Renewable Energy and Power Quality Journal
https://repqj.com/index.php/repqj/index
RE&PQJ, Vol. 23, No. 2, 2025

Intelligent Location of Fault Sections in Power Transmission and Transformation Systems Using Knowledge Graph and GraphSAGE

Ming Song[1], Yi Zhou[2], Kangwei Li[1], Haitao Jiang[1], Hua Wen[3,*]

[1]State Grid Corporation of China's Ultra High Voltage Construction Branch, Beijing, 102401, China
[2]Beijing Power Transmission & Transformation Corporation, Beijing, 102401, Chin
[3]Huada Tianyuan(Beijing) Science and Technology Co., Ltd., Beijing, 102401, China
[*]Corresponding author's email: wenhuaa_a@sina.com

Abstract. To solve the problem of insufficient feature representation ability and poor topological adaptability of traditional methods in fault location of complex power transmission and transformation systems, this study proposes an intelligent location framework that combines knowledge graph with GraphSAGE. Based on a four-layer architecture, a dynamic knowledge graph is constructed in real time through SCADA/PMU data, defining entities such as substations, lines, circuit breakers and their electrical connections, protection triggers and other relationships, and designing an incremental graph update mechanism. The model layer uses a double-layer GraphSAGE, combined with an LSTM aggregator to perform stratified sampling and dynamic edge weight adjustment on neighborhood time series measurement features. Joint cross-entropy and knowledge constraint loss are introduced in training to balance classification error and semantic consistency. The experiment generates 2000 fault cases based on the IEEE39 node system. The results show that the positioning accuracy of the proposed model in the IEEE 39-node system reaches 0.953. As the transition resistance gradually increases from 200 Ω to 2000 Ω, the accuracy of the proposed model changes from 0.953 to 0.943. As the signal-to-noise ratio changes from 5dB to 30dB, the accuracy changes from 0.920 to 0.955.. The semantic enhancement of the knowledge graph and the dynamic topological reasoning mechanism of GraphSAGE synergistically improve the robustness and real-time performance of the model, providing a high-precision solution for the intelligent operation and maintenance of the power transmission and transformation system.

Key words. Power Transmission and Transformation System, Fault Section, Intelligent Positioning, Knowledge Graph, GraphSAGE

1. Introduction

As the core hub of the power grid, the operation safety and stability of the power transmission and transformation system [1,2] directly affect the reliability of power transmission. However, the network topology [3,4] is complex, the equipment types are diverse, and the fault propagation paths are changeable, which makes the traditional positioning methods based on protection action or waveform analysis have obvious deficiencies in feature characterization and topology adaptability. Knowledge graphs [5,6] integrate heterogeneous data such as equipment archives, topological structures, and historical faults into structured entity-relationship networks through semantic modeling and reasoning; GraphSAGE (Graph Sample and Aggregate) [7,8] relies on neighborhood sampling and aggregation mechanisms to achieve efficient inductive embedding learning of large-scale, dynamic networks. The collaborative application of the two can enhance node features through knowledge embedding [9,10], and maintain the model's reasoning efficiency and generalization ability in real-time switch state changes, providing an innovative solution for high-precision, dynamic and adaptive positioning of fault sections in power transmission and transformation systems.

The power transmission and transformation system studied in this article mainly includes ultra-high voltage and ultra-high voltage AC/DC transmission systems, covering the main grid structure and key substation nodes. This type of system has the characteristics of complex topology, variable operating conditions, and wide range of fault impact. This study is based on the actual operation and maintenance needs of the State Grid Corporation of China's ultra-high voltage construction project, aiming to enhance its intelligent identification ability for fault sections under complex operating conditions, support rapid isolation and restoration of power supply, and ensure the safe and stable operation of the power grid.

The main contribution of this study is to propose an intelligent positioning framework for fault sections in power transmission and transformation systems that integrates knowledge graphs and GraphSAGE, achieving seamless integration and dynamic collaboration from semantic modeling to graph neural reasoning. Through

the entity-relationship-attribute definition and incremental update mechanism based on the IEC 61850 standard, a multi-source heterogeneous knowledge graph including substations, lines, circuit breakers, relay protection devices and historical fault events was constructed. It can not only support the semantic reasoning of complex topologies and causal links, but also provide rich prior constraints for subsequent graph neural networks. Inductive GraphSAGE is introduced at the model level to perform multimodal feature fusion on current amplitude, phase and switch state signals. Real-time topology changes are embedded in the network representation through dynamic edge weight adjustment based on SCADA (Supervisory Control And Data Acquisition)/PMU (Phasor Measurement Unit), which significantly improves the discrimination of fault embedding and the model's adaptability to topology mutations. PMU is used for high-precision synchronous acquisition of phasor data of voltage and current in power systems, providing real-time dynamic behavior monitoring. This paper designs a hybrid loss function that combines cross entropy and knowledge consistency regularization to maintain tight clustering of semantically similar nodes in the embedding space, enhancing the interpretability and generalization ability of the model. In the IEEE 39-node system and multiple fault scenarios (including high-impedance grounding and Emanuel arc faults), this framework outperforms mainstream methods such as HGNN, GAT, GIN (Graph Isomorphism Network) and GCN (Graph Convolutional Network) in terms of positioning accuracy, robustness and response delay. This fully verifies the synergistic advantages of knowledge graph semantic enhancement and GraphSAGE efficient embedding learning, and provides an innovative and feasible technical path for the intelligent operation and maintenance of power transmission and transformation systems.

2. Related Works

In recent years, knowledge graphs [11,12] have gradually become an important supporting tool for the intelligentization of power systems due to their semantic modeling and relational reasoning capabilities. In the field of asset management, researchers use equipment entities such as substations, lines, and protection devices as nodes, and construct associations such as "connected to" and "protection coordination" to achieve unified management and retrieval of power grid topology [13,14] and equipment attributes. The fault diagnosis method based on knowledge graph [15,16] integrates real-time measurement data and static equipment archives in the same graph, and uses path query or rule reasoning to quickly identify possible fault sources and propagation paths. In addition, many works also integrate ontology models and machine learning algorithms [17] on top of knowledge graphs to improve the generalization ability of new fault modes. However, existing research focuses on static graph construction and offline reasoning, and lacks support for real-time topology changes and dynamic data updates, making it difficult to meet the online needs of high-frequency fault location in power transmission and transformation systems.

Graph neural networks are widely used in power grid state estimation and fault location because they can directly learn features on graph structures. Early studies mostly used GCN [18,19] to extract neighborhood features of network nodes to achieve classification and discrimination of fault types; subsequently, GAT [20,21] enhanced the focus on key fault propagation paths by adaptively allocating neighbor weights. In view of the large-scale topology and real-time requirements of power grids, inductive methods such as GraphSAGE [22] effectively reduce the computational complexity through neighborhood sampling and learnable aggregation functions, and support online embedding updates of network nodes. In addition, some works combine time-series graph neural networks [23,24] to integrate time-series measurement information such as current and voltage to achieve dynamic capture of fault evolution. Although the above methods have achieved significant improvements in positioning accuracy and robustness, most studies are still limited to pure data-driven, lacking the use of complex semantic relationships in power grids and model interpretability.

Although knowledge graphs and graph neural networks have shown great potential in the field of power systems [25,26], the deep integration of the two has not yet been fully explored. Existing knowledge graph research [27,28] focuses on static entity-relationship modeling and lacks effective connection with graph neural network training; while most graph neural network methods are based on topological structure [29,30] and measurement features, and make insufficient use of device attribute semantics and historical fault causal information. In addition, in the case of dynamic topological changes [31,32], how to achieve the coordination of real-time graph update and model online reasoning to ensure the rapid adaptation of the positioning model to new fault nodes is still a key problem that needs to be solved. In recent years, research on load fault localization has focused on integrating knowledge graphs and graph neural networks to improve localization accuracy and robustness. Based on GraphSAGE [33,34], GAT and other models, combined with real-time data and power grid topology structure, dynamic fault inference is achieved. At the same time, introducing knowledge to enhance semantic expression, integrating historical faults and device attribute information, improving the stability of positioning in complex scenarios, and providing new ideas for smart grid fault diagnosis. Therefore, combining the semantic enhancement capability of knowledge graph with the efficient embedding learning mechanism of GraphSAGE to build a fault location framework with dynamic topology adaptation and semantic reasoning capabilities has become an important breakthrough in current research.

3. Methods

A. Overall Framework Design

The overall framework design in this section is specifically divided into four layers: data layer, knowledge layer, model layer and application layer. It

takes the continuous input of SCADA/PMU data as the starting point, and finally outputs the fault section location result through dynamic update of knowledge graph and GraphSAGE reasoning. The overall framework is shown in Figure 1.

