

Intelligent substation communication network fault location method based on dynamic spatiotemporal graph association perception

Ligang Ye^{1,*}, Wenzhang Li¹, Jing Zhao¹ and Yuanyuan Liu²

¹State Grid Inner Mongolia Eastern Power Co., LTD, Inner Mongolia ultra high pressure Branch Xilinhote, 026000, China

²State Grid Information & Telecommunication Group Co., LTD Beijing, China

Abstract

INTRODUCTION: Accurately locating faults in intelligent substation communication networks is crucial for power grid safety. Existing methods fail to fully capture dynamic fault characteristic evolution and complex dependencies within network topologies

OBJECTIVES: This paper aims to (1) model spatiotemporal fault features in communication networks, (2) enhance fault pattern capture through multi-view learning, and (3) improve fault location accuracy.

METHODS: We propose a multi-view spatiotemporal dynamic graph network. First, a multi-view graph neural network models spatial dependencies via cross-view comparative learning using topological and attribute data. Second, a gated recurrent unit with dynamic time windows extracts temporal evolution trends, focusing on local fault patterns and short-term dependencies.

RESULTS: Evaluations on a 220kV substation communication network show our method achieves higher fault location accuracy versus baselines, effectively capturing spatiotemporal fault characteristics.

CONCLUSION: The proposed framework addresses dynamic fault evolution and topological dependencies, providing a robust solution for intelligent substation fault diagnosis.

Keywords: Intelligent Substation, Communication Network, Fault Location, Dynamic Graph Neural Network, Spatiotemporal Modeling

Received on 12 February 2025, accepted on 20 July 2025, published on 10 October 2025

Copyright © 2025 L. Ye *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/ew.10423

*Corresponding author. Email: hellosgcc@qq.com

1. Introduction

The communication network of intelligent substations undertakes the responsibility of data transmission. By integrating various intelligent devices such as sensors, switches, and protection devices, it realizes the collection, monitoring, control, and information exchange of power communication data, which is crucial to ensuring the safe, stable, and efficient operation of the power grid [1,4]. The current intelligent substation communication network adopts a "three-layer two-network" topological structure, where nodes represent devices in the intelligent substation, such as intelligent terminals, merging units, switches, protection and measurement and control devices [2,6]. Edges represent the connection relationships between devices, i.e., network communication links, which can be fiber optic Ethernet or other forms of communication media for transmitting data and control signals [9]. When a fault occurs in the communication network, it generates massive alarm information, and accurately and quickly identifying the fault location from these data is a challenge.

Traditional fault location methods rely on the experience of operation and maintenance personnel, who judge based on message records and traffic monitoring in network packet analyzers. Reference [3] proposes a fault diagnosis method for generic object-oriented substation events (GOOSE), which manually builds a heuristic algorithm to decouple the fault characteristics of the intelligent substation packet loop analyzer, but the workload is huge and complex, making it difficult to determine the fault location in a short time. Reference [4] proposes a packet transmission path search algorithm to cross-locate the alarm information occurrence area on the communication link, which effectively improves the positioning efficiency of the potential fault area, but it relies on relatively single fault characteristic information and is difficult to accurately determine the exact fault location.

With the continuous development of the smart grid field, artificial intelligence technology has become an important force to solve the problem of fault location in intelligent substation communication networks. References [5,7] use deep belief networks to automatically learn the fault state characteristics of the original time-domain signals of the communication network, realizing fault location in intelligent substation communication networks and power grid fault type identification, respectively. Reference [8] first constructs a fault diagnosis model based on graph neural networks (GNNs) from a graph perspective, combining the topological mapping structure of the substation configuration description (SCD) with the information representation of device nodes, to capture the complex relationships between nodes in the intelligent substation

communication network, so as to effectively locate and identify faulty devices. Reference [10] proposes a fault location method based on graph filter neural networks, which enriches the representation form of fault characteristic information of different nodes through redundant detection of the fault state of the communication network.

Although these deep learning methods have improved the accuracy of fault location to a certain extent, they focus on modeling and analyzing the fault state of the communication network at a specific time point [9,24], ignoring the spatiotemporal evolution characteristics of communication network faults and the dynamic changes of fault states in the time dimension, resulting in the inability to identify the early signs of faults. Moreover, the intelligent substation communication network generates various types of data, including device status, communication link quality, alarm signals, etc. [12-13], which can reflect the operating status of the network from different perspectives. Currently, these fault location methods focus on single-view learning, limiting the understanding of data complexity and diversity, and thus unable to mine more comprehensive fault characteristic information.

Aiming at the problems existing in the current fault location methods for intelligent substation communication networks, this paper proposes a spatiotemporal graph association perception analysis method for intelligent substation communication network fault location to achieve real-time and accurate monitoring and early warning of faulty devices. The method first constructs three views based on the topological structure and characteristic attributes of the communication network to deeply explore the multi-view characteristics of faults in the communication network. Then, a multi-view graph neural network module is designed to deeply mine graph information through cross-view contrastive learning and capture the spatial dependencies of different views in the communication network. At the same time, a gated recurrent unit with a temporal dynamic window is proposed in the time dimension to capture the dynamic changes of communication network faults and adaptively assign weights to the importance of different time steps, thus effectively revealing the spatiotemporal evolution characteristics of faults and enhancing the ability to perceive early signs of faults. By fusing multi-view spatiotemporal information, this paper effectively improves the fault location effect, and the effectiveness of the method is verified through case studies.

2. Intelligent Substation Communication Network

The intelligent substation communication network

transmits information based on the "three-layer two-network" architecture. The three layers include the station control layer, bay layer, and process layer, and the two networks include the station control layer network and process layer network. The station control layer focuses on optimizing functional configuration and software applications to enable related devices to have more functions; the bay layer is responsible for changing the communication mode to improve the communication compatibility between devices; the process layer realizes collaboration through new devices and packet transmission; the station control layer network focuses on transmitting data between the station control layer and bay layer, providing two types of services and reports for corresponding data upload; the process layer network is unique to intelligent substations, transmitting packets for information interaction between process layer and bay layer devices. The clear layering and reasonable networking lay a solid foundation for efficient communication, functional collaboration, and safe and stable operation of devices in the intelligent substation.

