AEM JOURNAL.ORG

Fault Diagnosis of Drop-Out Fuses Based on Hidden Markov Model and Graph Embedding

Shoujun Bao, Shuhong Li, Ziyong Pan, Fuxiang Liang, Shenghai Han
State Grid Qinghai Province Electric Power Company, Yushu Power Supply Company, Yushu 815000, Qinghai, China
Corresponding author: Shoujun Bao (e-mail: 13132007930@163.com)

ABSTRACT Traditional fault diagnosis methods relying on manual inspection and static thresholds often struggle to identify subtle or evolving faults in drop-out fuses operating within increasingly complex power distribution environments, where electromagnetic interference and dynamic operating conditions may affect monitoring reliability. To address these challenges, this paper proposes a fault diagnosis method based on a hidden Markov model and graph embedding. The proposed framework consists of three stages: constructing a power equipment graph and employing an improved graph convolutional network to extract topology-aware embedding vectors; designing a four-state hidden Markov model using multidimensional time-series data, including current, voltage, and temperature, to characterize dynamic state evolution through the Baum–Welch algorithm; and integrating graph features with HMM state sequences for precise fault identification and real-time warning. Experimental results demonstrate that the proposed method achieves a fault diagnosis accuracy of 94.21% with an F1-score of 92.74%. For non-obvious faults, the F1-score remains between 0.887 and 0.945, while the average warning lead time exceeds 17 minutes. Practical deployment further reduces the manpower required for a single fault response process to 3.6 person-hours. The proposed framework significantly enhances fault detection accuracy and response efficiency, providing an effective intelligent perception solution for power distribution systems and offering technical support for reliable monitoring and maintenance of modern electromagnetic energy infrastructures.

INDEX TERMS hidden Markov model, graph embedding, drop-out fuse, fault diagnosis

I. INTRODUCTION

DROP-out fuses are fundamental for ensuring the safe and stable operation of medium- and low-voltage distribution systems as essential short-circuit and overload protection equipment [1,2]; their reliable operation is analogous to the critical electrical safeguards required for high-throughput machinery, where sudden faults could halt continuous production [3-5]. Such equipment is usually installed at key locations such as feeder branch points and the front end of distribution transformers. These devices have the advantages of a simple structure, low cost, and convenient maintenance. However, with the continuous expansion of the scale of distribution systems and the diversification of load types, their operating environment is becoming increasingly complex [6,7]. Fuses are easily affected by multiple factors, such as load fluctuations, environmental interference, and equipment aging, resulting in their faults becoming more hidden and diverse [8].

This study specifically examines drop-out fuses, which share common characteristics with various industrial overload devices (e.g., thermal relays or electronic protection modules in machinery) in terms of overload protection mechanisms, state evolution patterns, and monitoring data structures. These devices all exhibit progressive state degradation under multidimensional sensing signals. Consequently, the proposed method demonstrates potential for logical modeling transfer to other overload protection scenarios. However, the experimental validation in this paper is limited to distribution network fuses.

II. RELATED WORK

As key protection equipment in medium- and low-voltage power distribution systems, drop-out fuses are subject to various fault types, including blown fuses, arcing, and poor contact [9-11]. Traditional detection methods, such as fault diagnosis methods based on thermal imaging technology, can achieve a high detection accuracy when an obvious fault occurs [12]. Another type of research has improved the precision of fuse multi-fault diagnosis through deep learning models, which is suitable for high-frequency testing needs in industrial scenarios [13]. In summary, existing detection methods have significant limitations in dealing with load complexity and equipment state variability, and they cannot meet the needs of early warning and dynamic tracking for intelligent operation and maintenance of distribution networks.

In view of the deficiency of traditional methods in ignoring the topological structural information between distribution network equipment, some studies have successfully revealed the electrical dependencies between equipment by constructing a power grid topology map based on geographic information system (GIS) data, which effectively assisted in fault location and propagation analysis [14]. At the same time, some scholars have studied the application of graph neural networks (GNN) in power systems and developed a node anomaly detection model based on Graph Convolutional Networks (GCN) , which significantly improved the accuracy and robustness of fault identification [15]. However, these works mostly focus on the single perspective of static graph structure and fail to fully integrate the dynamic changes of equipment operation state, making it difficult to fully reflect the evolution of the equipment in complex environments. In addition, there are also studies that attempted to combine graph embedding with traditional machine learning methods, but due to the limitations of shallow embedding technology, the model's expressiveness and generalization performance still need to be improved [16].

III. IMPLEMENTATION OF DROP-OUT FUSE FAULT DIAGNOSIS METHOD COMBINING HIDDEN MARKOV MODEL AND GRAPH EMBEDDING

DATA COLLECTION AND PREPROCESSING

The data used in this study are from the actual operation monitoring system of a 10 kV distribution network in a certain area, covering 184 fuse units deployed across 12 typical distribution lines. The monitoring system continuously samples the equipment operation state in a 1-minute cycle. The collection cycle covers the 6-month summer and autumn peak operation period, spanning various climate ranges and load scenarios, with strong temporal integrity and environmental diversity. The training set, validation set, and test set are partitioned chronologically, corresponding to the data from the first 4 months, 5th month, and 6th month, respectively, ensuring no time overlap and preventing future information leakage. All data were sourced from the long-term online monitoring system of the actual distribution network, without using accelerated aging or simulation data. The Minor Anomaly status was composed of equipment samples marked with warning events in the operation and maintenance records that were later confirmed to have no fuse blowout but exhibited continuous temperature rise or current deviation. The collection indicators include multi-dimensional physical indicators such as three-phase current (Ia, Ib, and Ic), three-phase voltage (Va, Vb, and Vc), fuse temperature (T), equipment operation frequency (fm), and warning event trigger identification [17,18].

Since the initial data structure collected is a multi-dimensional time-series dataset constructed with equipment number and timestamp as the primary key, there is a certain proportion of missing measurements, repeated records, and short-term abnormal spikes, which may affect the stability and identification precision of subsequent modeling. Therefore, this study systematically preprocesses the data. All observed data are aligned uniformly according to the equipment number and time axis in the preprocessing stage, and equipment nodes with a missing rate of more than 30% are eliminated to avoid low-quality data from interfering with model training. To alleviate the interference caused by high-frequency fluctuations, a sliding window median filter is used to smooth short-term abnormal spikes to improve the stability and robustness of the sequence. When dealing with missing data, local linear interpolation is applied to fill in the missing time intervals, which are less than 5 minutes. Sequences with long missing time spans are directly eliminated from the modeling data to ensure the continuity and validity of the observed data. After completing the above operations, Z-score standardization is applied to all time-series features to eliminate inter-dimensional scale differences. The standardization formula is as follows:

$$x' = \frac{x-\mu}{\sigma} \tag{1}$$

In Formula (1), μ and σ denote the mean and standard deviation of each feature across all equipment, computed exclusively from the training-set observations. They remain fixed during the testing phase to prevent information leakage. The standardized data dimensions remain consistent, effectively eliminating the dimension effect and enhancing the model's sensitivity to feature fluctuations. After the above preprocessing, the final multi-dimensional time series data set contains 184 equipment nodes, and the cumulative time series sample volume is about 472,000, which constitutes the basic input data for the modeling and training phase of this study. Here, sample size refers to the number of effective sliding-window samples used for model training and validation. Each window contains multidimensional observations across multiple consecutive time steps, not the total number of original minute-level observations.

