MULTITEMPORAL DATA FUSION AND REASONING IN POWER GRID FAULT DECISION SUPPORT

Yipu ZHANG[1]*, Qingbiao LIN[1], Shengmin QIU[1], Kun ZHANG[1], Xiaogang WU[1], Xin GAO[1]

As the scale of power grids continues to expand, the volume of grid-related information increases significantly. Diagnosing the operational state of power systems by integrating real-time data streams with historical fault data has become a key focus in the field of electric power. However, existing fault diagnosis methods, though relatively mature, still struggle to handle the large volume of dynamic data and fail to accurately identify fault information. To address this challenge, this paper puts forward a fault reasoning model based on Knowledge Graphs. The model integrates real-time power grid data with historical fault data to perform intelligent analysis and provide decision support through a reasoning mechanism. Experimental results show that the proposed hybrid algorithm achieves an accuracy of 96.3%, a recall rate of 0.99, and an F1-score of 0.90 after 250 training iterations. These performance metrics significantly outperform those of the comparison algorithms. The results of the fault reasoning experiments further demonstrate that the proposed model achieves higher reasoning accuracy across different types of fault causes, with an average accuracy reaching 0.93, showing strong generalization ability. These empirical findings prove that the model effectively fuses multitemporal data, improves the timeliness and accuracy of fault handling, and provides intelligent technical support for power grid fault management.

Keywords: Power grid fault; Multitemporal data; Fault diagnosis; Convolutional neural network; Dynamic time warping

1. Introduction

Electric energy plays a vital role in daily life and the development of industrial and agricultural sectors. Ensuring the safety of power equipment and operations is essential for the efficient supply of electricity [1]. In recent years, to improve the reliability of power supply, fault diagnosis methods for power grids have become a key topic of academic research [2]. The rapid development of deep learning technologies has brought significant progress to the intelligentization of fault diagnosis methods. However, as power grids grow larger, the amount of operational data increases drastically. Traditional fault diagnosis methods can no longer meet the demand for real-time processing of large-scale data [3-4]. Therefore,

[1] Power Dispatching Control Center, China Southern Power rid Company Ltd. (CSG), Guangzhou, Guangdong 510000, China
*Correspondence author's e-mail: zyp1230813@163.com

there is an urgent need for an efficient and intelligent fault diagnosis approach to support decision-making for power system operators. To address this need, this study introduces an improved Graph Convolutional Network-Graph Attention Network (GGCN) based on Knowledge Graph (KG) technology. To solve the problem of misaligned sequences between real-time data streams and historical data, a Dynamic Time Warping (DTW) algorithm is further introduced, and its computational efficiency is optimized using a Sliding Window (SW). These techniques are combined to form a model named Sliding Window Dynamic Time Warping–Graph Convolutional Network-Graph Attention Network (SW-DTW-GGCN). This model aims to effectively integrate real-time data streams and historical data from power grid equipment, enhance the real-time performance and accuracy of fault diagnosis, and provide practical value for power grid fault decision-making.

2. Related works

In recent years, the research on decision support models for power grid faults has become a central focus in the academic community. Liu et al. constructed a KG for power grid faults by performing knowledge extraction and other processes to enhance the efficiency and accuracy of fault diagnosis. They then formed fault records through entity extraction and developed a fault diagnosis system based on the KG [5]. Kandasamy and Jaganathan combined continuous wavelet transform with deep learning methods to build an intelligent decision support scheme for microgrid faults. Their approach extracted key information from time series data using continuous wavelet transform and then identified and classified faults using deep learning techniques [6]. Biswal et al. pointed out the difficulty in detecting high-impedance faults in microgrids. To address this issue, they introduced dual-tree complex wavelet transform and data mining methods. They extracted wavelet features to separate high-impedance fault data, then conducted fault detection and classification, and finally tested the method through experimental design [7]. Zhu et al. aimed to solve the problems of high false alarm rates and low efficiency in traditional fault diagnosis methods. They modeled power grid state data using a computer simulation system, categorized various fault types, and applied an Extreme Learning Machine (ELM) model to perform diagnosis, testing its performance with multiple fault datasets [8]. Zhong et al. focused on improving the accuracy of reasoning strategies for power transformer faults. They first analyzed the structural similarity between target faults and historical cases, discussed attribute similarities, and finally calculated overall similarity to obtain reasoning strategies from historical cases [9].