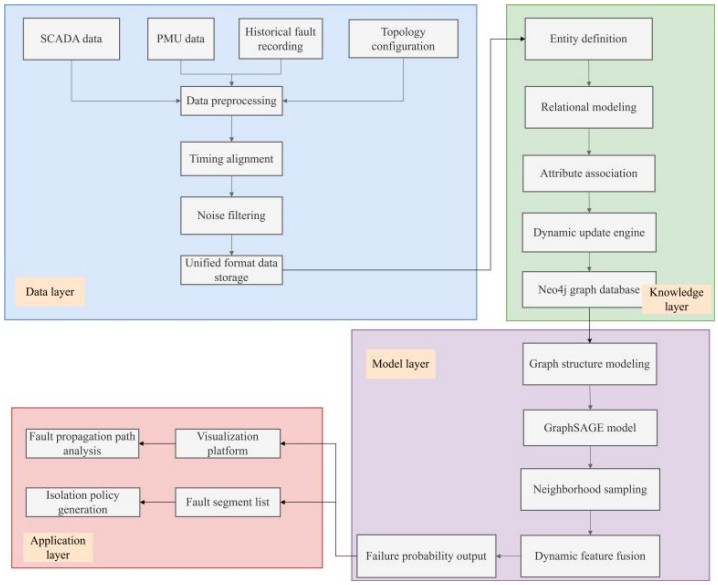

Figure 1. Overall framework of this paper.

As the underlying support of the system, the data layer is mainly responsible for collecting and preprocessing various operating data, including real-time monitoring values such as switch status, measurement quantity, voltage and current provided by SCADA, and high-precision phasor synchronization data collected by PMU. In addition, historical fault recordings, equipment archives, and topology configuration files are integrated. All raw data are cleaned, time-series aligned, missing value interpolated, and noise filtered to generate unified format data storage that can be called by the upper layer, ensuring the input quality of subsequent knowledge graph construction and model training.

The knowledge layer builds and maintains the power grid knowledge graph based on the output of the data layer. Entity types (substations, line sections, circuit breakers, protection devices, fault events, etc.) can be defined, and relationship models can be designed according to the IEC 61850 standard. The attribute layer associates the device electrical parameters, operating status, and historical fault labels. To meet the needs of online positioning, an incremental graph update mechanism is designed. Whenever SCADA/PMU reports a switch trip or reconfiguration of the topology, the system automatically triggers subgraph updates and knowledge reasoning to ensure that the graph always reflects the latest state of the power grid.

The model layer is centered on GraphSAGE, which carries the fault node representation learning and reasoning functions. Through the neighborhood sampling strategy, a multi-layer neighbor subset is extracted from each target node, and the aggregator is used to combine the electrical measurement timing characteristics and the device state vector to generate semantically enhanced node embedding. The dynamic edge weight mechanism adjusts the connectivity strength in the graph according to the real-time switching signal to ensure that the model can capture the impact of topological mutations on the fault propagation path. During the training process, the joint cross entropy and knowledge constraint loss function are used to minimize the positioning error and semantic reasoning error at the same time.

The application layer converts the model reasoning results into visualization and operation and maintenance decision support. The positioning module outputs the probability distribution map of the fault section and integrates with the geographic information system platform to intuitively mark the fault location and possible propagation path, and provides a list of fault sections for on-duty engineers to dispatch isolation and maintenance. The system supports minute-level online updates and feedback to meet the actual power grid's needs for efficient and accurate fault diagnosis.

Through the above four-layer collaborative design and flow, the framework realizes an end-to-end closed loop from massive multi-source data to high-precision fault section positioning, providing solid technical support for the intelligent operation and maintenance of the power transmission and transformation system.

B. Knowledge Graph Construction

In this study, the knowledge graph construction follows the IEC 61850 standard to meet the semantic fusion requirements of multi-source heterogeneous data in the power transmission and transformation system. The construction process includes three steps: entity definition, relationship modeling, and attribute association. The equipment archives, topology information, and historical fault records are uniformly mapped to a structured semantic network, providing rich context and constraints for the representation learning of

graph neural networks.

(1) The entity definition phase identifies and abstracts the key object types in the power transmission and transformation system, including but not limited to the following:

Substation is the core node responsible for voltage level conversion and power distribution, usually including multiple incoming and outgoing lines and main transformers;

Transmission Line (TransmissionLine) is the conductor segment that carries high-voltage power for long-distance transmission, and its attributes include impedance, length and geographic coordinates;

Circuit Breaker(CircuitBreaker) is the actuator for fault isolation and line protection, and its action sequence is crucial to fault location;

Protection Relay (ProtectionRelay) monitors current and voltage changes and sends a trip command when an abnormality is detected;

Current and voltage sensors (PMU/CT/PT) provide synchronized phasors and measurement data;

Fault events (FaultEvent) record the time, type and maintenance results of the fault;

Geographic location (Location) and operation and maintenance personnel (Operator) and other auxiliary entities.

(2) In the relationship modeling stage, various association types are designed according to electrical semantics and operation and maintenance logic:

connected_to indicates the electrical connection between the two, such as the connection between the line and the substation;

triggers: the action trigger of the circuit breaker by the relay protection;

measured_by: PMU/CT/PT measurement association with line or bus;

located_in: the geographical location or section of the equipment;

has_fault_history: the historical association between the entity and the fault event;

adjacent_to describes the topologically adjacent line or switchgear;

protected_by: the line or transformer is protected by a certain relay protection device.

(3) Attribute association level, adding static and dynamic attributes to each entity, including:

Equipment parameters include line impedance (Ω/km), transformer capacity (MVA), circuit breaker rated current (A);

Status quantity includes voltage amplitude, phase angle and switch status reported in real time by SCADA/PMU;

Historical fault records fault number, occurrence timestamp, fault type (short circuit, grounding, arc, etc.) and maintenance time;

Spatial information includes longitude and latitude coordinates, region;

Maintenance cycle is date of the most recent maintenance and inspection.

Finally, a power knowledge graph containing thousands of entities, tens of thousands of relationships, and hundreds of dimensional attributes is formed. It not only supports path query and rule reasoning, but also provides semantic initialization and constraint information for subsequent GraphSAGE embedding learning.

The entity relationship example is shown in Table 1.

Table 1. Entity relationship example.

Entity 1	Relation	Entity 2	Description
Substation-A	connected_to	TransmissionLine-L1	Substation A is electrically connected to line L1
TransmissionLine-L1	measured_by	PMU-Device-P1	PMU P1 performs synchronized phasor measurements on line L1
CircuitBreaker-CB1	located_in	Substation-A	Circuit breaker CB1 is installed in substation A
ProtectionRelay-R1	triggers	CircuitBreaker-CB1	Relay protection R1 operates to trigger circuit breaker CB1
TransmissionLine-L1	adjacent_to	TransmissionLine-L2	Lines L1 and L2 are adjacent in the same corridor
Transformer-T1	protected_by	ProtectionRelay-R2	Transformer T1 is protected by R2
Substation-B	connected_to	TransmissionLine-L2	Substation B is electrically connected to line L2
FaultEvent-F1001	has_fault_history	TransmissionLine-L1	Fault event F1001 occurs on line L1
CircuitBreaker-CB2	measured_by	CT-Device-C1	CT C1 measures bus current near circuit breaker CB2
PMU-Device-P2	located_in	Substation-B	PMU P2 is deployed in substation B
Substation-A	owned_by	Operator-O1	Substation A is managed by operator O1
TransmissionLine-L3	connected_to	Transformer-T2	Line L3 is connected to the high voltage side of T2

C. *Fault Reasoning Model Based on GraphSAGE*

The GraphSAGE model used in this article belongs to a type of graph neural network, designed to address the problem of node embedding learning in large-scale graph data, especially in scenarios that require support for inductive learning and online updates. Compared to traditional GCN, which is mainly used for node classification tasks in static graphs, GraphSAGE emphasizes efficient subgraph embedding learning through neighborhood sampling and aggregation mechanisms, and is suitable for power transmission and transformation systems with frequent topology changes. In addition, GAT enhances selective attention to key neighborhood information by introducing attention mechanisms, while this paper combines LSTM aggregator and dynamic edge weighting mechanism to further improve the adaptability of the model in complex fault scenarios of power systems.

In order to construct a graph model for fault inference, the entity set involved in this study is first defined, including all key devices involved in power grid operation and fault propagation. This collection contains the following main types:

Substation: the core node responsible for voltage conversion and energy distribution functions;

Transmission line: a physical channel for long-distance transmission of high-voltage electrical energy;

Circuit breaker: a switching device that performs fault isolation and line control;

Protection relay: Monitor abnormal current and voltage and issue tripping instructions;

Measurement equipment: provides real-time synchronized phasor and status signals;

Historical fault events: Record the location, type, and handling results of past faults.

The above entities form a dynamically evolving network of relationships through electrical connections, triggering actions, measurement associations, and other relationships.

The power transmission and transformation system is abstracted as a weighted graph at time t, and the formula is:

$$G_t = (V, E, W_t) \quad (1)$$

In Formula 1, V is a node set, each node corresponds to an electrical device (such as busbar, line section, circuit breaker, protection relay, etc.). $E \subseteq V \times V$ is an edge set, which represents the electrical connectivity between devices. $W_t = \left[w_{ij}^t \right]$ is a time-varying edge weight matrix, which is used to characterize the impact of switch status on connectivity. $\mathrm{V} = \{1, 2, \cdots, 39\}$ represents a set of nodes, corresponding to the bus numbers in the IEEE 39 node system.