The "three-layer two-network" architecture enables the intelligent substation communication system to connect a large number of secondary devices. These secondary intelligent devices can be divided into data sources, relay devices, destinations, and connection components according to their functions, and are distributed in the process layer, bay layer, and station control layer to complete information interaction of their respective functions, forming a complex graph-like topological structure, usually expressed as $G = (V, E, X)$, where $V = \{v_1, v_2, \dots, v_N\}$ represents the set of secondary device nodes, $E = \{e_1, e_2, \dots, e_L\}$ represents the set of links between secondary devices, and $X = \{x_1, x_2, \dots, x_N\}$ represents the status information of secondary devices, recording the switch status of merging units, protection devices, intelligent terminals, and measurement and control devices. The feature X of any device node in the graph can be defined as follows:

$$X = [X_{MU_1}, X_{MU_2}, \dots, X_{MU_n1}, X_{P_1}, X_{P_2}, \dots, X_{P_n2}, X_{IT_1}, X_{IT_2}, \dots, X_{IT_n3}, X_{MC_1}, X_{MC_2}, \dots, X_{MC_n4}] \quad (1)$$

n_1, n_2, n_3, n_4 are the total numbers of merging units, protection devices, intelligent terminals, and measurement and control devices connected to the device, respectively. $X_{MU}, X_P, X_{IT}, X_{MC}$ represent their transceiving packet information status. As shown in

Equation (1), when a communication fault occurs in the intelligent substation system, the operating status of the device and the network topology change, causing the power flow distribution in the communication network to change, and an alarm signal is issued simultaneously. Since when a device fails, the associated devices will all generate alarm signals, making it difficult to locate the faulty device. As shown in Figure 1, the optical fiber disconnection in the line bay causes local devices to generate alarm signals. Alarm signals are usually divided into switch quantity abnormal alarms, sampling abnormal alarms, and device abnormal alarms [20], where switch quantity abnormal alarms mainly include GOOSE-related alarm signals, sampling abnormal alarms mainly include sampled values (SV)-related alarm signals, and device abnormal alarms mainly include alarm signals when the device self-check is abnormal.

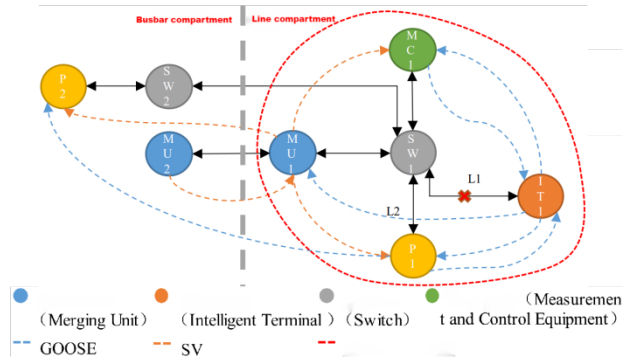


Figure 1. The schematic diagram of the range of communication fault alarm signals in substations

This paper uses graph neural networks (GNNs) to perform representation learning on the graph structure of the intelligent substation communication network. GNN is a deep learning model for processing graph structure data [11], which mainly uses the neighborhood message passing and aggregation mechanism to enable each node to collect information from its neighboring nodes and then update its own state. The specific learning process is shown in Equation (2):

$$h_i = \text{TRAN}(\text{AGG}(x_j | \forall j \in N(i))) \quad (2)$$

where $N(i)$ is the set of one-hop neighbor nodes of node i , h_i represents the potential representation of the i -th device node output by the GNN network, AGG is the aggregation operator for integrating the neighborhood information of the node, and TRAN is the nonlinear

parameterization of the transformation function.

3. Spatiotemporal Multi-View Dynamic Graph Model for Fault Location

3.1. Construction of Multi-View Communication Network Graph

Considering the complexity and diversity of the topological network structure constructed by communication device connections, single-view learning may lead to the omission of key fault discrimination information. To enhance the comprehensive understanding of the fault status and behavior of the communication network, this paper constructs an original view, an attribute view, and a global structure view.

Original View Construction: Since the SCD configuration file defines the physical connection relationship of substation communication devices, this paper parses the SCD configuration file to obtain the adjacency matrix of the communication network at the current moment according to the method in references [3,8]. The receiving and transmitting packet information of device ports can usually describe their operating conditions. This paper uses Equation (1) to encode the packet information of merging units, protection devices,

intelligent terminals, and measurement and control devices as the attribute features of the devices. X_i is the attribute feature of the i -th device node, and the calculation process is shown in Equation (3):

$$X_i = \text{Encoding} \left(\sum_{i=1}^{n_1} X_{MU_i} + \sum_{j=1}^{n_2} X_{P_j} + \sum_{m=1}^{n_3} X_{IT_m} + \sum_{k=1}^{n_4} X_{MC_k} \right) \quad (3)$$

Attribute View Construction: GNN requires that the input graph data follows the homogeneity assumption [14], that is, the attribute features of connected nodes in the graph are similar. However, communication faults sometimes cause the attribute differences between connected device nodes in the network. An attribute matrix X is used to construct a k -nearest neighbor graph to capture the attribute similarity relationship between device nodes, which is regarded as the attribute view. Suppose two device nodes V_i and V_j are given, and the cosine distance is used to measure their relationship in the attribute view, as shown in Equation (4):

$$A_{ij}^a = \begin{cases} 1 - \frac{X_i \cdot X_j}{\|X_i\| \|X_j\|} \geq \varepsilon \\ 0 \end{cases} \quad (4)$$

