledge of the switch states to maintain high detection accuracy across different configurations.

3) MISSING SENSOR DATA

In previous sections, we evaluated the model under the assumption of complete sensor data from all bus and line measurements. In this section, we assess the model's performance when some measurement data is missing from the network. To achieve this, we randomly reduced the available measurements in three distinct scenarios to test both the proposed and the benchmark models. Table 6 shows that the proposed model is highly effective in locating faults even in the presence of missing measurement data. The table also highlights the resilience of the proposed method compared to the Graph Convolutional Network (GCN) across varying

TABLE 6. Performance metrics (in %) for GCN and proposed method with different missing sensor rates.

Method	Missing Rate	Accuracy	Precision	F1 Score (%)	Coverage	Recall
GCN	0	88.20	87.81	89.47	80.94	91.19
	25	82.35	81.50	83.49	75.12	84.85
	55	75.87	74.62	76.46	72.19	79.23
	65	70.96	68.50	71.65	65.78	72.00
Proposed Method	0	98.50	97.63	98.75	97.52	99.89
	25	92.63	91.76	92.19	91.50	95.32
	55	85.26	84.69	85.56	84.59	88.48
	65	80.35	80.23	81.14	80.50	82.45

levels of sensor data loss. We should note that only GCN is used as a benchmark because it consistently provides the best solutions compared to other models. Specifically, the proposed method maintains a high level of accuracy even when a significant portion of the sensors are unavailable. The robustness of the proposed method is primarily due to its use of the propagation matrix to update the adjacency matrix, which enhances the model's ability to leverage information from nodes or buses that are distant from the anomaly origin. By propagating information across the network, the proposed method effectively integrates data from non-adjacent nodes and lines that are correlated with the anomaly.

This integration significantly improves its ability to detect anomalies, even when direct measurements are lacking. As a result, the proposed method is able to maintain high performance and accurate anomaly detection, ensuring network observability despite substantial data loss.

4) CYBERSECURITY CONCERNS

It is essential to emphasize that a substantial amount of data unavailability, especially from sensor inputs, may signify cybersecurity threats. Our approach is designed to maintain strong performance even when sensor data are inadequate or potentially hacked, often indicating security breaches. Although our existing techniques guarantee operating efficiency regardless of these problems, we now lack the ability to determine whether the disruptions are attributable to cyber attacks. We intend to improve our system to detect abnormalities and accurately ascertain whether they originate from cybersecurity threats. This advancement will enable us to more precisely identify the origins of data anomalies, thus enhancing the security and dependability of smart grid operations.

V. CONCLUDING REMARKS AND FUTURE WORK

Anomaly detection and isolation continue to be critical challenges in power system networks. The introduction of smart meters and advanced measurement technologies has allowed power system operators to more accurately pinpoint fault locations, quickly address anomalies, and improve the overall reliability of power networks. However, complexities such as topology changes, the integration of renewable energy sources, stochastic propagation of anomalies, and missing data require approaches that are sensitive to these evolving conditions. In this study, we presented an innovative method that combines a Graph Convolutional Network (GCN) with a modified probability propagation matrix and augmented dual graphs to detect anomalies using network sensors placed on both nodes and lines. Also, we developed an optimization model that not only identifies the sources of anomalies, but also provides the probability that each node and line is the source of faults or impacted by them. We also propose a filtering strategy to further remove false positives based on the detected source of anomalies. This optimization model effectively locates the most probable fault sources and considers fault propagation throughout the network, improving the interpretability and accuracy of the results. The proposed method was tested on the IEEE 118-bus system, with simulation results demonstrating that it achieves reasonable performance and significantly improves the network's resilience to topology changes and missing data. By incorporating the optimization model, our approach provides more precise and consistent anomaly location performance, offering probabilities that help operators prioritize responses based on the likelihood that nodes or lines are faulty. Although the model is specifically designed for power networks, its adaptable features make it suitable for other sensor-intensive networks or graph structures, such as transportation or social networks, where anomaly detection at nodes and edges is essential. In future work, we plan to extend and validate the proposed model on other benchmark datasets from different domains to demonstrate its generalizability and broader applicability. The optimization framework can be applied to these domains to identify sources of anomalies and understand their propagation throughout the network. In future work, we will explore dynamic scenarios in which the network topology changes over time and consider the presence of noise in the sensor data. The proposed model paves the way for further research in developing customized GCN models that better understand the unique characteristics of the networks under study, making them more adaptable and interpretable. For future work, incorporating spatiotemporal elements into our GCN framework is a promising direction. This enhancement would allow the model not only to immediately locate faults, but also to anticipate potential problems based on historical and real-time data trends. Developing a hybrid model that combines the strengths of both GCN and spatiotemporal graph neural networks could cover a broader spectrum of needs, from rapid response to anticipatory fault management. This approach would significantly advance our capabilities in fault detection and contribute to more robust and resilient power system operations. Future work also includes improving model efficiency, particularly through vectorization and parallelization, to ensure the model remains feasible for large-scale deployment.