GRAPH EMBEDDING FEATURE EXTRACTION

To mine the complex structural dependencies and semantic associations between equipment in the power distribution system, a graph structure model based on equipment connection relationships is constructed, and the embedded feature representation of nodes is extracted based on the graph neural network architecture. Compared with the limitations of traditional neural networks in processing data in Euclidean space, GCN allows information aggregation on arbitrary topological structures, significantly improving the model's adaptability to complex physical connection structures [20]. Standard GCNs employ static, symmetric adjacency normalization, which fails to capture asymmetric electrical interactions between distribution network devices (e.g., upstream fuses providing unidirectional protection to downstream components). To address this limitation, this study integrates Graph Attention Mechanism (GAT) into GCNs. By dynamically assigning learnable attention weights to model edge importance, the model can distinguish critical connections from redundant ones, thereby more accurately representing heterogeneous dependencies in power equipment networks. Graph structure modeling does not replace the thermoelectric fault mechanism. Instead, it provides differentiated contextual interpretations for identical current/temperature signals by encoding electrical adjacency relationships in distribution networks (e.g., upstream feeder capacity, downstream load density), thereby enhancing the discriminative power for local degradation states. In the graph construction process, self-loops are first added to the original adjacency matrix A to construct a new adjacency matrix $\widehat{A} = A + I$, ensuring that the node can consider its own information while updating the feature representation. To overcome the feature scaling offset caused by the difference in node degrees, symmetric normalization is applied, and the degree matrix $\widehat{D}$ is defined to construct a standardized convolution kernel. The symmetric normalization here is only used to construct the initial adjacency structure; in the subsequent graph attention layer, the edge weights are dynamically generated by learnable asymmetric attention coefficients, thus effectively modeling the directed electrical dependencies between devices. The feature update at the $(l+1)$-th layer is given by:

$$H^{(l+1)} = \sigma\left(\widehat{D}^{-\frac{1}{2}}\widehat{A}\widehat{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \tag{2}$$

In Formula (2), $H^{(l)}$ is the node feature matrix of the l-th layer. $W^{(l)}$ is the weight matrix to be trained. σ is the nonlinear activation function. In this study, ReLU is selected to enhance the nonlinear ability of feature expression.

Considering the impact of the number of GCN layers on model performance and training efficiency, this study adopts a 3-layer GCN structure, which not only ensures multi-hop neighbor information transmission but also avoids the gradient vanishing and overfitting problems caused by too deep networks. The number of layers is determined through cross-validation and empirical tuning, reflecting a good balance between model performance and complexity. In view of the heterogeneity of the impact of connection edges on equipment state in the distribution network, simple graph convolution cannot accurately capture the differentiated influence of key neighbor nodes; therefore, GAT is used. This mechanism achieves weighted aggregation of adjacent nodes through learnable attention weights, enhancing the model's ability to express important connection relationships. The mechanism first projects the node features into a new feature space through linear transformation and then concatenates the projection results of the central node i and its neighbor node j. The attention score is calculated as follows:

$$e_{ij} = \mathrm{LeakyReLU}\left(a^{\top}[Wh_i \| Wh_j]\right) \tag{3}$$

In Formula (3), $a^{\top}$ is the learnable attention vector; W is the shared weight matrix; $\|$ represents the vector concatenation operation. To ensure the normalization of neighborhood weights, the attention coefficients of neighbors of the same node are normalized through the softmax operation:

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k\in N_i}\exp(e_{ik})} \tag{4}$$

Finally, the updated features of node i are obtained by a weighted aggregation of its neighboring nodes:

$$h_i' = \sigma\left(\sum_{j\in N_i}\alpha_{ij}Wh_j\right) \tag{5}$$

To improve the expression stability and diversity of the attention mechanism, four parallel attention heads are used to aggregate neighbor information of different subspaces through a multi-head mechanism to achieve a more fine-grained local relationship characterization. The number of heads is determined after multiple rounds of debugging, considering both performance improvement and computational overhead. In addition, to enrich the node representation, the structural semantic embedding h_i' extracted by the fusion of graph convolution and attention mechanism is concatenated with static attribute vectors s_i such as equipment category, installation time, rated current, voltage level, etc., to form a unified feature expression that integrates structural semantics and attribute semantics. The specific concatenation operation is:

$$f_i = \mathrm{concat}(h_i', s_i) \tag{6}$$

Subsequently, the comprehensive feature f_i is input into a multi-layer fully connected neural network, and the interactive features of multi-dimensional information are further extracted through linear transformation and nonlinear activation functions, which improves the complementary expression ability between different types of data and the model's generalization performance. Finally, a high-dimensional feature vector with rich topological semantics and equipment attribute information is output, providing high-quality data input for subsequent state monitoring, fault diagnosis, and predictive analysis.

MULTI-SOURCE INFORMATION FUSION AND FAULT DIAGNOSIS

To improve the accuracy of fuse state identification and the sensitivity of anomaly detection in distribution systems, a fault diagnosis framework based on a multi-source information fusion strategy is constructed. The framework integrates structural information and time series evolution features to achieve precise identification and early warning of potential fuse faults.

In terms of graph structural information modeling, GCN is used to topologically embed each node (fuse) in the distribution network and extract its low-dimensional feature vector to reflect the structural importance of the node and the load connection attributes. Its formal expression is:

$$v_g = \mathrm{GCN}(X, A), \quad v_g \in \mathbb{R}^{d_1} \tag{7}$$

In Formula (7), X is the original feature matrix of the node. Under the combined effect of stable topological structure and dynamic load, the graph embedding vector v_g has a strong context expression ability, which effectively supplements the spatial dependency features that cannot be captured in traditional time series models.