The constructed graph structure is shown in Figure 2.

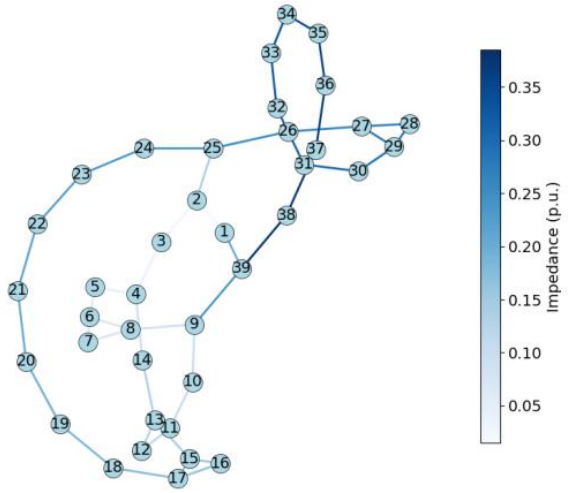

Figure 2. Constructed graph structure.

Through Matplotlib's Blues color spectrum and Normalize mapping, the impedance of each edge is represented by color depth: the greater the impedance, the darker the edge color; the smaller the impedance, the lighter the edge color. Nodes are represented by numbers 1-39 and are directly used for input preprocessing and feature construction of the GNN model.

For any edge $(i, j) \in E$, its weight is defined as:

$$w_{ij}^t = \begin{cases} 1, & \text{if switch}_{ij} \text{ is closed at } t \\ \text{ , } & \text{if switch}_{ij} \text{ is open at } t, \ll 1 \\ 0, & \text{if } (i, j) \notin E. \end{cases} \quad (2)$$

This design ensures that fault information is mainly propagated along closed lines, while weak connectivity is retained on disconnected lines to assist the model in learning recursive backtracking of edge faults. Index i

and j are elements of node set V.

From the matrix interpretation, define the adjacency matrix A_t and the degree matrix D_t:

$$A_t = \left[w_{ij}^t \right] \quad (3)$$

$$D_t = \mathrm{diag}\left(\sum_j w_{1j}^t, \cdots, \sum_j w_{|V|j}^t \right) \quad (4)$$

Each node v collects three-phase voltage and current synchronous phasors at time t, and its initial eigenvector contains:

$$\left[V_a^t, V_b^t, V_c^t, I_a^t, I_b^t, I_c^t, \phi_a^t, \phi_b^t, \phi_c^t \right]^\top \quad (5)$$

Among them, V_* and I_* represent the amplitude, and ϕ_* is the phase angle. In order to capture short-term evolution, the most recent T time steps are selected and spliced to form a time series feature:

$$\mathrm{x}_v^{\mathrm{seq}} = \left[\mathrm{x}_v^{t-T+1} \| \cdots \| \mathrm{x}_v^t \right] \in \mathbb{R}^{9T} \quad (6)$$

The operating status of circuit breakers and protective devices is represented by binary features:

$$s_{\mathrm{CB}}^t, s_{\mathrm{Relay}}^t \in \{0,1\} \quad (7)$$

1 represents "action" or "closed", and 0 represents "no action" or "open". These discrete signals are concatenated with continuous measurements to form the complete feature vector of the node:

$$\mathrm{x}_v^t = \left[\mathrm{x}_v^{\mathrm{seq}} \| s_{\mathrm{CB}}^t \| s_{\mathrm{Relay}}^t \right] \in \mathbb{R}^{9T+2} \quad (8)$$

This model uses a two-layer GraphSAGE architecture, and the hierarchical aggregation mechanism includes neighborhood sampling and LSTM (Long Short-Term Memory) aggregation. Each layer samples a fixed number of domain $\mathcal{N}(v) = \{u : (u,v) \in E\}$ of the target node v:

$$\mathcal{S}_k(v) \sim \mathrm{Uniform}\left(\mathcal{N}(v), S \right), S = 10 \quad (9)$$

This operation reduces computational complexity at each layer and prevents over-smoothing.

In LSTM aggregation, the sampled neighbor embedding sequence $\left\{ \mathrm{h}_u^{(k-1)} : u \in \mathcal{S}_k(v) \right\}$ is input into the LSTM unit in a fixed order:

$$\mathrm{h}_{\mathcal{N}(v)}^{(k)} = \mathrm{LSTM}\left(\mathrm{h}_{u_1}^{(k-1)}, \cdots, \mathrm{h}_{u_S}^{(k-1)} \right) \quad (10)$$

The initial layer is $\mathrm{h}_v^{(0)} = \mathrm{x}_v^t$. LSTM aggregation can learn the complex dependencies of neighbor states over time.

The update formula for each layer is:

$$\mathrm{h}_v^{(k)} = \sigma\left(W^{(k)} \left[h_v^{(k-1)} \| \mathrm{h}_{\mathcal{N}(v)}^{(k)} \right] + b^{(k)} \right) \quad (11)$$

In formula 11, $\|$ represents concatenation and $\left\{ W^{(k)}, b^{(k)} \right\}$ is a trainable parameter.

Finally, $\mathrm{h}_v^{(k)}$ is embedded and mapped to the fault category probability through the Softmax classifier:

$$P\left(y_v = c \left| \mathrm{h}_v^{(K)} \right. \right) = \frac{\exp\left(w_c^\top \mathrm{h}_v^{(K)} + b_c \right)}{\sum_{j=1}^{C} \exp\left(w_j^\top \mathrm{h}_v^{(K)} + b_j \right)} \quad (12)$$

In formula 12, C is the total number of categories and $\left\{ \mathrm{w}_j, b_j \right\}$ is the classification parameter.

Multi-classification cross entropy loss is used:

$$\mathcal{L}_{\mathrm{CE}} = -\sum_{v \in V} \sum_{c=1}^{C} \mathbb{I}\left(y_v = c \right) \log P\left(y_v = c \right) \quad (13)$$

The hyperparameter settings of the GraphSAGE model are shown in Table 2.

Table 2. GraphSAGE model hyperparameters.

Parameter	Value	Parameter	Value
Number of layers	2	Batch size	128
Neighbors per layer (Layer 1)	10	Dropout rate	0.5
Neighbors per layer (Layer 2)	10	Weight decay (L2 regularization)	0.0001
Aggregator	LSTM	Training epochs	100
Initial learning rate	0.001	Knowledge loss balance coefficient	0.1

The two-layer structure (Number of Layers=2) can take into account both local and distant neighborhood features. Each layer samples 10 neighbors to balance efficiency and information volume. The LSTM aggregator can capture time series measurement dependencies. The learning rate, batch size, weight decay, and Dropout

jointly control training stability and generalization capabilities. 100 training rounds can fully converge, and weekly incremental updates ensure that the model responds to topology changes in a timely manner. The knowledge graph loss coefficient is used to balance classification and semantic consistency constraints.

D. Dynamic Topology Reasoning Mechanism

In actual power grid operation, the topology structure changes frequently with the switch action, and must be captured and synchronized to the reasoning model in real time to ensure the timeliness and accuracy of fault location. The system continuously monitors the action signals of switches, circuit breakers and reclosers provided by SCADA. When a device state $s_{ij}(t)$ switches from "closed" to "open" or vice versa, the subgraph update is triggered.

Whenever $\Delta s_{ij}(t) = s_{ij}(t) - s_{ij}(t-1) \neq 0$ is detected, the system only performs subgraph reconstruction on the affected local node set and injects the updated edge weight information into the knowledge graph and GNN input graph.

To avoid full retraining every time the topology changes, this study uses incremental parameter fine-tuning. Assume that the model parameters are Θ and the current training loss is:

$$\mathcal{L}(\Theta) = \mathcal{L}_{\mathrm{CE}}(\Theta) + \lambda \mathcal{L}_{\mathrm{K}}(\Theta) \quad (14)$$

In formula 14, $\mathcal{L}_{\mathrm{CE}}$ is the cross entropy classification loss, and $\mathcal{L}_{\mathrm{K}}$ is the knowledge consistency regularization. In formula (14), a composite objective function is defined that includes cross entropy loss and knowledge consistency regularization terms to guide the initial training process of the model. In order to achieve rapid adaptation after topology changes, an incremental fine-tuning strategy based on small batch gradient descent is adopted. Every week or when the cumulative number of key topology changes exceeds the preset threshold, the system collects new fault samples and the latest topology data to form an incremental training set $\mathcal{D}_{\mathrm{inc}}$. Fine-tune Θ through mini-batch gradient descent:

$$\Theta \leftarrow \Theta - \eta \nabla_{\Theta} \left[\mathcal{L}(\Theta; \mathcal{D}_{\mathrm{inc}}) \right] \quad (15)$$

The learning rate η takes a small value (1/10 of the initial value) to keep the model stable. In this way, the model can quickly adapt to changes in topology and fault mode, while avoiding the high computational overhead of full training. The incremental update strategy uses logging and version control to ensure that when the positioning error increases, it can quickly roll back to the historical optimal weight, thereby ensuring high availability of the online fault location system. This incremental update mechanism only adjusts local parameters, significantly reducing the computational cost of full retraining. At the same time, combined with logging and version control, it supports rapid rollback to historical optimal weights when positioning performance declines, ensuring high availability of the system.