REFERENCES

[1] S. Borlase, "Smart Grids: Infrastructure, technology, and solutions," in *Electric Power and Energy Engineering*. Boca Raton, FL, USA: CRC Press, 2017.

[2] A. E. Labrador Rivas and T. Abrão, "Faults in smart grid systems: Monitoring, detection and classification," *Electr. Power Syst. Res.*, vol. 189, Dec. 2020, Art. no. 106602.

[3] J. De La Cruz, E. Gómez-Luna, M. Ali, J. C. Vasquez, and J. M. Guerrero, "Fault location for distribution smart grids: Literature overview, challenges, solutions, and future trends," *Energies*, vol. 16, no. 5, p. 2280, Feb. 2023.
[4] X. Ren, Y. Pan, M. Hou, R. Liang, L. Su, Q. Wang, and P. Zhang, "Fault location of the renewable energy sources connected distribution networks based on time differences of the modal traveling waves," *IEEE Access*, vol. 11, pp. 129671–129682, 2023.
[5] A. Al-Hinai, A. Karami-Horestani, and H. H. Alhelou, "A multi-objective optimal PMU placement considering fault-location topological observability of lengthy lines: A case study in OMAN grid," *Energy Rep.*, vol. 9, pp. 1113–1123, Dec. 2023.
[6] Y. Liu, D. Lu, S. Vasilev, B. Wang, D. Lu, and V. Terzija, "Model-based transmission line fault location methods: A review," *Int. J. Electr. Power Energy Syst.*, vol. 153, Nov. 2023, Art. no. 109321.
[7] K. Chen, C. Huang, and J. He, "Fault detection, classification and location for transmission lines and distribution systems: A review on the methods," *High Voltage*, vol. 1, no. 1, pp. 25–33, Apr. 2016.
[8] T. Takagi, Y. Yamakoshi, M. Yamaura, R. Kondow, and T. Matsushima, "Development of a new type fault locator using the one-terminal voltage and current data," *IEEE Power Eng. Rev.*, vols. PER–2, no. 8, pp. 59–60, Aug. 1982.
[9] F. H. Magnago and A. Abur, "Fault location using wavelets," *IEEE Trans. Power Del.*, vol. 13, no. 4, pp. 1475–1480, Apr. 1998.
[10] Z. Galijasevic and A. Abur, "Fault location using voltage measurements," *IEEE Power Eng. Rev.*, vol. 21, no. 11, p. 58, Nov. 2001.
[11] S. M. Brahma and A. A. Girgis, "Fault location on a transmission line using synchronized voltage measurements," *IEEE Trans. Power Del.*, vol. 19, no. 4, pp. 1619–1622, Oct. 2004.
[12] C.-W. Liu, K.-P. Lien, C.-S. Chen, and J.-A. Jiang, "A universal fault location technique for N-terminal (N = 3) transmission lines," *IEEE Trans. Power Del.*, vol. 23, no. 3, pp. 1366–1373, Jul. 2008.
[13] A. Gopalakrishnan, M. Kezunovic, S. M. McKenna, and D. M. Hamai, "Fault location using the distributed parameter transmission line model," *IEEE Trans. Power Del.*, vol. 15, no. 4, pp. 1169–1174, Apr. 2000.
[14] Q. Jiang, X. Li, B. Wang, and H. Wang, "PMU-based fault location using voltage measurements in large transmission networks," *IEEE Trans. Power Del.*, vol. 27, no. 3, pp. 1644–1652, Jul. 2012.
[15] 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.
[16] A. Anwar and A. N. Mahmood, "Anomaly detection in electric network database of smart grid: Graph matching approach," *Electr. Power Syst. Res.*, vol. 133, pp. 51–62, Apr. 2016.
[17] N. M. Khoa, M. V. Cuong, H. Q. Cuong, and N. T. T. Hieu, "Performance comparison of impedance-based fault location methods for transmission line," *Int. J. Electr. Electron. Eng. Telecommun.*, vol. 11, no. 3, pp. 234–241, 2022.
[18] H. Wei, W. Pengfei, Z. Haifeng, and D. Yuanjie, "Evaluation method for arrival time of traveling wave in fault location of distribution network," in *Proc. IEEE 6th Adv. Inf. Manage., Communicates, Electron. Autom. Control Conf. (IMCEC)*, May 2024, pp. 390–394.
[19] J. Yan, G. Song, Z. Chang, X. Gao, and C. Zhang, "Fault location method for MMC-HVDC systems based on numerical Laplace transform," *IEEE Trans. Power Del.*, vol. 39, no. 5, pp. 1–12, Oct. 2024.
[20] A. P. Alves da Silva, A. H. F. Insfran, P. M. da Silveira, and G. Lambert-Torres, "Neural networks for fault location in substations," *IEEE Trans. Power Del.*, vol. 11, no. 1, pp. 234–239, Jan. 1996.
[21] A. Yadav and Y. Dash, "An overview of transmission line protection by artificial neural network: Fault detection, fault classification, fault location, and fault direction discrimination," *Adv. Artif. Neural Syst.*, vol. 2014, pp. 1–20, Dec. 2014.
[22] Y. Zhao, Q. Liu, D. Li, D. Kang, Q. Lv, and L. Shang, "Hierarchical anomaly detection and multimodal classification in large-scale photovoltaic systems," *IEEE Trans. Sustain. Energy*, vol. 10, no. 3, pp. 1351–1361, Jul. 2019.