The time series modeling is based on HMM, which models the operation state of each fuse in the observation period and outputs the state sequence and the corresponding probability distribution, which is specifically expressed as:

$$\begin{aligned} P_{\mathrm{HMM}} &= \{p_{t,i}\}_{t=1}^{T}, \\ p_{t,i} &= P(s_t = i \mid O_{1:t}), \\ P_{\mathrm{HMM}} &\in \mathbb{R}^{T\times d_2} \end{aligned} \tag{8}$$

In Formula (8), $O_{1:t}$ represents the observation sequence up to time step t, and $p_{t,i}$ is the probability that the hidden state at time t is i. In the multi-source feature fusion stage, the concatenation fusion strategy is used to jointly encode the structural information and the time series state evolution information. For each node, the GCN structural feature vector v_g is first extracted, and then it is expanded and concatenated with the hidden-state probability matrix P_{HMM} of the past T time steps to obtain the final input feature vector:

$$v_{\mathrm{diag}} = [v_g \| \, \mathrm{flatten}(P_{\mathrm{HMM}})] \in \mathbb{R}^{d_1 + T\cdot d_2} \tag{9}$$

The classifier uses an ensemble learning method combining random forest and XGBoost to train and predict fusion feature v_{diag}. The Fully Connected Network (FCN) is employed solely as a non-linear feature enhancement and compression module. It is trained independently using cross-validation and its parameters are frozen thereafter. The concatenated feature vector is fed into a two-layer fully connected network (with a 512-dimensional hidden layer and ReLU activation) before being passed through an ensemble classifier. The Random Forest and XGBoost models are trained independently and do not share gradients with the FCN. For XGBoost, learning rate 0.1, maximum depth 6, subsampling rate 0.8, number of trees 300, early stopping rounds 30 (based on the validation set F1); the Random Forest uses 100 trees, with a maximum number of features of $\sqrt{d}$ (where d is the input dimension). All experiments were performed on an Intel Xeon Gold 6248R CPU + NVIDIA A100 GPU. The four types of operation state labels consistent with the above are used, and the model hyperparameters are optimized through grid search to improve the model's generalization ability. The four-state Hidden Markov Model (HMM) employs a dual calibration approach combining historical fault data with expert rules to define operational thresholds: Normal operation is confirmed when current $\leq$ 1.1 times rated value, voltage fluctuation < $\pm$ 5%, and temperature < 70 °C. Minor anomalies occur when any parameter exceeds normal limits but remains below critical thresholds (current $\leq$ 1.3 times rated value, voltage < $\pm$ 10%, temperature < 85 °C). Critical anomalies are identified by further deterioration of parameters without tripping. Fault status is confirmed by either a tripping signal or a temperature reading $\geq$ 105 °C. The emission probability of the HMM adopts a three-component Gaussian mixture model (GMM). The initial parameters of each state are statistically estimated by a subset of expert thresholds and semisupervised optimization is performed on the complete observation sequence using the Baum-Welch algorithm, rather than directly hard-assigning states based on thresholds.

To achieve early warning of fault conditions, this study establishes a criterion based on state transition probability. It is assumed that the HMM hidden state transition matrix is $A = [a_{ij}]$, where:

$$a_{ij} = P(s_{t+1} = j | s_t = i) \tag{10}$$

When the current hidden state s_t belongs to the abnormal category and the cumulative probability of transitioning to the fault state s_f in the next H steps satisfies:

$$P_{cum} = P\left(\bigcup_{k=1}^{H} s_{t+k} = s_f \mid s_t\right) > \theta \tag{11}$$

a fault warning signal is triggered, and the threshold value θ in this study is 0.9. The selection of this threshold is based on the comprehensive balance between the sensitivity and false positive rate of the early warning system. A high threshold can effectively reduce the false positive rate and avoid affecting the stable operation of the system and the accuracy of operation and maintenance decisions due to the frequent triggering of early warnings. At the same time, the threshold of 0.9 can still maintain a high ability to capture potential faults, ensuring the timeliness and reliability of early warning signals. In addition, the early warning confidence interval is estimated through cross-validation errors to assist in determining the stability of the early warning and further enhance the robustness of the model in practical applications.

The overall diagnosis process integrates the structural topology perception ability of the graph neural network and the state evolution representation ability of the hidden Markov model and combines the integrated classifier to complete the final judgment, realizing multi-dimensional and refined identification and early warning of the operation state of the fuse. Figure 1 shows the overall fault diagnosis process. Actual HMM modeling uses a four-state system (Normal, Minor Anomaly, Severe Anomaly, Fault). In Figure 1, the Warning state represents a combination of the Minor Anomaly and Severe Anomaly states and is used only for flowchart simplification:

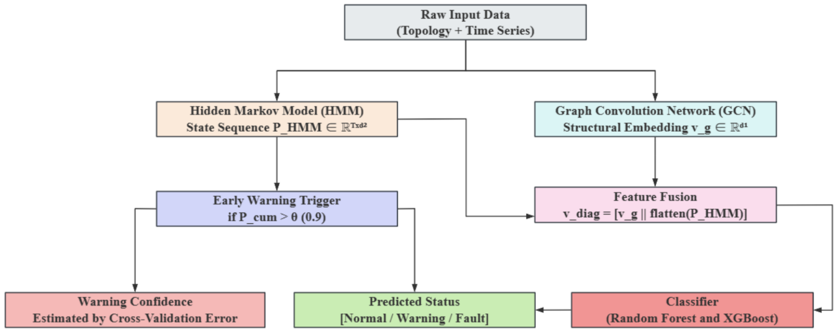

FIGURE 1. Overall fault diagnosis flow chart

IV. EXPERIMENTAL PROCESS AND RESULT ANALYSIS

(1) Fault identification performance and false positive and false negative analysis

To fully verify the effectiveness and stability of the drop-out fuse fault diagnosis method combining the hidden Markov model and graph embedding proposed in this study in actual application scenarios, a performance comparison experiment including three groups of control methods is designed. The experimental analysis is conducted on the collected dataset. In terms of fault identification performance evaluation, the identification accuracy, precision, recall, and F1-score of each method on the data set are statistically analyzed. Table 1 lists the results.

TABLE 1. Performance comparison of different methods in the task of detecting drop-out fuse faults

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Experimental Group	94.21	93.87	91.64	92.74
Control Group 1	76.83	74.12	69.45	71.71
Control Group 2	89.34	88.05	85.91	86.97
Control Group 3	91.26	89.67	88.42	89.04

In the task of identifying drop-out fuse faults, the proposed diagnosis method that integrates hidden Markov models and graph embedding mechanisms shows significant performance advantages. In terms of accuracy, this method reaches 94.21%, which is significantly higher than the 76.83% of control group 1 based on traditional threshold rules, 89.34% of control group 2 based on LSTM, and 91.26% of control group 3 based on GraphSAGE. Further observing the performance of each method in terms of precision and recall, this method reaches 93.87% and 91.64%, respectively. It is not only more robust in false positive control but also has more advantages in the comprehensiveness of fault sample identification. The final F1-score is 92.74%, which is the most balanced and superior performance among all methods.