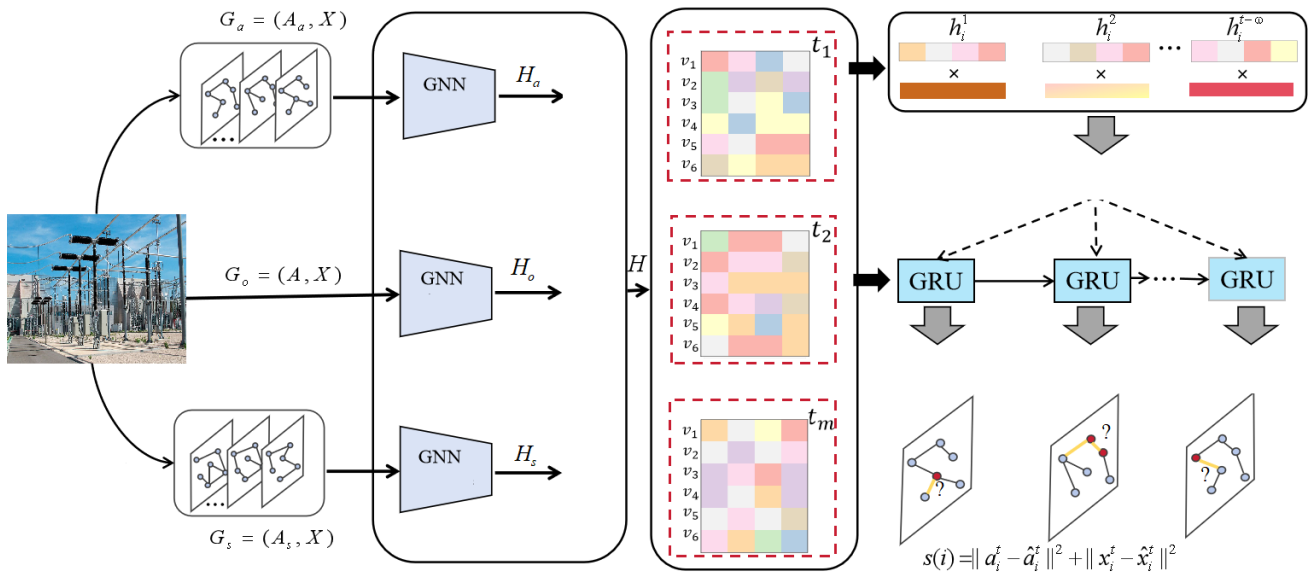


Figure 2. Overall framework for fault location of communication network in intelligent substation

where ε is a non-negative hyperparameter that can

control the number of neighbors of each node to ensure that the obtained new adjacency matrix satisfies sparsity.

Global Structure View Construction: The neighborhood aggregation and transmission mechanism adopted by graph network learning can only understand the low-order local connection relationship. Obtaining high-order node information requires providing a more comprehensive global perspective to help better understand the structure and function of the entire network and take measures in advance to prevent fault propagation. Although stacking more GNN layers can be used, it will cause the over-smoothing problem [16], that is, the originally distinguishable fault nodes will gradually absorb the characteristics of adjacent healthy nodes, resulting in smaller and smaller feature differences. To effectively capture global structure information, this paper uses a graph diffusion strategy as an enhanced method to obtain the global structure view, and the specific process is shown in Equation (5):

$$A^s = \sum_{k=0}^K \lambda(1-\lambda)^k A^k \quad (5)$$

A^k represents the k -th power of the adjacency matrix of the communication network graph, meaning that each node can capture information from k -hop neighbor devices, and λ is a hyperparameter coefficient used to control the diffusion amplitude of each step. This paper accumulates adjacency matrices of different orders, which can consider both close and distant neighbor nodes, thereby capturing multi-scale information and more comprehensively evaluating the propagation range and impact degree of faulty device nodes.

As shown in Figure 2, The original graph, attribute graph, and structure graph constructed are respectively marked as G_o , G_a , and G_s . At the same time, three independent graph encoders are constructed, and the constructed communication network views are input into the corresponding encoder modules for learning. Each encoder consists of two layers of GNN, and the specific process is shown in Equation (6):

$$\left. \begin{aligned} H_o &= A_o \sigma(A_o X W_o^1) W_o^2 \\ H_a &= A_a \sigma(A_a X W_a^1) W_a^2 \\ H_s &= A_s \sigma(A_s X W_s^1) W_s^2 \end{aligned} \right\} \quad (6)$$

where, H_o, H_a, H_s re the learned representations of communication nodes under three different views, σ represents the activation function, and W_o, W_a, W_s are the learnable weight parameter matrices in the three views, respectively. Considering that each view provides different factors for capturing faults in the substation communication network, their contributions to fault location may be different. This paper introduces an attention mechanism to adjust the weights of each view to balance their influences, ensuring that the model can

obtain the best fault location performance on different views. The final node representation of the substation communication network in the spatial dimension is shown in Equation (7):

$$\left. \begin{aligned} H &= \sum_{i=o,a,s} \gamma_i H_i, \gamma_i \geq 0 \\ \gamma_i &= \frac{\exp(\beta_i^T H_i + b_i)}{\sum_{j=o,a,s} \exp(\beta_j^T H_j + b_j)} \end{aligned} \right\} \quad (7)$$

where γ is the attention coefficient assigned to the representations of the three graph views, which is non-negative. β is a learnable vector used to assign appropriate weights to each element in the feature matrix, so that the model can capture the spatial information most relevant to the communication network fault nodes.

In the multi-view graph space of the substation communication network, it is necessary to measure the generality and difference of the representation of communication device nodes under different views. Contrastive learning is an unsupervised learning method [18] that only needs to use the structure and features of the data itself for learning, suitable for scenarios with lack of label data. This paper designs two contrastive loss functions using the idea of contrastive learning. One is the loss function based on information noise contrastive estimation (InfoNCE) [19], which ensures that the representations of the same node under different views are as close as possible; the other is the loss function based on triplets [21], which is used to enhance the difference between different views. The InfoNCE method constructs positive sample pairs (different views from the same node) and negative sample pairs (the same view from different nodes), allowing the model to learn universal feature representations. Given any device node V_i in the communication network, its InfoNCE loss under the original, attribute, and structure views is shown in Equation (8):

$$\begin{aligned} \ell_{INC}(V) = & -\log \frac{e^{\phi(H_o(V), H_a(V))/\tau}}{\sum_{n=1}^N e^{\phi(H_o(V_i), H_a(V_n))/\tau}} + \\ & -\log \frac{e^{\phi(H_o(V), H_s(V))/\tau}}{\sum_{n=1}^N e^{\phi(H_o(V_i), H_s(V_n))/\tau}} \end{aligned} \quad (8)$$

where $\phi(\cdot)$ is the cosine distance function for measuring the proximity of two node representations, N is the number of communication device nodes, and τ is the

temperature coefficient for adjusting the similarity distribution and convergence speed.