[23] C. Chen, Q. Li, L. Chen, Y. Liang, and H. Huang, "An improved GraphSAGE to detect power system anomaly based on time-neighbor feature," *Energy Rep.*, vol. 9, pp. 930–937, Mar. 2023.
[24] A. Gholami and A. K. Srivastava, "Comparative analysis of ML techniques for data-driven anomaly detection, classification and localization in distribution system," in *Proc. 52nd North Amer. Power Symp. (NAPS)*, Apr. 2021, pp. 1–6.
[25] O. A. Alimi, K. Ouahada, and A. M. Abu-Mahfouz, "A review of machine learning approaches to power system security and stability," *IEEE Access*, vol. 8, pp. 113512–113531, 2020.
[26] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang, "A review of graph neural networks and their applications in power systems," *J. Modern Power Syst. Clean Energy*, vol. 10, no. 2, pp. 345–360, Mar. 2022.
[27] A. Takiddin, R. Atat, M. Ismail, K. Davis, and E. Serpedin, "A graph neural network multi-task learning-based approach for detection and localization of cyberattacks in smart grids," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2023, pp. 1–5.
[28] E. Vincent, M. Korki, M. Seyedmahmoudian, A. Stojcevski, and S. Mekhilef, "Detection of false data injection attacks in cyber–physical systems using graph convolutional network," *Electric Power Syst. Res.*, vol. 217, Jun. 2023, Art. no. 109118.
[29] K. Chen, J. Hu, Y. Zhang, Z. Yu, and J. He, "Fault location in power distribution systems via deep graph convolutional networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 1, pp. 119–131, Jan. 2020.
[30] W. Khan and N. Ebrahim, "ANOGAT-Sparse-TL: A hybrid framework combining sparsification and graph attention for anomaly detection in attributed networks using the optimized loss function incorporating the Twersky loss for improved robustness," *Knowl.-Based Syst.*, vol. 311, Feb. 2025, Art. no. 113144.
[31] W. Khan, S. Abidin, M. Arif, M. Ishrat, M. Haleem, A. A. Shaikh, N. A. Farooqui, and S. M. Faisal, "Anomalous node detection in attributed social networks using dual variational autoencoder with generative adversarial networks," *Data Sci. Manage.*, vol. 7, no. 2, pp. 89–98, Jun. 2024.
[32] W. Khan, A. Mohd, M. Suaib, M. Ishrat, A. A. Shaikh, and S. M. Faisal, "Residual-enhanced graph convolutional networks with hypersphere mapping for anomaly detection in attributed networks," *Data Sci. Manage.*, vol. 8, no. 2, pp. 137–146, Jun. 2025.
[33] W. Khan, M. Ishrat, A. Neyaz Khan, M. Arif, A. A. Shaikh, M. M. Khubrani, S. Alam, M. Shuaib, and R. John, "Detecting anomalies in attributed networks through sparse canonical correlation analysis combined with random masking and padding," *IEEE Access*, vol. 12, pp. 65555–65567, 2024.
[34] Information Trust Institute. *IEEE 118-bus System*. Accessed: Sep. 2024. [Online]. Available: https://icsegiti.illinois.edu/ieee-118-bus-system/
[35] L. Liang, H. Zhang, S. Cao, X. Zhao, H. Li, and Z. Chen, "Fault location method for distribution networks based on graph attention networks," in *Proc. 3rd Int. Conf. Energy Electr. Power Syst. (ICEEPS)*, Jul. 2024, pp. 573–577.

MAHDI ZARIF (Member, IEEE) received the Ph.D. degree in electrical engineering from Ferdowsi University of Mashhad, in 2012. He is currently a Postdoctoral Researcher at the Institute for Sensing and Embedded Network Systems Engineering (I-SENSE), Florida Atlantic University, working on smart grids and smart cities. Previously, he was a Postdoctoral Associate at the University of Miami, for two years, focusing on energy analytics and sustainability. His research interests include power system analysis, renewable energy, smart grids, data analysis, deep learning, and energy management.

RAMIN MOGHADDASS (Member, IEEE) received the Ph.D. degree from the Department of Mechanical Engineering, University of Alberta, in 2008. He is currently an Associate Professor with the Department of Industrial and Systems Engineering, University of Miami. He is also the Director of the University of Miami Industrial Training and Assessment Center, which is supported by the Department of Energy. He is the Director of the University of Miami Institute for Clean Energy. He then spent two years with Massachusetts Institute of Technology as a Postdoctoral Research Fellow. His research focuses on data analytics, stochastic modeling, and decision making under uncertainty.