Control group 1 uses the traditional method based on artificial threshold rules. Due to the lack of modeling ability for dynamic state features, its recall is only 69.45%, which cannot deal with complex disturbances and implicit faults. Control group 2 uses LSTM to model time dependencies, and its performance improves. Its F1-score reaches 86.97%, but its modeling ability is still limited in the absence of structural information between nodes. Control group 3 uses GraphSAGE to achieve graph structure learning, and its performance is further improved. Its F1-score reaches 89.04%, indicating that graph neural networks have certain advantages in modeling local topology and inter-node associations. However, in comparison, the method of this study not only applies hidden Markov models to model time series state transitions but also combines graph embedding to achieve high-order representation of multi-dimensional features, successfully bridging the gap between graph structure modeling and time dynamic analysis, showing higher fault identification precision and stability, and verifying its wide applicability and superiority in actual power equipment fault diagnosis scenarios.

To further analyze the control ability of each method for false positives and false negatives, the false positive rate (FPR) and false negative rate (FNR) of each method on the data set are calculated. Both FPR and FNR are calculated at the sample level (observations per minute) and the average of 10 independent trials is reported. Figure 2 presents the results.

Figure 2 shows that the proposed method achieves the lowest values in both FPR and FNR, with an FPR of 3.12% and an FNR of 4.69%. In contrast, the FPR of the LSTM method is 4.73%, slightly higher than that of the method in this paper. The GraphSAGE model performs better in terms of FNR, which is 5.37%, but its FPR still reaches 5.01%. The traditional threshold method performs poorly in both indicators, with both FPR and FNR exceeding 8%, showing that the ability to identify abnormal patterns under complex working conditions is limited and there are significant deficiencies in reliability and actual deployment value.

(2) Early warning effectiveness and implicit fault identification capability

To further evaluate the early warning response capability and implicit fault identification performance of the proposed method in actual power operation scenarios, experiments are conducted from two dimensions: early warning time-

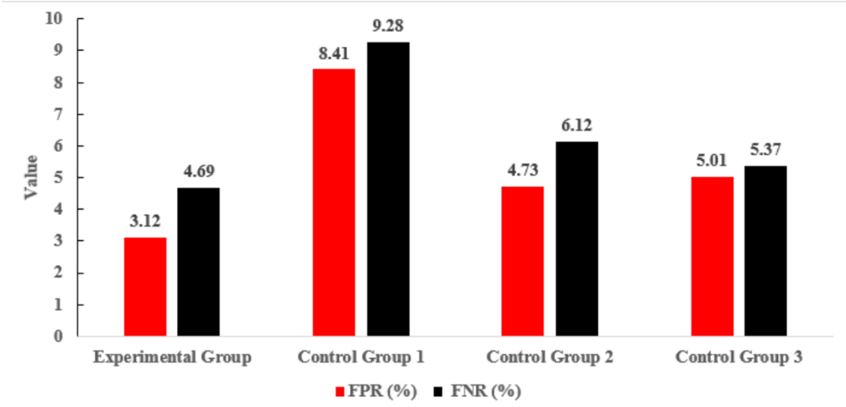

FIGURE 2. Comparison of FPR and FNR of different methods in the drop-out fuse detection task

liness and non-explicit anomaly detection. First, based on the sample fragments with clear fault timestamps in the collected data, the time point when each model first issues a high-confidence warning before the fault occurs is counted, and the time difference between it and the actual fault occurrence time is calculated as the early warning lead time. During the experiment, 10 rounds of test data from different substation combinations are collected to obtain stable early warning performance evaluation results. Each round of testing corresponds to an independent real fault event and its preceding time window, selecting a total of 10 typical fault cases from different substations and equipment. Figure 3 shows the results.

In Figure 3, different methods have significant differences in their ability to respond to early warnings. Among them, the experimental group performs best overall, achieving an average early warning lead time of at least 17 minutes in most test rounds, and even up to 20 minutes in some rounds. This advantage shows that the method proposed in this paper has strong abnormal perception and trend prediction capabilities at the early stage of fault evolution, and it can capture potential risk signals in advance. In contrast, the average early warning lead time of control group 1 is low overall, generally less than 6 minutes and less than 4 minutes in some rounds, indicating that it has obvious lag when facing complex time series backgrounds and multisource disturbances. Combined with the principles of the method in this group, it is mostly based on traditional threshold trigger mechanisms and lacks the ability to model implicit dynamics, so it responds slowly in the initial weak abnormal state. The performance of control groups 2 and 3 is at an intermediate level. The former maintains an average of 10-14 minutes and has a certain advance warning capability. Its modeling method is mainly based on strategies such as sequence learning or Bayesian regression, and it can maintain a certain sensitivity to specific types of progressive faults. The latter shows relatively higher volatility, and the lead time in some rounds is close to that of the experimental group, but its prediction ability decreases significantly in some disturbance scenarios. It is speculated that its model is more dependent on the activation of explicit features.

In terms of implicit fault detection, based on the constructed sub-datasets containing four typical nonobvious faults, including small fluctuations, low-frequency disturbances, intermittent anomalies, and multisource noise superposition, fault diagnosis is performed using each group of methods. The four categories of implicit fault labels were mutually exclusively annotated by domain experts based on post-event operation and maintenance records and multi-source signal cross-validation, and were not automatically generated by rules. During the diagnosis process, the F1-scores of each group under different faults are recorded. Figure 4 shows the results.

In Figure 4, the F1-scores of the proposed method for all four types of implicit faults are significantly better than those of control groups, showing a strong ability to identify implicit faults. Specifically, for low-frequency disturbances and small fluctuations, the experimental group achieves F1-scores of 0.945 and 0.887, respectively, showing that the hidden Markov model combined with the graph embedding method can effectively capture these weak and complex abnormal signals. This is due to the advantages of the method in this paper in time series feature modeling and graph structure expression, which can mine the subtle dynamic changes implied before and after the fault. The traditional threshold method has a relatively rough threshold setting and performs significantly insufficiently, especially in the identification of intermittent anomalies and low-frequency disturbances, with the lowest score.

In addition, during the experiment, this study also counts the response time and implicit fault identification accuracy of each method for different types of faults in the task of implicit fault detection. Response time refers to the model's average inference latency (in seconds) for a single sample, excluding data acquisition or transmission time. Table 2 lists the results.