To enhance the difference between different views, this paper constructs a view triplet, taking the node of the original view as the anchor point, the attribute view as the positive sample, and the structure view as the negative sample, and the positions of the latter two can be swapped. The optimization goal of the view triplet loss function is to make the distance between the target node sample and the normal node sample as small as possible, while the distance between the target node sample and the fault node sample as large as possible, thereby increasing the difference between positive and negative samples. The triplet loss of any device node V_i in the given communication network can be obtained by Equation (9):

$$\ell_{Trip}(V) = \max(\underbrace{\varphi(H_o(V), H_a(V))}_{\varphi(H_o(V), H_s(V))} - \text{margin}, 0) \quad (9)$$

where margin is a hyperparameter used to control the distance interval between positive and negative views.

3.2. Temporal Dynamic Communication Network Learning

The temporal dynamic evolution process of the intelligent substation communication network can be described as a set of discrete graph sequences $\Phi = \{G^1, G^2, \dots, G^T\}$, where G^t represents the communication network topology graph at time t . Multi-view graph contrastive learning can only model the spatial dependencies of communication device nodes at a certain moment and cannot capture the dynamic characteristics of the communication network changing over time. To capture this dynamic change, this paper designs a gated recurrent unit (GRU) that fuses temporal context information. Through its gating mechanism, it effectively controls the flow of information and captures the long-term and short-term dependencies of communication device node representations in the time dimension. This paper constructs a temporal dynamic window to allow each node in the communication network to extract its short-term state information within a specific time period. The attention mechanism is used to weight the temporal context information of the local window, highlighting key information, thereby focusing on capturing local, rapidly changing fault patterns and short-term dependencies, as shown in Equation (10):

$$\left. \begin{aligned} c_{h,i}^t &= [h_i^{t-\omega}, \dots, h_i^{t-1}, h_i^t] \\ e_{h,i}^t &= r^T \tanh(q_h(c_{h,i}^t)^T) \\ \hat{h}_i^t &= \text{softmax}(e_{h,i}^t)(c_{h,i}^t)^T \end{aligned} \right\} \quad (10)$$

where h_i^t represents the spatial representation of the i -th device node at time t , ω is the window size, r and q_h are trainable parameters to optimize the attention weight scores of the window context, and finally, the temporal context information about the i -th device is integrated into $\hat{h}_i^t$, further enhancing the model's ability to capture short-term dependencies.

GRU is a common variant of recurrent neural networks, which performs well in modeling long-term dependencies through a unique gating mechanism. This paper uses GRU to integrate the historical state of nodes into the comprehensive representation at the current moment, retaining the long-term dependency information before the fault occurs, while quickly responding to the short-term changes when the fault occurs, thereby capturing the dynamic evolution trend of faults, as shown in Equation (11):

$$\left. \begin{aligned} P^t &= \sigma(W_p \cdot [Z^{t-1}, \hat{H}^t]) \\ R^t &= \sigma(W_r \cdot [Z^{t-1}, \hat{H}^t]) \\ \tilde{Z}^t &= \tanh(W_z \cdot [R^t \times Z^{t-1}, \hat{H}^t]) \\ Z^t &= (1 - P^t) \times Z^{t-1} + P^t \times \tilde{Z}^t \end{aligned} \right\} \quad (11)$$

where P^t and R^t are the outputs of the update gate and reset gate of GRU at time t , respectively, $\hat{H}^t$ is the input fusing the short-term context information at time t , and finally, the embedding representation Z^t at time t is obtained. This state can perceive the early sign features of faults, provide key information for fault location, and realize accurate prediction and timely response to faults.

Accurate Fault Location for Intelligent Substation Communication Network

The goal of fault location is to calculate the anomaly score of nodes in the dynamic graph; the higher the score, the greater the probability of a fault. Considering the high cost of obtaining fault labels, this paper designs an end-to-end training framework for unsupervised optimization, which learns the fault locator without any fault sample labels in the training set. The specific process is that after obtaining the embedding representations of graph nodes at each moment of the communication network dynamic graph, two fully connected layer-based decoders ξ_s and ξ_a are added to

reconstruct their topological structure and node attributes in the original view. The reconstruction loss is defined as shown in Equation (12):

$$\ell_{\text{Rec}} = \|A^t - \hat{A}^t\|^2 + \|X^t - \hat{X}^t\|^2 \quad (12)$$

where $\hat{A}^t = \sigma(M^t M^{tT})$ represents the reconstructed adjacency matrix at time t , $M^t = \xi_s(Z^t)$, and $\hat{X}^t = \xi_a(Z^t)$ is the reconstructed node attributes at time t . Finally, the overall loss function of this paper is defined as shown in Equation (13):

$$\ell_{\text{all}} = \ell_{\text{INC}} + \ell_{\text{Trip}} + \ell_{\text{Rec}} \quad (13)$$

For a device node in the intelligent substation communication network, if its original structure and attribute information can be approximated by the reconstruction decoder, the probability of it being faulty is very low. On the contrary, if the graph network pattern cannot be well reconstructed, it means that it does not conform to the pattern of most normal nodes, indicating that the node may have a fault. The node can be marked as a potential fault point for further inspection and maintenance. The fault locator in this paper uses the reconstruction error to calculate the anomaly score of each communication device node. The calculation process is shown in Equation (14):

$$s(V_i) = (1 - \alpha) \|a_i^t - \hat{a}_i^t\|^2 + \alpha \|x_i^t - \hat{x}_i^t\|^2 \quad (14)$$

where α is an important control parameter to balance the influence of structure reconstruction and attribute reconstruction. The algorithm flow of the fault location method proposed in this paper is summarized in Algorithm 1.