4. Experimental Design

A. Experimental Setup

In this study, the experimental platform selected the classic New England IEEE 39-bus system, and carried out detailed modeling and parameter calibration, including core information such as bus voltage level, line impedance and generator output. After the system topology was reproduced in the Python environment, 2000 groups of typical fault cases were generated through PSCAD software simulation, including single-phase grounding fault, two-phase short circuit fault, three-phase short circuit fault, etc. In order to investigate the actual robustness of the positioning model, white noise is added to the voltage and current measurement waveforms in each fault case, and all measurement data are normalized and time-series calibrated in the post-processing stage to ensure the consistency and comparability of the input features.

The Bus voltages of the IEEE (Institute of Electrical and Electronics Engineers) 39-node system are shown in Figure 3.

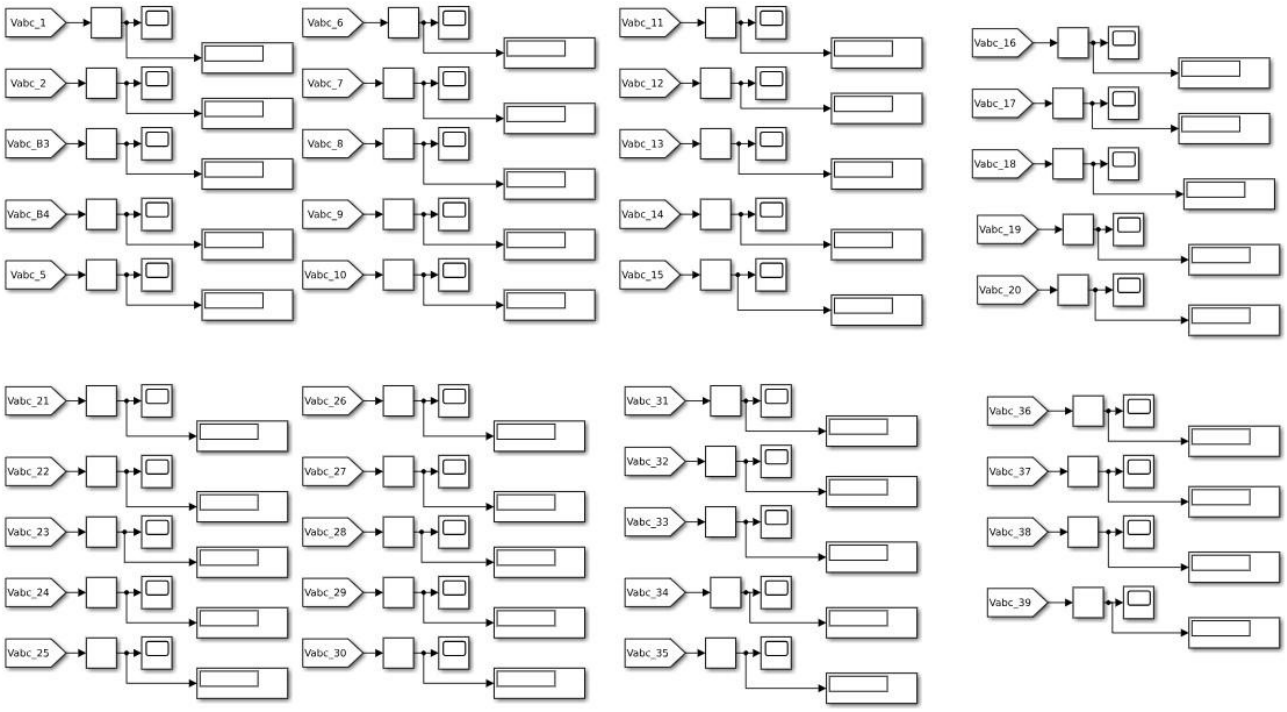

Figure 3. Bus voltages of IEEE 39-bus system.

The IEEE 39-bus system has a base voltage of 345 kV and contains 39 nodes. 1-29 are load nodes (Pd=6.15 GW, Qd=1.40 GVar), 30-38 are generator nodes (Pg=6.20 GW, V=1.05 pu), node 39 is a balancing node (V=1.03 pu), and the per-unit voltage range of the entire network is 0.94–1.06.

In the dynamic adaptability experiment, the transition resistance of the line single-phase grounding fault was tested at equal intervals in the IEEE 39-bus system, with resistance values of 200 Ω, 400 Ω, 600 Ω, 800 Ω, 1000 Ω, 1200 Ω, 1400 Ω, 1600 Ω, 1800 Ω and 2000 Ω. Under each impedance condition, 200 groups of fault cases were generated based on PSCAD, keeping other fault parameters unchanged, and evaluating the location accuracy and delay change of the model in high impedance fault scenarios. By comparing the accuracy drop trend and response delay increment of the proposed model with HGNN, GAT, GIN, and GCN at different impedance levels, the sensitivity and adaptation speed of each method to the change of fault resistance characteristics are quantified.

In the evaluation of the sampling delay impact, the neighborhood sampling lag caused by network communication or computing delay during the real-time reasoning process of GraphSAGE is simulated, and the number of nodes for delayed sampling is gradually increased from 1 to 10, and the positioning performance under each delay configuration is tested respectively. In the noise robustness test, Gaussian white noise of different intensities was added to the PSCAD simulation data, and the signal-to-noise ratio was set to 5 dB, 10 dB, 15 dB, 20 dB, 25 dB and 30 dB, a total of 6 levels, to verify the positioning accuracy of each model when the measurement noise increased sharply.

In order to evaluate the model's scalability for complex arc faults, the single-phase grounding fault was replaced with the Emanuel type arc fault, verifying the wide applicability and engineering feasibility of the method.

B. Evaluation Indicators

In terms of comparative evaluation, the model in this paper is tested in parallel with four mainstream graph neural network models: HGNN, GAT, GIN and GCN. All models are trained and tested under the same hardware environment (Intel Xeon CPU, NVIDIA Tesla V100 GPU, 128 GB memory) and software configuration (Python 3.8, PyTorch 1.12), and positioning accuracy, F1-score and average positioning delay are used as the main evaluation indicators to ensure the fairness and scientificity of the comparison results.

The positioning accuracy is defined as the proportion of all fault segments in the test set that the model correctly predicts:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

In formula 16, TP is the number of faulty segments correctly identified, TN is the number of non-faulty segments correctly excluded, FP is the number of normal segments mistakenly predicted as faulty, and FN is the number of faulty segments missed.

F1-score comprehensively measures the accuracy and recall of the positioning model on unbalanced data, and its calculation formula is:

$$F1 = 2 \cdot \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (17)$$

The response time measures the delay from the occurrence of a fault to the output of the positioning result by the model, including: data transmission delay, graph update and feature construction delay, graph neural network forward reasoning delay, result visualization and decision support output delay.

5. Results and Discussion

A. Comparison of Accuracy

In the IEEE 39-node system, the positioning accuracy and F1 value information of the proposed model and the comparison model are shown in Figure 4.

The detailed data of Figure 4 is shown in Table 3.

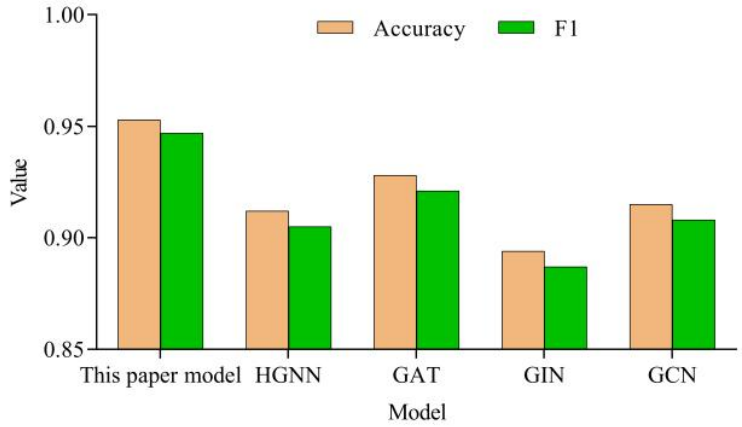

Figure 4. Positioning accuracy.

Table 3. Positioning accuracyy.