In Table 2, in the implicit fault detection task, the response time and identification accuracy of each method under different types of faults show obvious differences. From the overall trend, the experimental group achieves a high

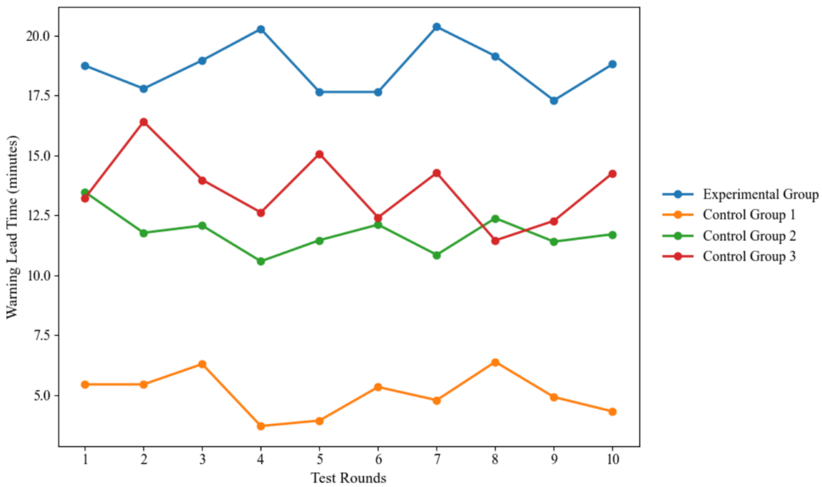

FIGURE 3. Results of 10 rounds of early warning tests for each method

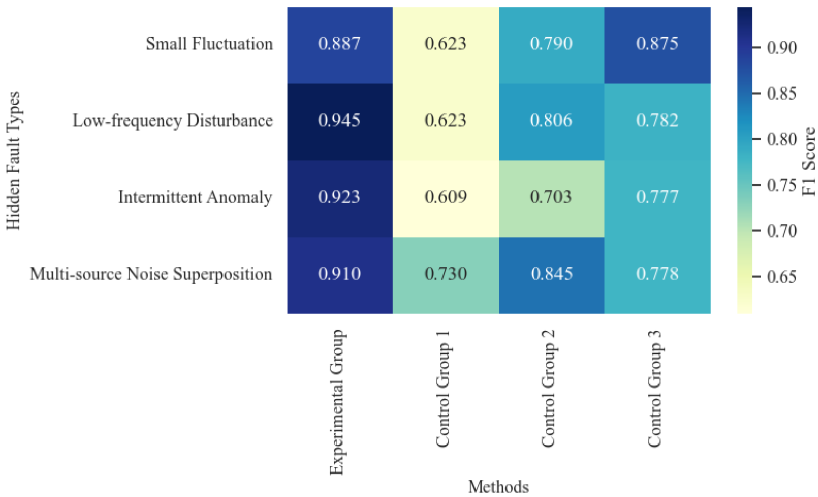

FIGURE 4. F1-score comparison heat map of different methods under different implicit faults

TABLE 2. Comparison of response time and implicit fault identification performance of each method under different fault types

Fault type	Metric	Experimental Group	Control Group 1	Control Group 2	Control Group 3
Small Fluctuation	Response Time (s)	3.2 ± 0.5	2.8 ± 0.6	4.1 ± 0.7	3.9 ± 0.6
Small Fluctuation	Identification Accuracy (%)	91.3	64.7	78.2	86.5
Low-frequency Disturbance	Response Time (s)	3.6 ± 0.4	2.9 ± 0.5	4.3 ± 0.6	4.0 ± 0.5
Low-frequency Disturbance	Identification Accuracy (%)	94.8	65.2	79.3	81.1
Intermittent Anomaly	Response Time (s)	3.4 ± 0.5	3.1 ± 0.7	4.5 ± 0.8	4.2 ± 0.7
Intermittent Anomaly	Identification Accuracy (%)	92.5	62.3	72.9	80.2
Multi-source noise superposition	Response Time (s)	3.8 ± 0.6	3.0 ± 0.6	4.7 ± 0.7	4.3 ± 0.6
Multi-source noise superposition	Identification Accuracy (%)	90.7	69.8	83.6	79.5

identification accuracy in all four types of fault scenarios, with an average accuracy of 92.3%, significantly better than 65.5% of the control group 1, 78.5% of the control group 2, and 81.8% of the control group 3. This shows that this method has stronger robustness and discrimination ability when dealing with non-explicit modes such as small fluctuations and low-frequency disturbances.

Further observations show that the experimental group's identification accuracy for typical weak signal faults, such as low-frequency disturbances and intermittent anomalies, is 94.8% and 92.5%, respectively, which is particularly outstanding. This result not only verifies the effectiveness of the proposed method in modeling non-explicit dynamic modes but also demonstrates its potential for application in actual power system operation monitoring.

(3) Adaptability evaluation under equipment type and topology

To systematically evaluate the generalization ability of the proposed method in complex power grid scenarios, a transfer test environment is constructed from two dimensions: topology complexity and node equipment heterogeneity, to verify the model's adaptability under non-training domain conditions.

In the topology test, three typical power grid structures are constructed: low-density tree topology, mediumdensity mesh topology, and high-redundancy ring topology. Each type of structure maintains a similar number of nodes to control the interference of scale factors. The above topologies are abstractly reconstructed based on real distribution network connection patterns. A zero-shot transfer strategy was used during testing, meaning the trained model was directly applied to the new topology without any fine-tuning. In each topology scenario, the proposed method and three control methods are run, respectively, and five rounds of fault diagnosis experiments are performed. The average F1-score of each method in each topology is recorded and counted. Figure 5 presents the experimental results.

In Figure 5, the experimental group maintains high stability in the three types of power grid topologies, with F1-scores of 0.91, 0.93, and 0.92, respectively, which is better than other methods overall. Control group 1 performs relatively weakly in the three topologies, especially in the ring structure, where the F1-score drops to 0.79, showing insufficient adaptability to complex structures. Control group 2 has small fluctuations under different structures, with the F1-score maintained between 0.84 and 0.86, showing certain advantages in time series modeling. Control group 3 achieves high accuracy in all topologies, with F1-scores of 0.87, 0.88, and 0.86, respectively, showing the effectiveness of graph structure learning in dealing with network complexity. However, compared with control group 3, the experimental group still shows better diagnosis capabilities in mesh and ring topologies, indicating that the proposed fusion method of hidden Markov model and graph embedding has stronger generalization ability and robustness in scenarios with irregular structures or enhanced connectivity.