4. Experiments and Analysis

4.1. Experimental Setup

Taking the 220kV intelligent substation communication network as an example, this paper deeply analyzes the proposed communication network fault location method to verify the effectiveness of the proposed method. For the dataset, this paper collects the packet information flow of the communication network in the operating state of the intelligent substation every 30 minutes and parses the packet files to obtain the subscription relationship between devices, collecting about 20,000 pieces of temporal communication network data, recording the communication situation in the bus and transformer bay areas, among which 2,406 samples are labeled as fault types. 70% of the total data samples are used as the training set, 10% as the validation set, and 20% as the test set. To comprehensively evaluate the effect of the

proposed method, six baseline methods are selected for a comprehensive comparative study. The parameter settings of these baselines are as follows:

Table 1. Experimental Comparison Results 1

Time	Model	AUC	F1-score	GMean
T=30min	RNN	53.53	23.51	23.34
	LSTM	60.63	20.12	34.56
	DBN	69.35	56.52	33.93
	ProDNN	76.81	58.56	38.10
	SCGNN	83.39	61.63	43.82
	TADDY	86.22	69.27	50.04
	Ours	91.63	78.28	59.50
T=60min	RNN	35.64	14.46	10.86
	LSTM	41.60	18.34	12.53
	DBN	56.23	39.38	19.06
	ProDNN	60.11	55.73	26.27
	SCGNN	77.40	60.53	55.99
	TADDY	82.62	63.08	56.58
	Ours	89.24	74.53	62.81
T=120min	RNN	25.64	10.46	8.03
	LSTM	21.60	12.34	9.54
	DBN	39.23	19.38	11.89
	ProDNN	54.11	43.73	28.23
	SCGNN	62.40	50.53	39.66
	TADDY	70.62	59.08	43.30
	Ours	82.24	65.53	58.21

RNN: A unidirectional RNN model is adopted, with the number of model layers set to 2, the number of hidden layer neurons set to 48, no dropout is used to avoid excessive information loss, and the tanh activation function is used.

LSTM: A bidirectional LSTM model is adopted to extract bilateral information from the dynamic communication network time series. The number of model layers is set to 2, the number of hidden layer neurons is set to 96, the dropout value is set to 0.3, and the bias parameter is set to True.

DBN: Three layers of restricted Boltzmann machines are stacked, the number of neurons in the two hidden layers is set to 64 and 128, respectively, the learning rate is set to 0.001, and a momentum of 0.8 is introduced during training to accelerate the training speed and improve the stability of the model.

ProGNN: Two layers of graph neural networks are stacked, each layer of the graph network uses a mean aggregation function, the hidden layer parameter is 128, the learning rate decay range is set between 0.001 and 0.1, a dynamic learning rate scheduling strategy is adopted, the batch-size is set to 16, and SGD stochastic

gradient descent is used as the optimizer during training. SCGNN: High-order Chebyshev polynomials are used to approximate the graph convolution kernel, the order of each polynomial is set to 3, the number of filters is 36, the max-pooling mechanism is used in the convolution process, and L2 regularization is used to limit the size of parameters to prevent model overfitting.

TADDY: A 4-layer graph attention neural network is adopted, each layer of the attention network uses LSTM as the aggregation function to capture remote information, the dropout value is 0.5, the batch-size is set to 64, and the three hidden layer parameters are 256, 64, and 256 neurons, respectively.

Algorithm 1: Substation Communication Network Fault Location Algorithm Flow

Input: Dynamic graph training set

$\Phi = \{G^t\}_{t=1}^T$, Number of training epochs E ,

Neighbor sampling size for attribute view ε ,

Diffusion hops for structural view K ,

Temporal window size ω .

Output: Graph Node Fault Discrimination Score.

1. Random parameter initialization

2. **for** $epoch=1$ to E **do**:

3. **for** $G^t = (A^t, X^t) \in \{G^t\}_{t=1}^T$ **do**:

4. $G_o^t, G_a^t, G_s^t \leftarrow G^t$

5. $H^t \leftarrow (H_o^t, H_a^t, H_s^t)$, GNN Learn and fuse spatial representations from multi-view graphs

6. $\hat{H}^t \leftarrow [H^{t-\omega}, \dots, H^{t-1}, H^t]$, Capture short-term temporal information within window size using attention mechanisms

7. $Z^t \leftarrow [Z^{t-1}, \hat{H}^t]$, Utilize GRU units to capture both long- and short-term spatiotemporal historical information

8. $(\hat{A}^t, \hat{X}^t) \leftarrow \xi_s(Z^t), \xi_a(Z^t)$,

Decoding in both attribute and structural spaces

9. Compute fault scores for each graph node based on reconstruction tasks $s(i)$

10. Compute the loss function ℓ_{all} and backpropagate to update parameters

11. **end for**

12. **end for**

Table 2. Experimental Comparison Results 2

Size	Model	AUC	F1-score	GMean
5%	RNN	47.73	32.20	24.49
	LSTM	56.29	38.61	31.84
	DBN	61.69	42.30	35.22
	ProDNN	63.33	43.72	43.72
	SCGNN	69.57	47.25	46.71
	TADDY	73.29	52.48	51.06
	Ours	79.58	59.60	54.18
20%	RNN	50.29	35.56	29.47
	LSTM	62.77	44.92	30.57
	DBN	65.65	48.18	37.32
	ProDNN	74.22	55.19	55.19
	SCGNN	77.53	52.82	51.93
	TADDY	78.27	58.99	53.37
	Ours	82.24	60.32	56.80
40%	RNN	53.60	42.54	32.48
	LSTM	69.60	49.34	38.72
	DBN	74.67	51.56	42.24
	ProDNN	80.84	58.30	43.45
	SCGNN	83.48	60.53	49.11
	TADDY	85.46	59.38	59.38
	Ours	86.90	62.24	58.36