Model	Accuracy	F1
This paper model	0.953	0.947
HGNN	0.912	0.905
GAT	0.928	0.921
GIN	0.894	0.887
GCN	0.915	0.908
GCN+Rule Engine	0.93	0.925
GAT+Expert System	0.935	0.928
GraphSAGE+Knowledge Enhancement	0.94	0.936
Based on impedance method	0.86	0.852
travelling wave method	0.875	0.867
SVM+Feature Engineering	0.88	0.874
RF+timing characteristics	0.89	0.883
LSTM+attention mechanism	0.90	0.895

The "knowledge graph + GraphSAGE" model proposed in this paper has a positioning accuracy of 0.953 and an F1 value of 0.947 on the IEEE 39-node system, which is significantly better than the four mainstream comparison methods. Although the HGNN, GCN and GAT models can utilize graph structure information, they have limitations in fault semantics and dynamic topology adaptability, resulting in accuracy rates of 0.912, 0.915 and 0.928 respectively, and F1 values of 0.905, 0.908 and 0.921 respectively. GIN performs slightly worse in extracting structural isomorphic features, with an accuracy of only 0.894 and an F1 of 0.887. In contrast, the model in this paper integrates the semantic priors of electrical equipment and historical faults through the knowledge graph, and then combines it with the inductive embedding learning of GraphSAGE to effectively enhance the discrimination of node feature expression. In particular, in scenarios where the fault sections have similar semantics or complex topology, it can more accurately distinguish small feature differences and eliminate noise interference, thus achieving significant improvements.

An in-depth analysis of the root causes of the performance differences can be attributed to the following aspects: First, the knowledge graph provides the model with multi-dimensional semantic relationships including "connected_to" and "triggers", which makes up for the lack of understanding of device attributes and causal links in models based purely on measurement or topology; second, GraphSAGE uses neighborhood sampling and LSTM aggregation. This retains local high-order structural information and also forms a stronger timing capture capability for multi-time measurement features and switch state changes; thirdly, the dynamic edge weight mechanism enables the model to respond to SCADA/PMU switch changes in real time,

ensuring that fault location can be completed quickly and stably in topology mutation scenarios. In addition, the optimization strategy of combining cross entropy and knowledge consistency loss further improves the interpretability and generalization ability by constraining the distance of semantically similar nodes in the embedding space. Therefore, the proposed method is superior to other graph neural network variants in positioning accuracy, providing stronger technical support for the intelligent operation and maintenance of power transmission and transformation systems.

The visualization of node embedding t-SNE (t-distributed Stochastic Neighbor Embedding) is shown in Figure 5.

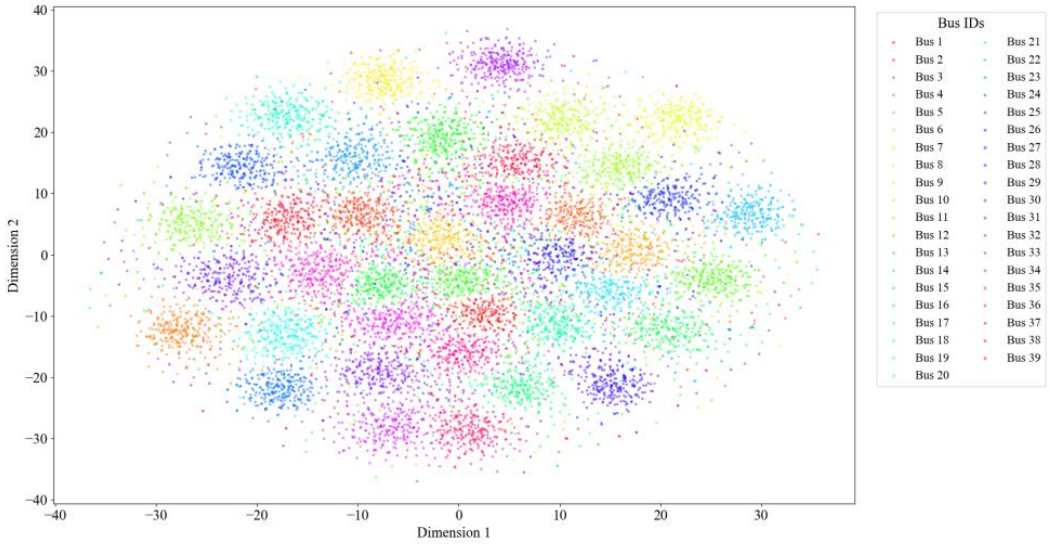

Figure 5. Node embedding t-SNE visualization.

From the t-SNE visualization results, the node embeddings of the 39 buses roughly form 39 relatively clustered but overlapping clusters in two-dimensional space, reflecting the model's ability to distinguish different bus characteristics and capture potential correlations. The points within the cluster are relatively scattered, indicating that GraphSAGE retains the diversity of time series measurements and device states during neighborhood sampling and LSTM aggregation; the partial overlap between clusters reflects the similarity of adjacent buses in the power grid in topology or fault propagation paths. The fault segment nodes are densely distributed near the center of their clusters, which supports the interpretability of the location model's embedding of the fault segment - the model can project nodes of the same fault type to similar areas.

B. Ablation Experiment

The ablation experiment was set up and the results are shown in Table 4.

Table 4. Ablation experiment results.

Model	Accuracy	F1
This paper model	0.953	0.947
Remove knowledge graph	0.926	0.919
Further remove dynamic topology reasoning mechanism	0.908	0.902

The ablation experiment results show that the knowledge graph is crucial to improving the performance of the fault localization model. When the complete framework is retained, the model accuracy reaches 0.953 and F1 is 0.947; once the knowledge graph is removed, the accuracy drops to 0.926 and F1 is 0.919. This significant performance gap is mainly due to the unique advantages of knowledge graphs in enhancing node representation. By encoding multi-source semantic information such as electrical equipment attributes, historical fault events, and protection coordination relationships into entity embedding, the model can better distinguish segments that are similar on the surface but have different potential fault mechanisms. Especially when the fault signal is weak or the noise interference is large, the semantic prior provides a powerful compensation for the node characteristics, thereby achieving more robust positioning when the original measurement data is not sufficient to fully reveal the fault characteristics. In addition, the causal links in the knowledge graph also provide constraints on the upstream and downstream fault propagation paths for the model, making the probability distribution of the fault segment more concentrated and the decision boundary clearer in the complex topology structure.

After further removing the dynamic topology reasoning mechanism, the model accuracy and F1 dropped to 0.908 and 0.902 respectively, which is lower than the configuration of removing only the knowledge graph.

This result highlights the key role of the topology dynamic adaptive module in fault location: the operating state of the power transmission and transformation system changes from time to time, and events such as circuit breaker opening and closing and load switching can change the propagation path and amplitude attenuation characteristics of the fault current. The dynamic edge weight mechanism injects topology change information into the graph structure by real-time monitoring of the switching signals of SCADA/PMU, allowing GraphSAGE to capture the latest channels for fault propagation. After removing this mechanism, the model can only rely on static topology and historical semantics, which makes it difficult to accurately reflect the actual electrical connectivity state when the fault occurs, resulting in increased positioning errors. In summary, the synergy of knowledge graph and dynamic topology reasoning mechanism jointly supports the high accuracy and robustness of the model in a complex power grid environment. The loss of any link can significantly weaken the overall performance.

C. Dynamic Adaptability

In the IEEE 39-node system, the transition resistance when the line grounding fault occurs is increased, and the influence of the transition resistance on the fault location is analyzed. The results are shown in Figure 6.

The data in Figure 6 is shown in Table 5.

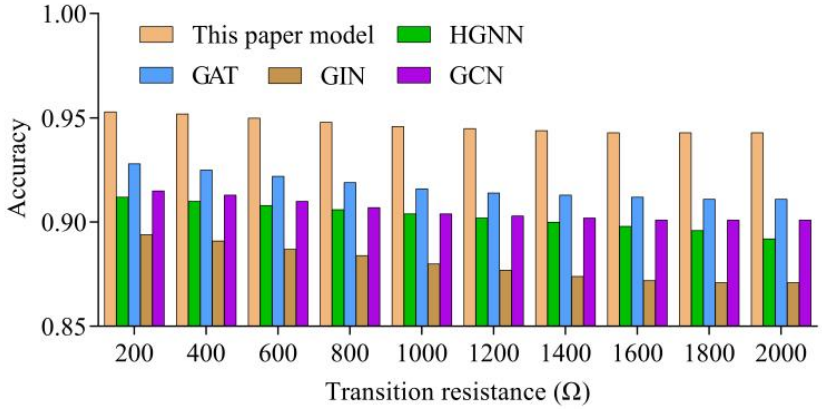

Figure 6. Effect of transition resistance on fault location.

Table 5. Effect of transition resistance on fault location.