In the equipment type test, multiple types of new equipment nodes that do not participate in model training are applied on the basis of the original power grid, including modular circuit breakers, high-speed sampling and metering equipment, distributed power interface equipment, and intelligent reactive power compensation controllers. An equal number of test nodes are randomly embedded under each equipment type, and disturbance samples are uniformly generated to detect the adaptability of each method to different equipment types. The diagnostic accuracy of each group of methods under different equipment nodes is statistically analyzed. Figure 6 presents the results.

In Figure 6, in the new equipment type adaptability test, each method shows obvious differences in the four types of equipment nodes that do not participate in the training. The experimental group maintains a diagnostic accuracy above 0.89 under different equipment types, which are 0.92, 0.90, 0.89, and 0.91, respectively, showing good cross-equipment generalization ability. Control group 1 has the lowest accuracy among the four types of equipment, with a minimum of 0.72 and a maximum of no more than 0.78, reflecting its extremely limited adaptability to changes in structure and data form. Control group 2 performs relatively balanced across all types of equipment, with an accuracy ranging from 0.81 to 0.85, reflecting a certain generalization ability of time series modeling. Control group 3 has a diagnostic accuracy of more than 0.85 on all equipment, with a maximum of 0.88, showing the performance advantage of graph neural networks in heterogeneous node scenarios. However, the experimental group achieves leading accuracy in various new types of equipment, further verifying the collaboration advantages of hidden Markov models and graph embedding fusion methods in complex equipment feature extraction and fault identification.

To further study the overall transfer ability of the model under conditions outside the training domain, this study also counts the three key indicators of the average diagnostic accuracy, average F1-score, and average single inference delay of each method in the above test environment from a macro level. Table 3 summarizes the corresponding experimental results.

The data in Table 3 show that in the evaluation of the transfer performance of topological structures outside the training domain, the performance differences of various methods are significant. From the perspective of average accuracy, the experimental group reaches 0.915, which is much higher than 0.798 of control group 1, 0.846 of control group 2, and 0.878 of control group 3. This shows that the experimental group method has a stronger generalization ability in unseen topology scenarios and can precisely capture the implicit fault features in different topological structures. The average F1-Score of the experimental group is 0.922, which is also better than that of the control groups, which means that the experimental group achieves a better balance between the accuracy and comprehensiveness of fault identification. It can effectively reduce false positives and false negatives. In terms of average inference delay, although control group 1 has an advantage of 6.3ms, its accuracy, and F1-score are low, indicating that its fast response comes at the expense of diagnostic accuracy. The 12.4ms inference delay of the experimental group ensures high diagnostic performance while still having a fast response speed, which is more practical.

(4) Robustness test and operation and maintenance benefit simulation analysis

In the robustness test and operation and maintenance benefit simulation analysis, this study focuses on evaluating the anti-interference ability of the method in this paper in non-ideal environments and the operation and maintenance benefits after deployment. For robustness evaluation, this study simulates four types of abnormal transmission scenarios that may be encountered in the actual deployment of the online monitoring system of the drop-out fuse, including sensor noise interference, key data loss, timing delay disturbance, and data drift. In each of the above abnormal situations,

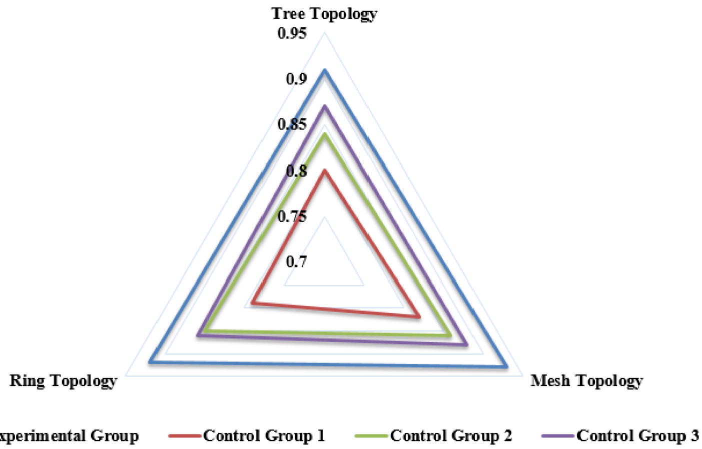

FIGURE 5. F1-scores of various methods for different topological structures

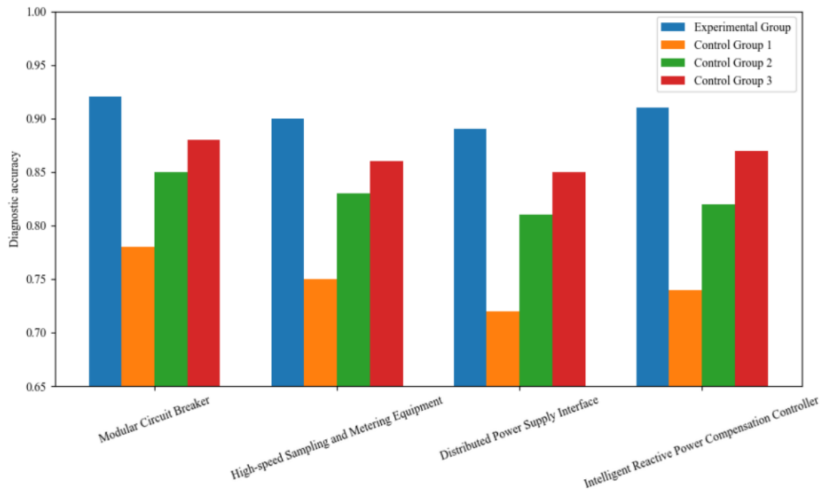

FIGURE 6. Diagnostic accuracy of each group for different equipment types

TABLE 3. Statistics of the transfer performance of each method on topologies outside the training domain

Method	Average Accuracy	Average F1-Score	Average Inference Delay (ms)
Experimental Group	0.915	0.922	12.4
Control Group 1	0.798	0.805	6.3
Control Group 2	0.846	0.852	24.1
Control Group 3	0.878	0.884	35.7

fault identification is performed using the experimental group method and the control group methods, and accuracy and recall are comprehensively considered as evaluation indicators. Table 4 lists the experimental results:

TABLE 4. Fault identification results of each method under different interferences