4.2. Experimental Results

Tables 1 and 2 provide a detailed comparison with the six baseline methods to show their fault location performance. Table 1 shows the fault prediction results of these baseline methods for predicting faults within the next 30 minutes, 60 minutes, and 120 minutes. Table 2 shows the fault prediction results of these baseline methods under 5%, 20%, and 40% training sets. Through in-depth comparative analysis of the experimental results of these methods, the following conclusions can be drawn: First, the method proposed in this study outperforms all baseline methods in the three main evaluation indicators of accuracy (5.92%), recall (3.62%), and F1 score (2.50%). This remarkable result not only proves the effectiveness of the method in fault location of substation communication networks but also highlights its potential and value in practical applications. This advantage may stem from the fact that the method can better capture and utilize the complex patterns and relationships in the data. Second, we note that recurrent neural networks (RNN) and long short-term memory networks (LSTM) perform the worst in fault location performance. After analysis, we believe this is mainly because they rely on time series data and can only extract time features, unable to effectively use the valuable

spatial information in the graph topology, which limits their performance in complex network fault prediction. In contrast, graph-based prediction methods such as deep belief networks (DBN) and property GNN (ProGNN), while considering the spatial dependencies between nodes, lose time information during processing. The loss of this information may lead to their inability to accurately capture the dynamic changes of fault occurrence, thereby affecting the accuracy of prediction.

Table 3. Multi view comparison results

View	AUC	F1-score	GMean
I	79.04	63.86	51.73
II	75.83	60.45	48.39
III	71.53	58.34	47.42
I+II	86.32	77.74	55.62
I+III	84.03	75.33	54.79
II+III	83.76	74.69	52.81
I+II+III	91.63	78.28	59.50

Finally, compared with self-attention coding GNN (SCGNN) and transformer anomaly detection dynamic graphs (TADDY), the method proposed in this paper performs well in long-distance prediction and location. This outstanding performance indicates that the method proposed in this paper can effectively combine time dimension and spatial topology information, which is particularly important for long-distance prediction. In summary, the experimental data in the two tables jointly support the experimental conclusions. Table 1 focuses on the time dimension, showing the fault prediction capabilities of each baseline method for short-term, medium-term, and long-term faults, allowing us to understand the performance of each method under different time spans. Table 2 starts from the proportion of the training set, reflecting the adaptability and performance changes of each method under different data volumes. The method of this study can organically combine the time dimension and spatial topology information, with obvious advantages in long-distance prediction and location, which is reflected in the long-term prediction of Table 1 and different data volumes of Table 2. It is analyzed that the performance of each baseline method is limited due to the defect in the use of spatiotemporal information, highlighting the advantage of this method in combining spatiotemporal information, which is crucial for timely responding to and handling faults in the substation communication network.

4.3. Performance Comparison of Spatiotemporal Graph Networks

To evaluate the specific contribution of different views to fault location performance, this paper constructs three views: original view (I), attribute view (II), and structure view (III), and sets up seven comparison groups. The experimental results are summarized in Table 3. When only any one of the three views is used, the accuracy, F1-score, and GMean of fault location are lower than the combination of two or three views. This finding emphasizes the importance of multi-view analysis in improving fault location performance. The original view performs better for certain types of faults because it directly presents the most authentic communication mode of the intelligent substation communication network, truthfully reflecting the path and frequency of information transmission between devices and the real-time operating status of devices. For faults directly related to the original communication link, such as faults caused by line connection interruption or communication protocol abnormality, the original view can quickly locate the fault location and accurately identify the fault type by virtue of its direct presentation of the original network status. The attribute view focuses on capturing the similarity of device attributes in the network. When faults related to device attributes occur, such as faults caused by device aging or performance mismatch, the attribute view can quickly find abnormal devices by comparing the differences between device attributes, thereby improving the fault identification ability. The structure view focuses on mining the dependency of the network structure, which can analyze the structural features such as the position of devices in the network topology and connection relationships. When faults are caused by unreasonable network structures, such as unbalanced loads or key node failures, the structure view can accurately identify the root cause of the fault based on its in-depth understanding of the network structure, thereby improving the identification accuracy of such faults. In addition, the experiment also found that different fault types have different sensitivities to view combinations, which provides a basis for designing more optimized view combinations for specific fault types in the future. By optimizing the view combination for different fault types, we can further improve the accuracy and efficiency of fault location. This experiment not only verifies the effectiveness of the multi-view graph network in fault location but also reveals the specific contribution of different views to fault location performance. These findings provide valuable guidance for future research and practice, especially in designing and optimizing fault location systems, on how to select appropriate view combinations according to specific fault types and network

characteristics. By comprehensively using different views, a more powerful and flexible fault location system can be constructed to cope with the increasingly complex fault detection challenges in intelligent substation communication networks.