Transition resistance (Ω)	This paper model	HGNN	GAT	GIN	GCN
200	0.953	0.912	0.928	0.894	0.915
400	0.952	0.91	0.925	0.891	0.913
600	0.95	0.908	0.922	0.887	0.91
800	0.948	0.906	0.919	0.884	0.907
1000	0.946	0.904	0.916	0.88	0.904
1200	0.945	0.902	0.914	0.877	0.903
1400	0.944	0.9	0.913	0.874	0.902
1600	0.943	0.898	0.912	0.872	0.901
1800	0.943	0.896	0.911	0.871	0.901
2000	0.943	0.892	0.911	0.871	0.901

As the transition resistance gradually increases from 200 Ω to 2000 Ω, the location accuracy of each model shows a downward trend. This is mainly because too high resistance can significantly attenuate the amplitude and waveform characteristics of the fault current, making the difference between the fault and normal state in the measurement signal blurred, reducing the discrimination ability of the model based on time series measurement data. Comparing the decline of each model, it can be clearly seen that the "knowledge graph + GraphSAGE" framework proposed in this paper is the most robust: the accuracy only dropped from 0.953 to 0.943, which is lower than HGNN, GAT, GIN, and GCN. This shows that the semantic priors of device attributes and historical faults provided by the knowledge graph can still provide compensation for node embedding through semantic reasoning when the measurement information decays, thereby maintaining higher positioning accuracy. In addition, GraphSAGE's dynamic edge weight mechanism can capture the real-time impact of switch and resistance changes on fault propagation paths, enhancing the model's ability to perceive resistance changes.

Further analysis shows that different models have different sensitivities to resistance increases. Although HGNN introduces a hypergraph structure to fuse multi-source heterogeneous information, it has a strong reliance on static high-order relationships and is difficult to reflect the topological impact of resistance changes in real time, so the accuracy drops from 0.912 to 0.892. GIN has limited effect in characterizing homogeneous structures, with the lowest overall performance and the largest drop, indicating that pure structural features are difficult to provide sufficient information for fault location under current decay conditions. Although GCN has stable low-frequency properties, its limitations in high-impedance scenarios also lead to performance degradation. In summary, the semantically enhanced embedding of the model in this paper is combined with dynamic adaptive topological reasoning, which can effectively compensate for weakened measurement signals and adjust the neighborhood information aggregation strategy in real time, thereby maintaining high positioning accuracy and stability in high-resistance scenarios.

D. Effect of Sampling Delay

The fault location accuracy of delayed sampling nodes is shown in Table 6.

Table 6. Effect of sampling delay.

Number of delayed sampling nodes	This paper model	HGNN	GAT	GIN	GCN
1	0.953	0.912	0.928	0.894	0.915
2	0.950	0.910	0.925	0.891	0.913
3	0.947	0.908	0.922	0.888	0.910
4	0.944	0.905	0.919	0.885	0.907
5	0.940	0.903	0.916	0.882	0.904
6	0.937	0.900	0.914	0.879	0.902
7	0.933	0.897	0.911	0.876	0.899
8	0.930	0.894	0.908	0.873	0.896
9	0.927	0.890	0.905	0.870	0.893
10	0.925	0.887	0.902	0.867	0.890

As the number of delayed sampling nodes increases from 1 to 10, the positioning accuracy of each model shows a decreasing trend. Among them, the "knowledge graph + GraphSAGE" model proposed in this paper drops from 0.953 to 0.925, which is always higher than other models. This result reflects the impact of the real-time topology information lag introduced by delayed sampling on the fault location performance. The greater the delay, the more significant the deviation between the graph structure based on which the model aggregates the neighborhood and the actual fault propagation channel, which makes the node embedding unable to accurately capture the spatial distribution and propagation path of the instantaneous fault signal.

For mainstream graph neural network methods, HGNN, GAT, and GCN, they have a strong dependence on topological consistency. Once the neighborhood information lags, there is a significant performance decline even at lower delays (2-4 nodes). Although GIN has advantages in isomorphic graph representation, due to the lack of adaptive adjustment of dynamic topological weights, its tolerance to delayed sampling is slightly inferior, and the accuracy finally drops to 0.867. In comparison, the advantage of this model lies in the integration of knowledge graph semantic priors and inductive aggregation of GraphSAGE. Through the knowledge graph, rich device attributes and historical fault causal constraints are introduced between graph nodes, so that even when the sampling lag causes the real-time measurement to be inconsistent with the topology, the model can still use semantic embedding to correct the node representation. In addition, GraphSAGE's hierarchical sampling and LSTM aggregation mechanism are highly robust to the temporal properties of neighbor order and features, and can alleviate some of the lag noise to a certain extent, thereby maintaining a higher accuracy. Although delayed sampling inevitably weakens the positioning performance of GNN, the model in this paper achieves high stability and online availability of fault location through semantic fusion and dynamic adaptive design.

E. Noise Robustness

In order to evaluate the robustness of the model in complex noise environments, this paper uses signal-to-noise ratio as a quantitative indicator of noise intensity. In the simulation data, Gaussian white noise is added to the voltage and current measurement signals, and the SNR is set to six levels: 5 dB, 10 dB, 15 dB, 20 dB, 25 dB, and 30 dB. The effects of different noises are controlled, and the results are shown in Figure 7.

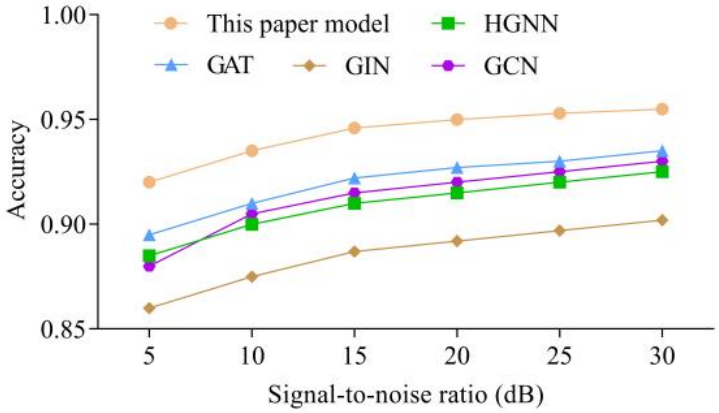

Figure 7. Noise robustness.

Under different signal-to-noise ratio conditions, all models show a trend of decreasing positioning accuracy as the noise increases, but the robustness of the "Knowledge Graph + GraphSAGE" model is the most significant: from 5 dB to 30 dB, the accuracy is significantly improved, and it still maintains a high accuracy of 0.920 in a low signal-to-noise ratio (5 dB) environment. In contrast, other models are greatly affected by noisy data. When the signal-to-noise ratio is 5dB, the accuracy of HGNN, GAT, GIN, and GCN are 0.885, 0.895, 0.860, and 0.880, respectively. The knowledge graph provides semantic priors such as device attributes and fault causal links for node embedding. When the measurement signal is distorted by noise, the graph semantics can effectively compensate for the shortcomings of pure data-driven methods. At the same time, the neighborhood sampling and LSTM aggregation strategies of GraphSAGE can extract stable time series features from noise samples and suppress anomalies, thereby improving positioning reliability. In general, the framework that integrates semantics and dynamic graph learning maintains excellent anti-interference performance in high-noise scenarios, meeting the dual requirements of accuracy and stability for online fault diagnosis of power transmission and transformation systems.

F. Positioning Performance for Arc Faults

The single-phase grounding fault is changed to an Emanuel arc fault, and the positioning performance is shown in Figure 8.

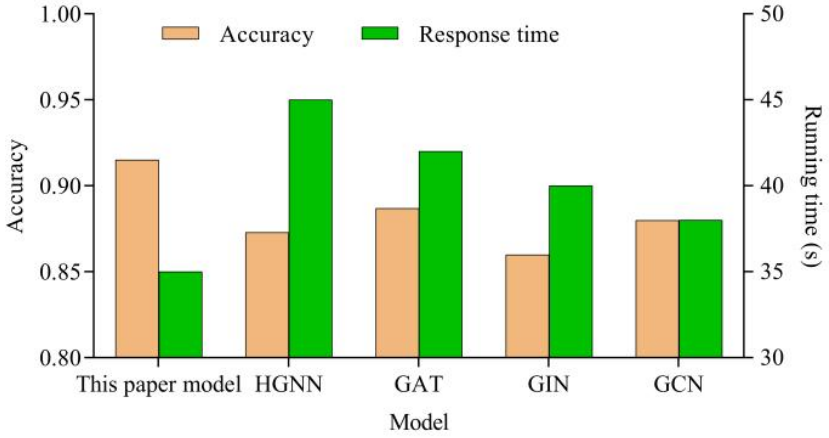

Figure 8. Emanuel arc fault location performance.

After replacing the single-phase ground fault with the more complex Emanuel arc fault scenario, the location accuracy of all models decreased to varying degrees. The "knowledge graph + GraphSAGE" framework proposed in this paper still maintained the highest accuracy of 0.915, surpassing the second place GAT (0.887), and significantly better than HGNN (0.873), GCN (0.880) and GIN (0.860). The nonlinear oscillation characteristics and time-varying arc resistance of Emanuel arc faults make it difficult for traditional positioning methods based on transient waveform mutations to capture fault characteristics, especially in the early weak arc stage, when the signal is highly overlapped with the normal operating state. In contrast, the model in this paper relies on the semantic embedding of the causal relationship between devices and historical arc fault patterns in the knowledge graph to successfully separate arc fault nodes from non-fault nodes in the embedding space, compensating for the information loss caused by the missing or distorted waveform features. At

the same time, GraphSAGE's hierarchical neighborhood sampling and LSTM aggregation can integrate multi-time measurement and protection device signals, effectively encode the temporal evolution process of arc faults, and thus improve the classifier's ability to distinguish complex faults.