Exception Type	Method	Accuracy	Recall
Sensor Noise Interference	Experimental Group	0.926	0.901
Sensor Noise Interference	Control Group 1	0.741	0.685
Sensor Noise Interference	Control Group 2	0.878	0.832
Sensor Noise Interference	Control Group 3	0.895	0.870
Key Data Loss	Experimental Group	0.918	0.886
Key Data Loss	Control Group 1	0.731	0.670
Key Data Loss	Control Group 2	0.841	0.812
Key Data Loss	Control Group 3	0.881	0.854
Timing Delay Disturbance	Experimental Group	0.917	0.897
Timing Delay Disturbance	Control Group 1	0.743	0.692
Timing Delay Disturbance	Control Group 2	0.888	0.825
Timing Delay Disturbance	Control Group 3	0.909	0.873
Data Drift	Experimental Group	0.902	0.875
Data Drift	Control Group 1	0.748	0.685
Data Drift	Control Group 2	0.857	0.828
Data Drift	Control Group 3	0.879	0.853

The data in Table 4 shows that from the perspective of the fault identification performance of each method in different abnormal situations, the experimental group method shows significantly better identification ability than control groups in four typical interference environments and has strong robustness. Under the condition of sensor noise interference, the accuracy and recall of the experimental group method reach 0.926 and 0.901, respectively, which are significantly higher than those of the traditional threshold method and also better than the LSTM-based sequence modeling method and the GraphSAGE graph neural network method, reflecting strong noise resistance. In the face of the situation of key data loss, the accuracy of the experimental group method remains at 0.918, and the recall is 0.886, which is still better than the various methods in control groups, indicating that it can still stably extract fault correlation features under the condition of incomplete information. For timing delay disturbances, the experimental group method achieves an accuracy of 0.917 and a recall of 0.897, maintaining the highest level among all methods, reflecting its dynamic modeling advantages under time series disturbances. Even with typical data drift phenomena, the experimental group still achieves an accuracy of 0.902 and a recall of 0.875, which is better than the performance of 0.748 and 0.685 of the control group 1 method and higher than the results of other advanced graph models. Overall, the experimental group method shows higher fault identification precision and stability in multiple types of abnormal transmission environments, verifying its robustness advantages in real deployment scenarios.

In terms of intelligent operation and maintenance benefits, combined with the emergency repair management process of drop-out fuses in the power grid, a simulation system is built to simulate fault trigger response, operation and maintenance scheduling, manual dispatch, and other links. The simulation is based on the historical operation and maintenance data of the power grid in a certain area, and a unified fault trigger rule and dispatch strategy are set. During the simulation period, the average emergency repair response cycle (hours), daily dispatch frequency (times/day), and manpower input required per unit fault (person-hours/time) of each method under the same fault intensity are counted. Table 5 summarizes the simulation results:

TABLE 5. Changes in the repair cycle, dispatch frequency, and manpower input of each group for the triggering of drop-out fuse faults

Method	Average Repair Cycle (hours)	Average Daily Dispatch Frequency (times/day)	Manpower Input Per Unit Fault (person-hours/time)
Experimental Group	3.4	1.2	3.6
Control Group 1	5.7	2.4	6.2
Control Group 2	4.1	1.8	4.7
Control Group 3	3.7	1.4	4.0

From the operation and maintenance response results triggered by drop-out fuse faults, the experimental group method shows significant advantages in both emergency repair efficiency and manpower input. In terms of the average emergency repair cycle, the response time of the experimental group is 3.4 hours, which is significantly shorter than the 5.7 hours of control group 1, 4.1 hours of control group 2, and 3.7 hours of control group 3, indicating that the proposed method can complete fault diagnosis and dispatch response more quickly. In terms of the average daily dispatch frequency, the experimental group achieves 1.2 times/day, which is lower than that of the other groups, especially compared with the 2.4 times/day of control group 1, and its dispatch frequency is almost reduced by half, indicating that this method has practical results in reducing FPR and avoiding invalid dispatch. In terms of manpower input per unit fault, practical deployment has proven its feasibility, reducing the total manpower required for a single fault response process to 3.6 person-hours, which is also lower than that of control group 1 (6.2), control group 2 (4.7), and control group 3 (4.0), reflecting its potential in saving human resources and improving the economic efficiency of operation and maintenance. Overall, the experimental group method not only ensures the timeliness of fault response but also significantly optimizes resource allocation, providing reliable support for the intelligent operation and maintenance of drop-out fuses.

V. CONCLUSION

The paper proposes an intelligent fault diagnosis method that integrates a Hidden Markov Model (HMM) and Graph Embedding Technology, leveraging spatiotemporal correlation to capture dynamic node interactions and fault evolution stages, thereby surpassing the limitations of traditional static analysis. While this approach is currently limited by single-domain data and edge computing constraints on real-time response, its future trajectory—including adopting federated learning and digital twin technology—is highly relevant to improving the cross-regional, collaborative fault prediction and maintenance efficiency required for large, interconnected manufacturing networks. This approach uses the spatiotemporal correlation between the power grid's physical topology and equipment operation state, employing dynamic directed graph construction to capture real-time node interactions and a nonhomogeneous HMM to characterize fault evolution stages. This breaks the limitations of traditional static and isolated analyses.

However, current limitations include reliance on single-domain data, resulting in insufficient cross-regional collaborative diagnosis capability, and the real-time constraints of dynamic graph updates due to edge computing resources, potentially affecting millisecond-level fault response.

Future work will focus on:

1) Developing a federated learning distributed collaborative diagnosis model to balance data privacy and knowledge sharing;

2) Optimizing dynamic graph updates for edge deployment realtime performance using lightweight graph neural networks and adaptive sampling;
3) Integrating digital twin technology for a fault evolution simulation platform, enhancing interpretability and decision credibility through virtual-physical verification.

Expanding this method to active distribution networks with high new energy access will also be a key research area.

AUTHOR CONTRIBUTIONS

Conceptualization –Shoujun Bao, Shuhong Li, Ziyong Pan, Fuxiang Liang and Shenghai Han; methodology – Shoujun Bao, Shuhong Li, Ziyong Pan, Fuxiang Liang and Shenghai Han; investigation – Shoujun Bao, Shuhong Li, Ziyong Pan and Fuxiang Liang; writing-original draft preparation – Shoujun Bao, Shuhong Li, Ziyong Pan, Fuxiang Liang and Shenghai Han. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by,
1. State Grid Qinghai Electric Power Company Science and Technology Project Funding, 522824240001, Research on Digital Monitoring and Operation of Drop type Fuses in High cold Regions Based on Beidou and Micro motion Technology.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] J. Łukasiak, M. Wiśnios, and A. Rosiński, “Method for Evaluating the Effectiveness of Electrical Circuit Protection with Miniature Fuse-Links,” Energies, vol. 16, no. 2, pp. 960-988, 2023, doi: 10.3390/en16020960.