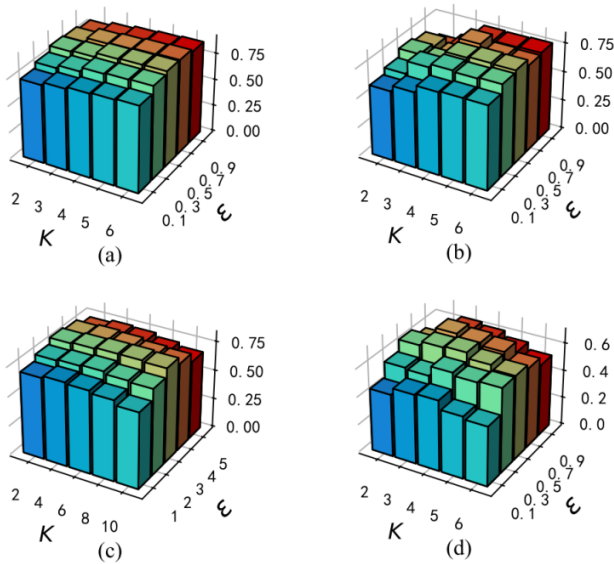


Figure 3. The relationship between K and ϵ

4.4. Analysis of Key Parameters for Multi-View Construction

This paper deeply discusses the specific impact of two key hyperparameters in constructing the attribute view and structure view on fault location performance. Through the experimental results shown in Figure 3, the influence of these parameters on the model performance can be intuitively observed. First, as the number of node neighbors decreases, the accuracy of fault location increases. The reason for this phenomenon is that reducing the number of neighbors helps faulty device nodes establish closer connections with nodes with similar attributes. This setting helps better follow the homogeneity assumption in the graph learning process, that is, similar nodes tend to be connected together, so that the spatial attribute relationship between device nodes can be more effectively captured. Second, the experimental results show that the accuracy of fault location increases with the increase of the diffusion distance. This indicates that the ability to obtain high-order node information is crucial for providing a more comprehensive global perspective, which helps better understand the structure and function of the entire communication network. In this way, the model can more

accurately identify and locate potential fault points. However, when the diffusion distance exceeds 4, the performance improvement slows down. This is because although the neighborhood aggregation mechanism of the graph neural network enables each communication node in the substation to diffuse and transmit information to other communication nodes through edges, to a certain extent, the increase of the diffusion distance allows the node to collect more abundant and diverse information. However, when the diffusion distance is excessively increased, a series of negative effects will be brought. As the diffusion distance continues to increase, a large amount of information not directly related to the core representation of the target communication node will also pour in. In the actual operation scenario of the substation, the processing capacity of each communication node is limited. The input of too much non-key information will cause a sharp increase in the processing burden of the node, and a large amount of computing resources will be consumed in processing irrelevant information, resulting in a significant decrease in information processing efficiency. Moreover, these redundant information may also interfere with the model's accurate capture of key information. In applications such as substation fault diagnosis, the model needs to accurately identify the information propagation path and characteristics related to faults. However, the noise of too much irrelevant information will cover up the truly valuable fault signals, making it difficult for the model to accurately determine the root cause and propagation range of the fault, thus affecting the accuracy and timeliness of fault diagnosis. In addition, excessive information diffusion may also destroy the originally reasonable information hierarchy in the graph neural network, making it difficult to distinguish the importance of node information at different levels, and finally leading to the degradation of model performance, unable to give full play to the advantages of the graph neural network in the substation communication network. In summary, the adjustment of these two hyperparameters has a significant impact on the performance of fault location, which not only improves the accuracy of fault detection but also enhances the adaptability and robustness of the model to complex network environments.

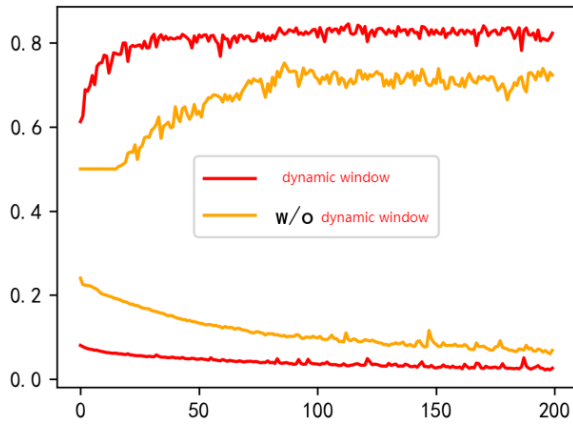


Figure 4. The Function of temporal dynamic window

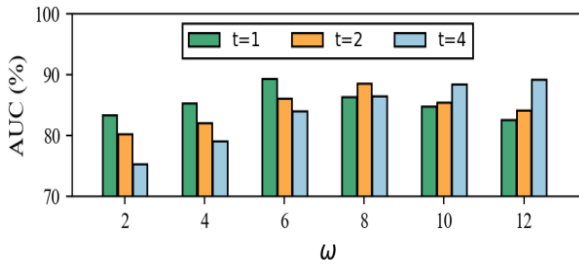


Figure 5. The influence of window size

4.5. Analysis of Temporal Dynamic Network

To study the impact of the temporal dynamic network on capturing the time-dimensional information of communication device node representations, especially the role of the dynamic window in fault location performance. First, we studied whether the increase of the dynamic window affects the fault location performance. To evaluate this, we plotted the training accuracy and loss with and without the temporal dynamic window, as shown in Figure 4. From the trend of these curves, it can be seen that after introducing the temporal dynamic window, the training accuracy of the model has been more significantly improved, and the loss decreases faster. This phenomenon proves that the temporal dynamic window can effectively focus on capturing local, rapidly changing fault patterns and short-term dependencies. In this way, the model can more quickly learn the dynamic changes when a fault occurs, thereby improving the accuracy of fault detection. Second, we analyzed the impact of the dynamic window size on fault

prediction and location accuracy through Figure 5. Based on the experimental results, we can conclude that when the dynamic window is small, the accuracy of the model in predicting long-distance faults decreases. This may be because the small window cannot fully capture the complete characteristics of the fault pattern, resulting in the model losing important time series information during the learning process. The loss of this information may affect the model's ability to understand and predict the changes before and after the fault occurs. However, when the dynamic window size increases to a certain extent, the accuracy rebounds. This is because the larger window provides more time series information, helping the model better understand the long-term dependencies of the fault pattern.

5. Conclusions

Aiming at the problem of fault location in intelligent substation communication networks, this paper proposes a fault location method for intelligent substation communication networks based on a multi-view spatiotemporal dynamic graph network, which overcomes the insufficient utilization of the dynamic changes of fault characteristics over time and the effective capture of complex dependencies in the communication network topology. The main work and conclusions are as follows:

Starting from the topological structure and characteristic attributes of the intelligent substation communication network, three views are constructed respectively to deeply explore the multi-view characteristics of faults in the communication network and enhance the comprehensive understanding of the fault status and behavior of the communication network.