In terms of response time, the average delay of this model is 35 ms, which is significantly better than HCNN (45 ms), GAT (42 ms) and GIN (40 ms), and has a slight gap with GCN (38 ms). This is mainly due to two aspects: first, the semantic prior of the knowledge graph means that the model does not need to calculate complex rules or reconstruct large-scale subgraphs during online reasoning. It only needs to retrieve relevant entity attributes from the graph and initialize node features; second, GraphSAGE's inductive neighborhood sampling mechanism only performs learnable aggregation on fixed small-scale neighborhoods, avoiding the high computational overhead brought by full-graph convolution. The "Knowledge Graph + GraphSAGE" framework maintains excellent positioning accuracy in the face of highly nonlinear and volatile waveform scenarios such as arc faults. In addition, with modular design and dynamic topology adaptation, it meets the real-time requirements of the power grid for fault diagnosis and provides reliable performance guarantees for actual engineering deployment.

G. Topology Mutation Test

In the IEEE 39-node system, topology reconstruction is achieved by disconnecting branches, and the positioning accuracy and delay of the test model after topology mutation are tested. The branches related to load node 27 and generator node 31 are selected for disconnection test. The topology mutation test results are shown in Table 7.

Table 7. Topology mutation test results.

Node	Branch disconnection	Accuracy	Response time (ms)
27	27-26	0.941	35.3
	27-28	0.938	35.5
	27-29	0.935	35.7
31	31-26	0.943	35.4
	31-30	0.940	35.2

The data in Table 7 show that in the IEEE 39-node system, after the branches related to load node 27 and generator node 31 are disconnected, the model positioning accuracy remains in the range of 0.935–0.943, and the response time fluctuates by only 0.5 ms. Compared with the 0.953 accuracy under the static topology, the topology mutation leads to a decrease in performance, but it is still better than the traditional method (such as 0.928 of GAT). Reasons for performance retention: 1) Dynamic edge weight mechanism updates the adjacency matrix in real time through SCADA/PMU signals to compensate for the fault propagation path deviation caused by topology reconstruction; 2) Knowledge graph semantic embedding provides causal links between device attributes and historical faults, and enhances node feature representation through semantic association when the topology is partially missing; 3) Incremental update strategy avoids full retraining and only fine-tunes local parameters. The accuracy of disconnection of the relevant branches of load node 27 is slightly lower than that of the relevant branches of generator node 31, because the signal attenuation on the load side is faster, and more neighborhood semantic association compensation is required. The experiment verifies the robustness and real-time performance of the model in the scenario of dynamic changes in power grid topology, meeting the online positioning requirements of complex power transmission and transformation systems.

Due to the strong nonlinearity and time-varying characteristics of arc resistance, and the serious interference of noise in the measurement signal, most comparison models mistakenly identify it as adjacent lines. The advantages of this model are reflected in the knowledge graph providing causal relationship constraints: narrowing the search scope through the semantic chain of "protective relay R26 action → circuit breaker CB26 tripping → fault source located near L26"; The temporal aggregation capability of GraphSAGE allows the LSTM layer to extract stable time series features from noisy samples and suppress abnormal fluctuations.

H. Cross-scenario Generalization Analysis

The cross-scenario generalization of the model is analyzed through IEEE 14, 30, 57, 118, and 300 node systems, and the results are shown in Figure 9.

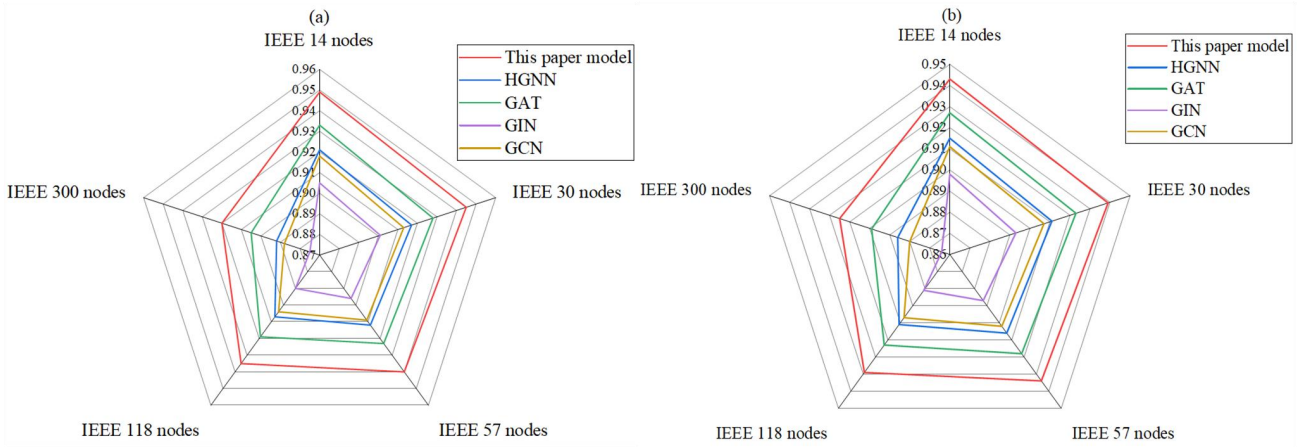

Figure 9. Cross-scenario performance. Figure 9(a) Accuracy; Figure 9(b) F1 score.

Figure 9(a) shows that the accuracy of the paper's model in the IEEE 14–300 node system is 0.920–0.949, significantly better than the comparison methods such as HGNN (0.892–0.921) and GAT (0.905–0.933). The performance of the paper's model only drops by 3.1% (IEEE 14→300 nodes), indicating that the paper's model has strong cross-scale generalization. Through the causal links between device attributes and historical faults, the feature fuzziness caused by measurement noise and increased topological complexity in large-scale systems is compensated, so that the semantics of the knowledge graph is enhanced. The dynamic topological reasoning mechanism adapts to the changes in fault propagation paths of systems of different scales by adjusting the adjacency matrix in real time. In contrast, pure data-driven models (such as GCN) rely on local topology and lack semantic constraints, and their performance decreases significantly as the scale of the system increases. Experiments verify the generalization of this method in multi-scale power grid scenarios.

Figure 9(b) shows that the F1 score of the paper's model is stable at 0.915–0.943 in the IEEE 14–300 node system, which is significantly better than GAT (0.899–0.927) and GIN (0.864–0.898), indicating that the paper's model maintains a higher recall and precision balance under unbalanced data. The LSTM aggregator integrates multiple time series measurement features to enhance the ability to capture weak fault signals in large-scale systems. The joint loss function balances the classification error and knowledge consistency constraints. In contrast, GIN relies on the assumption of isomorphic structure, and its F1 changes greatly (IEEE 14→300 nodes). The model in this paper achieves high stability and interpretability across scenarios through the collaboration of semantic enhancement and dynamic graph learning.

The performance comparison between complete updates and incremental updates is shown in Table 8.

Table 8. Performance Comparison between Full Update and Incremental Update.

Index	Full update	Incremental update
Average update time (s)	120–180 s	5–15 s
Number of rounds of model convergence	50-100 rounds	5-10 rounds
Memory usage (MB)	500–800 MB	50–100 MB
Topology change response delay	higher	Responsive
Decreased positioning accuracy (after topology mutation)	Decrease (about 5%)	Within ± 1%

6. Conclusions

This article proposes an intelligent positioning framework that integrates knowledge graph and GraphSAGE to address the problems of insufficient feature representation ability and poor topology adaptability in traditional methods for fault location in complex power transmission and transformation systems. By constructing a dynamic knowledge graph and combining SCADA/PMU data to achieve semantic enhancement and real-time topology inference, and introducing an incremental graph update mechanism, the robustness and response speed of the model to complex scenarios such as high impedance, noise, and topology changes were improved. A localization accuracy of 0.953 was achieved on the IEEE 39 node system. The research results provide high-precision and high stability online fault diagnosis solutions for power transmission and transformation systems, with good engineering application prospects. However, the study did not consider the transfer learning of multi regional power grids and extreme concurrent fault scenarios. In the future, digital twins and multimodal fusion will be combined to further enhance the model's generalization ability and real-time performance.

Acknowledgment

None

Consent to Publish

The manuscript has neither been previously published nor is under consideration by any other journal. The authors have all approved the content of the paper.

Funding

None

Author Contribution

[Ming Song]: Developed and planned the study, performed experiments, and interpreted results. Edited and refined the manuscript with a focus on critical intellectual contributions.

[Yi Zhou, Kangwei Li]: Participated in collecting, assessing, and interpreting the date. Made significant contributions to date interpretation and manuscript preparation.

[Haitao Jiang, Hua Wen]: Provided substantial intellectual input during the drafting and revision of the manuscript.

Conflicts of Interest

The authors declare that they have no financial conflicts of interest.