Data fusion technologies have been widely applied across various domains. Kuras et al. addressed the problem of fusing hyperspectral and LiDAR data in urban environments. They proposed an unsupervised optimization method for segmenting spectral components and LiDAR data and further updated regional maps by manually supplementing uncertain objects and materials [10]. Huang et al. applied data fusion methods to traffic flow prediction by integrating convolutional neural networks and predictive recurrent neural networks to process dynamic spatiotemporal data. They also used a multi-opinion mechanism to extract and fuse key information [11]. Zhang et al. developed an intelligent surveillance approach to dynamically fuse video data and Automatic Identification System (AIS) information in order to manage waterways more effectively. They introduced a filter-based method for fusing video and AIS data to enable real-time processing of large-scale information [12]. Gulshin and Kuzina proposed a time fusion transformer model for wastewater treatment. This model combined historical data with current predictions and was further developed into a neural basis expansion analysis model to provide interpretable time series forecasting [13]. Melgar et al. designed a predictive model for streaming sequence data based on information fusion. They constructed the model using historical values and their nearest future values. As real-time data streams arrived, the model was updated through a buffer containing streaming data [14].

In summary, most existing studies on power grid fault decision support systems have focused on improving fault detection methods. At the same time, research on data fusion techniques has achieved meaningful progress, and their practical value has been widely demonstrated. However, few studies have addressed the fusion of real-time and historical data to improve the timeliness and efficiency of power grid fault diagnosis models. In response to this gap, this paper introduces an improved GGAT based on KG technology. To enable the effective integration of real-time and historical data, the study further incorporates DTW algorithm and optimizes it using a SW, ultimately forming a SW-DTW-GGAT model. The proposed improvements are expected to enhance the efficiency and accuracy of power grid fault diagnosis and provide intelligent support for operation and maintenance decision-making.

3. Construction of the multitemporal power grid fault data fusion and reasoning model

3.1. Construction of the KG-GGAT fault decision reasoning model

Due to the diversity of power grid fault patterns and fault types, and the strong correlation among fault locations, making fault decisions requires substantial manual effort to record abnormal information. This process involves searching for historical fault causes and corresponding solutions, which makes the decision-making process complex and increases the cost for grid departments. The KG technology shows high efficiency in organizing, managing, and applying large-scale data, making the handling and application of knowledge much easier [15]. An effective KG contains comprehensive fault entity types and fault relationship links, which enables deep mining of fault knowledge in power grids and offers feasible solutions for fault diagnosis and strategy reasoning. The construction process of the KG is shown in Fig. 1.

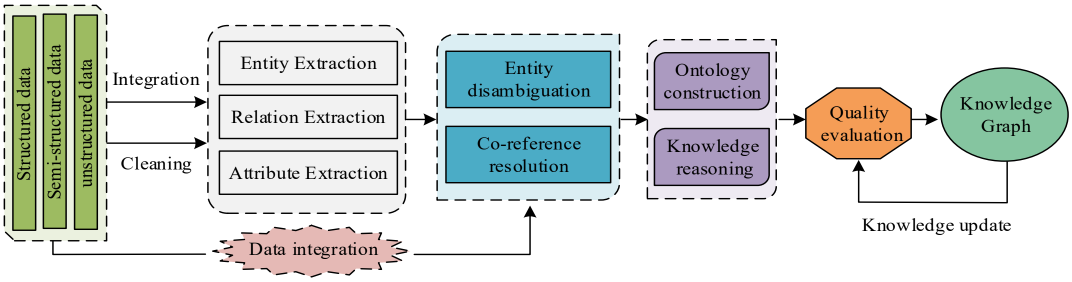


Fig. 1. Schematic diagram of the KG construction process.

As shown in Fig. 1, the data must be preprocessed before constructing the KG. This includes operations such as field alignment, missing value imputation, and outlier correction to ensure data consistency [16–17]. Then, data entities, relationships, and attributes are extracted to complete the knowledge extraction process, followed by establishing associations between entities based on the characteristics of power grid faults. Each entity is a data model, and is represented by S, as shown in Equation (1).

$$S = \langle C, H, R, A \rangle \tag{1}$$

In Equation (1), C denotes the association among entities, representing a set of relationships. H represents the set of categorical relationships between fault entities, R is the set of non-categorical relationships, and A is the set of rules. To solve the heterogeneity among different knowledge sources, knowledge fusion techniques are adopted, with knowledge fusion and entity linking being the two key steps. The equation for knowledge fusion is shown in Equation (2).