A multi-view spatial graph neural network is constructed in the spatial dimension, and cross-view contrastive learning is used to model the spatial dependencies of the communication network. Through the detailed extraction and deep fusion of features from different views, the complex spatial dependencies between nodes in the communication network are accurately captured, which not only effectively enhances the differences between views but also fully excavates the complementarity, realizes the complementary advantages of information, and comprehensively improves the ability to understand and represent the network spatial structure.

In the time dimension, the gated recurrent unit and self-attention mechanism are combined to accurately control the flow direction and intensity of temporal context information, thereby effectively capturing the long-term and short-term dependencies of communication device node representations in the time dimension, providing key information for fault location,

and realizing accurate prediction and timely response to faults.

This paper aims to reveal the spatiotemporal evolution characteristics of intelligent substation communication network faults and verifies that fusing multi-view spatiotemporal information can better enhance the ability to perceive early signs of faults, making the fault location accuracy reach 91.3%. At present, fault location mainly focuses on common types, and the ability to handle complex faults is still insufficient. In the future, the characteristics of complex faults will be deeply studied, the model architecture will be improved, and the ability to locate complex faults will be enhanced.

Acknowledgements

This work was supported by State Grid Inner Mongolia Eastern Power Co., LTD under grant No.SGMDJX00JJS1800357.

References

- [1] Li H, Zhang X, Pan H, et al. Research on the reliability of communication networks in intelligent substations. *Power Syst Prot Control*. 2021;49(9):165-171.
- [2] Wang Y, Zhang W, Wang C, et al. Numerical simulation research on electromagnetic response of grounding grid in intelligent substation based on multiple grid division technology. *Electr Meas Instrum*. 2024;61(12):97-103.
- [3] Zhou H, Huang J, Zhang C, et al. GOOSE loop fault diagnosis in intelligent substations based on Petri nets. *South Power Syst Technol*. 2017;11(6):8-16.
- [4] Zhang Y, Cai Z, Long P, et al. Real-time fault diagnosis model and method for communication networks in intelligent substations. *Power Syst Technol*. 2016;40(6):7-14.
- [5] Yang W, Yin K, Bao Y, et al. Power grid fault type identification based on deep belief networks. *Electr Power Eng Technol*. 2021;40(2):169-177.
- [6] Zhou L, Zhou Y, Shen W, et al. Reactive power and voltage optimization control for new power systems based on deep reinforcement learning. *Electr Meas Instrum*. 2024;61(9):182-189.
- [7] Ren B, Li J, Zheng Y, et al. Research on fault location of process-level communication networks in smart substations based on deep neural networks. *IEEE Access*. 2020;8:109707-109718.
- [8] Zhang C, Zheng Y, Lu J, et al. Research on fault location of secondary circuits in intelligent substations based on graph neural networks. *Power Syst Prot Control*. 2022;50(11):10-20.
- [9] Lin Y, Chen N, Cai J. Security control technology for distribution network dispatching data based on multi-source information fusion. *Electr Meas Instrum*. 2024;61(6):118-132.
- [10] Zou D, Hong Z, Dai W, et al. Transformer short-circuit current calculation method based on graph neural networks. *J Glob Energy Interconnect*. 2024;7(3):303-311.
- [11] Wu Z, Pan S, Chen F, et al. A comprehensive survey on graph neural networks. *IEEE Trans Neural Netw Learn Syst*. 2020;32(1):4-24.
- [12] Feng X, Feng X, Zhang Z, et al. Measurement error detection method for intelligent electricity meters based on maximum likelihood method and decision tree. *Electr Meas Instrum*. 2024;61(12):205-211.
- [13] Qi F, Yao W, Zhang C, et al. Fault diagnosis of power communication networks based on convolutional neural networks. *Inf Technol*. 2023;(6):119-128.
- [14] Chen D, Lin Y, Li W, et al. Measuring and relieving the over-smoothing problem for graph neural networks from the topological view. *Proc AAAI Conf Artif Intell*. 2020;34(4):3438-3445.
- [15] Wang H, Chen P, Li N, et al. Bi-level optimal operation of virtual power plants considering multi-type reserve of source-load. *Electr Meas Instrum*. 2024;61(11):31-39.
- [16] Tang H, Liang X, Guo Y, et al. Diffuse and Smooth: Beyond Truncated Receptive Field for Scalable and Adaptive Graph Representation Learning. *ACM Trans Knowl Discov Data*. 2023;17(5):1-25.
- [17] Zhang Q. Research on fault point location method for wired communication networks based on verification twin neural networks. *Yangtze Inf Commun*. 2024;37(8):169-171.
- [18] Lin S, Zhou P, Hu Z, et al. Prototypical graph contrastive learning. *IEEE Trans Neural Netw Learn Syst*. 2022;35(2):2747-2758.
- [19] Xu D, Cheng W, Luo D, et al. Infogcl: Information-aware graph contrastive learning. *Adv Neural Inf Process Syst*. 2021;34:30414-30425.
- [20] Zhu W, Hu Y, Nie Y, et al. Research on power grid false data intrusion detection method based on deep machine learning. *Electr Meas Instrum*. 2022;61(10):67-73.
- [21] Zhou H, Qin Q, Hou J, et al. Deep global semantic structure-preserving hashing via corrective triplet loss for remote sensing image retrieval. *Expert Syst Appl*. 2024;238:122105.
- [22] Aimit A, Bai X, Jiang D, et al. Research on power communication network fault identification and location technology based on Beidou and two-stage. *Artif Intell Sci Eng*. 2024;(3):75-83.
- [23] Cai L, Chen Z, Luo C, et al. Structural temporal graph neural networks for anomaly detection in dynamic graphs. *Proc 30th ACM Int Conf Inf Knowl Manag*. 2021:3747-3756.
- [24] Zhu W, Jiang W, Zhou Z, et al. Research on network security situation awareness method based on power grid dispatching system. *Electr Meas Instrum*. 2024;61(7):21-27.
- [25] Liu Y, Pan S, Wang Y, et al. Anomaly detection in dynamic graphs via transformer. *IEEE Trans Knowl Data Eng*. 2021;35(12):12081-12094.