[2] M. Yousaf, A. Jalilian, K. M. Muttaqi, and D. Sutanto, “An adaptive overcurrent protection scheme for dual-setting directional recloser and fuse coordination in unbalanced distribution networks with distributed generation,” IEEE Transactions on Industry Applications, vol. 58, no. 2, pp. 1831-1842, 2022, doi: 10.1109/TIA.2022.3146095.

[3] Y. M. Nsaif, S. H. Lipu M, A. Ayob, Y. Yusof, and A. Hussain, “Fault detection and protection schemes for distributed generation integrated to distribution network: Challenges and suggestions,” IEEE Access, vol. 9, pp. 142693-142717, 2021, doi: 10.1109/ACCESS.2021.3121087.

[4] M. Taleb, B. Fani, and G. Shahgholian, “Fuse Maintenance to Solve Protection Problems in Active Distribution Networks in the Presence of Distributed Generation Sources,” Signal Processing and Renewable Energy, vol. 7, no. 4, pp. 1-12, 2023, [Online]. Available: https://oiccpress.com/spre/article/view/7900.

[5] A. Latorre, T. B. Soeiro, R. Geertsma, A. Coraddu, and H. Polinder, “Shipboard DC systems—A critical overview: Challenges in primary distribution, power-electronics-based protection, and power scalability,” IEEE Open Journal of the Industrial Electronics Society, vol. 4, pp. 259-286, 2023, doi: 10.1109/OJIES.2023.3294999.

[6] H. Bisheh, B. Fani, G. Shahgholian, and I. Sadeghkhani, “Fuse saving coordination scheme for active distribution systems: State-of-the-art and a novel quasi-voltage current based scheme,” IET Generation, Transmission & Distribution, vol. 18, no. 4, pp. 729-755, 2024, doi: 10.1049/gtd2.13110.

[7] A. Leśniak, M. Szewczyk, T. Daszczyński, and P. Piekarski, “Development tests and FEM simulations of maximum load characteristics of medium voltage fuse links,” Przegląd Elektrotechniczny, vol. 2025, no. 1, pp. 37-42, 2025, doi: 10.15199/48.2025.01.08.

[8] H. Peng, S. Li, L. Shangguan, Y. Fan, and H. Zhang, “Analysis of wind turbine equipment failure and intelligent operation and maintenance research,” Sustainability, vol. 15, no. 10, pp. 8333-8368, 2023, doi: 10.3390/su15108333.

[9] J. Li, M. Yu, J. Wu, J. Li, Y. Wan, and C. Xue, “Study of failure mode and failure load of a civil aircraft pylon doublelayer fuse pin,” International Journal of Crashworthiness, vol. 29, no. 6, pp. 1107-1117, 2024, doi: 10.1080/13588265.2024.2349356.

[10] Y. Miyaoka, M. Maeyama, Y. Yamano, and Y. Inada, “Internal State of Fuse Arc Differentiating Interruption Success and Failure in Fuse-Semiconductor DC Circuit Breaker,” IEEJ Transactions on Electrical and Electronic Engineering, vol. 19, no. 1, pp. 41-50, 2024, doi: 10.1002/tee.23928.

[11] N. A. F. M. N. Ong, M. A. Sadiq, M. S. M. Said, and J. S. Kristensen, “Fault tree analysis of fires on rooftops with photovoltaic systems,” Journal of Building Engineering, vol. 46, pp. 103752-103773, 2022, doi: 10.1016/j.jobe.2021.103752.

[12] Z. Tang and X. Jian, “Thermal fault diagnosis of complex electrical equipment based on infrared image recognition,” Scientific Reports, vol. 14, no. 1, pp. 5547-5563, 2024, doi: 10.1038/s41598-024-56142-x.

[13] H. E. A. Adam, J. K. Kimotho, and J. G. Njiri, “Multiple faults diagnosis for an industrial robot fuse quality test bench using deep-learning,” RESULTS in Engineering, vol. 17, pp. 101007-101023, 2023, doi: 10.1016/j.rineng.2023.101007.

[14] A. I. Khalyasmaa, A. I. Stepanova, S. A. Eroshenko, and P. V. Matrenin, “Review of the digital twin technology applications for electrical equipment lifecycle management,” Mathematics, vol. 11, no. 6, pp. 1315-1338, 2023, doi: 10.3390/math11061315.

[15] Y. Wu, H. N. Dai, and H. Tang, “Graph neural networks for anomaly detection in industrial Internet of Things,” IEEE Internet of Things Journal, vol. 9, no. 12, pp. 9214-9231, 2021, doi: 10.1109/JIOT.2021.3094295.

[16] J. Zhang and X. He, “Compound-fault diagnosis of integrated energy systems based on graph embedded recurrent neural networks,” IEEE Transactions on Industrial Informatics, vol. 20, no. 3, pp. 3478-3486, 2023, doi: 10.1109/TII.2023.3306935.

[17] J. Liang, K. Zhang, A. Al-Durra, and D. Zhou, “A multi-information fusion algorithm to fault diagnosis of power converter in wind power generation systems,” IEEE Transactions on Industrial Informatics, vol. 20, no. 2, pp. 1167-1179, 2023, doi: 10.1109/TII.2023.3271358.

[18] X. Huang, P. Li, Y. Wang, X. Ren, Z. Zhao, and G. Li, “Knowledge Graph-Augmented ERNIE-CNN Method for Risk Assessment in Secondary Power System Operations,” Energies, vol. 18, no. 8, pp. 2104-2119, 2025, doi: 10.3390/en18082104.

[19] T. Zhang, X. Yan, R. Zhang, Q. Ye, and J. Ma, “Distributed architecture of power grid asset management and future research directions,” Ieee Access, vol. 10, pp. 57588-57595, 2022, doi: 10.1109/ACCESS.2022.3179119.

[20] M. Xing, W. Ding, T. Zhang, and H. Li, “STCGCN: a spatio-temporal complete graph convolutional network for remaining useful life prediction of power transformer,” International journal of web information systems, vol. 19, no. 2, pp. 102-117, 2023, doi: 10.1108/IJWIS-02-2023-0023.

[21] A. Deep, S. Zhou, D. Veeramani, and Y. Chen, “HMM-based joint modeling of condition monitoring signals and failure event data for prognosis,” IEEE transactions on reliability, vol. 72, no. 3, pp. 878-888, 2022, doi: 10.1109/TR.2022.3193353.

[22] W. Li and C. Zhang, “A hidden markov model for condition monitoring of time series data in complex network systems,” IEEE Transactions on Reliability, vol. 72, no. 4, pp. 1478-1492, 2023, doi: 10.1109/TR.2023.3234907.