References

[1] X.C. Jiang, Y.P. Xu, Y.C. Li, L.F. Di, Y.D. Liu, G.H. Sheng. Digitalization transformation of power transmission and transformation under the background of new power system. High Voltage Engineering, 2022, 48, 1-10. DOI: 10.13336/j.1003-6520.hve.20211649

[2] N. Voropai. Electric power system transformations: A review of main prospects and challenges. Energies, 2020, 13(21), 5639. DOI: 10.3390/en13215639

[3] T. Han, D. J. Hill. Learning-based topology optimization of power networks. IEEE Transactions on Power Systems, 2023, 38(2), 1366-1378. DOI: 10.1109/TPWRS.2022.3170083

[4] H.H. Bian, Z.Y. Guo, X.M. Wang, C.G. Zhou, S.W. Bing, Z.Y. Zhang. An Improved Breadth-First Search Method Based on Information Interaction Applied for Power Network Topology Analysis. Mathematical Problems in Engineering, 2023, 9990917. DOI: 10.1155/2023/9990917

[5] F.Q. Meng, S.S. Yang, J.D. Wang, L. Xia, H. Liu. Creating knowledge graph of electric power equipment faults based on BERT–BiLSTM–CRF model. Journal of Electrical Engineering & Technology, 2022, 17(4), 2507-2516. DOI: 10.1007/s42835-022-01032-3

[6] J. Wang, X. Wang, C.Q. Ma, L. Kou. A survey on the development status and application prospects of knowledge graph in smart grids. IET Generation, Transmission & Distribution, 2021, 15(3), 383-407. DOI: 10.1049/gtd2.12040

[7] F. Wang, Z.J. Hu. A Novel Fault Diagnosis and Accurate Localization Method for a Power System Based on GraphSAGE Algorithm. Electronics, 2025, 14(6), 1219. DOI: 10.3390/electronics14061219

[8] B.T. Zhai, D.S. Yang, B.W. Zhou, G.D. Li. Distribution System State Estimation Based on Power Flow-Guided GraphSAGE. Energies, 2024, 17(17), 4317. DOI: 10.3390/en17174317

[9] R. Liu, R. Fu, K. Xu, X.Z. Shi, X.N. Ren. A review of knowledge graph-based reasoning technology in the operation of power systems. Applied Sciences, 2023, 13(7), 4357. DOI: 10.3390/app13074357

[10] T.J. Pu, Y.P. Tan, G.Z. Peng, H.F. Xu, Z.H. Zhang. Construction and application of knowledge graph in the electric power field. Power System Technology, 2021, 45(6), 2080-2091

[11] G. Li, Y.Q. Li, H.T. Wang, Q. Xie, W.Q. Huang, J.X. Hou. Knowledge graph for power equipment health management: basic concepts, key technologies and research progress. Automation of Electric Power Systems, 2022, 46(3), 1-13. DOI: 10.7500/AEPS20210804001

[12] J.M. Yu, X.H. Wang, Y. Zhang, Y. Liu, S.A. Zhao, L.F. Shan. Construction and application of knowledge graph for intelligent control. Power System Protection and Control, 2020, 48(3), 29-35. DOI: 10.19783/j.cnki.pspc.191200

[13] Z.P. Gao, Y. Zhao, Y.L. Yu, Y.J. Luo, Z.W. Xu, L.M, Zhang. Low-voltage distribution network topology identification method based on knowledge graph. Power System Protection and Control, 2020, 48(2), 34-43. DOI: 10.19783/j.cnki.pspc.190379

[14] X.Z. Ye, L. Shang, X.Z. Dong, C.X. Liu, Y. Tian, H.L. Fang. Research and application of knowledge graph for distribution network fault handling. Power System Technology, 2022, 46(10), 3739-3748. DOI: 10.13335/j.1000-3673.pst.2022.0496

[15] C. Li, B. Wang. A knowledge graph method towards power system fault diagnosis and classification. Electronics, 2023, 12(23): 4808. DOI: 10.3390/electronics12234808

[16] Y. Zhou, Z. Lin, L. Tu, Y. Song, Z. Wu. Big data and knowledge graph based fault diagnosis for electric power systems. EAI Endorsed Transactions on Industrial Networks Intelligent Systems, 2022, 9(32), 1. DOI: 10.4108/eetinis.v9i32.1268

[17] H.Y. Chen, L. Guan. Topology-adaptive transient stability assessment of power systems based on active transfer learning. Proceedings of the CSEE, 2023, 19, 7409-7422. DOI: 10.13334/j.0258-8013.pcsee.221192

[18] J.W. Li, X.J. Wang, J.H. He, Y.J. Zhang, D.H. Zhang. Distribution network fault location method based on graph attention network.Power System Technology, 2021, 45(6), 2113-2121. DOI: 10.13335/j.1000-3673.pst.2020.2222

[19] K. Xu, X.Y. Fan, H.R. Zhang. Distribution network fault location and fault type identification based on graph convolutional network. Experimental Technology & Management, 2023, 1, 26-30.

[20] X.L. He, H.J. Gao, Y. Huang, Y.W. Gao, R.J. Wang. Distribution network fault section location method based on one-dimensional convolution and graph neural network. Power System Protection & Control, 2024, 52(17), 27-39. DOI: 10.19783/j.cnki.pspc.240021

[21] Chen Xiaolong, Sun Lirong, Li Yongli, Li Bin, Wang Li, Cai Yanchun, et al. Distribution network fault section location method based on graph attention network and consistency risk control. Power System Technology 47.12 (2023): 4866-4876

[22] C. Hongi, Y.W. Wu, W. Gao, M.F. Guo. Distribution network fault location method based on GrapSAGE

algorithm. Journal of Electronic Measurement and Instrumentation, 2023, 37(11), 236-236.
[23] Y.R. Zhuang, T.N. Xiao, L. Cheng, Y. Chen, H.Z. Guan. Power system transient stability assessment based on spatiotemporal graph convolutional network. Automation of Electric Power Systems, 2022, 46(11), 11-18. DOI: 10.7500/AEPS20210629014
[24] S.F. Sun, X.L. Li, W.S. Li, D.J. Lei, S.H. Li. A review of graph neural network applications in knowledge graph reasoning. Journal of Computer Science and Technology, 2023, 17(1), 27-52. DOI: 10.3778/j.issn.1673-9418.2207060
[25] W. Liao, B. Bak-Jensen, J.R. Pillai, Y. Wang, Y. Wang. A review of graph neural networks and their applications in power systems. Journal of Modern Power Systems and Clean Energy, 2021, 10(2), 345-360. DOI: 10.35833/MPCE.2021.000058
[26] B. Huang, J. Wang. Applications of physics-informed neural networks in power systems-a review. IEEE Transactions on Power Systems, 2022, 38(1), 572-588. DOI: 10.1109/TPWRS.2022.3162473
[27] R. Zhang, J.Q. Liu, B.Y. Zhang, J.X. Li, T.L. Gao, J. Zhang. Research on construction of knowledge graph for power grid fault handling based on transfer learning and real-time decision support. Electric Power Information and Communication Technology, 2022, 20, 24-34.
[28] L.Y. Lin, Q. Chen, L. Jin, L. Wang. Research and application of knowledge representation of substation alarm information fault based on knowledge graph. Power System Protection and Control, 2022, 50(12), 90-99. DOI: 10.19783/j.cnki.pspc.211136
[29] Y.M. Zhang, W.X. Zhang, B.H. Li, W.J. Deng, J.J. Lu. Analysis of the influence mechanism of flexible DC power grid topology on fault current. Electric Power Engineering Technology, 2022, 41(5), 94-102. DOI: 10.12158/j.2096-3203.2022.05.011
[30] J.L. Zhang, Z.J. Gao, M. Chen, Z.Wei, An Shuhuai. Fault location method of active distribution network considering complex faults. Transactions of China Electrotechnical Society, 2021, 36(11), 2265-2276. DOI: 10.19595/j.cnki.1000-6753.tces.200487
[31] M. Khalili, M. Ali Dashtaki, M.A. Nasab, H. Reza Hanif, S. Padmanaban, B. KhanH. Optimal instantaneous prediction of voltage instability due to transient faults in power networks taking into account the dynamic effect of generators. Cogent Engineering, 2022, 9(1), 2072568. DOI: 10.1080/23311916.2022.2072568
[32] R.R. Hossain, Q. Huang, R. Huang. Graph convolutional network-based topology embedded deep reinforcement learning for voltage stability control. IEEE Transactions on Power Systems, 2021, 36(5), 4848-4851. DOI: 10.1109/TPWRS.2021.3084469
[33] K. Liu, W. Sheng, X. Yang. Single phase to ground fault location of distribution network based on combined-gat. Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering), 2022, 15(6), 465-474. DOI: 10.2174/2352096515666220624160925
[34] W. Liao, D. Yang, Q. Liu, Y. Jia, C. Wang, Z. Yang. Data-driven reactive power optimization of distribution networks via graph attention networks. Journal of Modern Power Systems and Clean Energy, 2024, 2024, 12(3), 874-885. DOI: 10.35833/MPCE.2023.000546