$$RankScore(a,k) = \sqrt{Semantic(a,k) * Lexical(a,k)} \tag{2}$$

In Equation (2), a represents book knowledge, and k stands for keywords. After knowledge fusion, the data is further processed and finally stored as a KG. To ensure the quality of the graph, a dynamic maintenance mechanism must be established. First, the existing knowledge is verified, and then new knowledge is continuously integrated to maintain the timeliness and completeness of the graph. The KG is defined as shown in Equation (3).

$$KG = \langle E, R, G \rangle \tag{3}$$

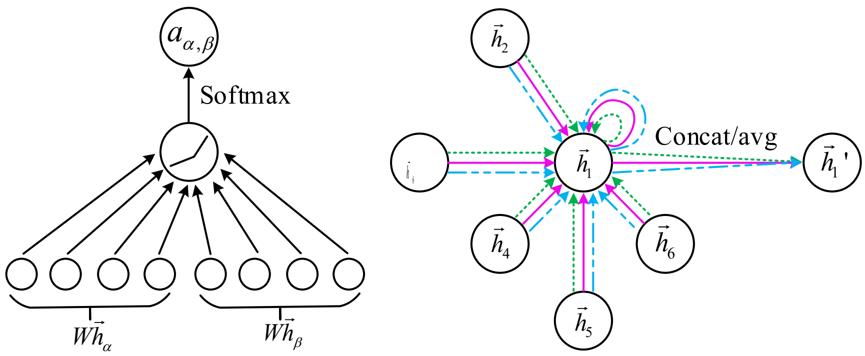
In equation (3), E denotes the entity set, and G denotes the triplet fact set.

The construction of the KG depends on the analysis of historical power grid fault cases. However, problems such as incomplete entity information and sparse expert knowledge often arise. To address this, this study introduces GCN. First, entities and relationships in the KG are vectorized. Then, a reasoning and decision-making algorithm is used to reduce the dimensionality of the high-dimensional vectors. Finally, the results are applied to fault decision reasoning tasks in power grids. The core operation of GCN is based on spectral graph convolution using the symmetrically normalized Laplacian matrix, as shown in Equation (4).

$$\varsigma^{Laplace} = \tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2} = \tilde{D}^{-1/2}\left(\bar{A}+1\right)\tilde{D}^{-1/2} \tag{4}$$

In Equation (4), $\bar{A}$ denotes the adjacency matrix, $\bar{A}$ represents the original adjacency matrix, I is the identity matrix, and $\bar{D}$ represents the degree matrix of $\bar{A}$. This layer takes as input the initial node feature matrix and adjacency matrix of the knowledge graph, with its core operation being the computation of the normalized Laplacian matrix and feature propagation. While GCN performs well in reasoning over static graphs, it cannot handle reasoning tasks over dynamic graphs, and it struggles to assign different learning weights to neighboring entities. To solve this problem, this study further introduces GAT, which integrates fault entity information within neighborhoods by applying an attention mechanism. The principle of GAT is shown in Fig. 2.

As shown in Fig. 2, the input to the GAT network consists of node features output by the GCN. Its core operation primarily involves two processes: calculating attention coefficients and performing weighted summation. The equation for calculating attention coefficients is shown in Equation (5).



(a) Compute the attention coefficient (b) Weighted sum

Fig. 2. Schematic diagram of the GAT principle

$$a_{\alpha,\beta} = soft\max\left(Leaky\,\mathrm{Re}LU\left(e_{\alpha,\beta}\right)\right) = \frac{\exp\left(Leaky\,\mathrm{Re}LU\left(e_{\alpha,\beta}\right)\right)}{\sum_{k\in N_\alpha}\exp\left(Leaky\,\mathrm{Re}LU\left(e_{\alpha,k}\right)\right)} \tag{5}$$

In Equation (5), $Leaky\,\mathrm{Re}LU$ denotes a nonlinear activation function, $a_{\alpha,\beta}$ represents the attention coefficient, α is the target entity, N_α denotes the set to which the entity belongs, and β represents neighboring entities. The equation for the weighted summation is shown in Equation (6).

$$Q_\alpha = \left\| \sigma\left[\frac{1}{K}\sum_{k=1}^{K}\sum_{\beta\in N_\alpha} a_{\alpha,\beta} W Q_\beta\right]\right. \tag{6}$$

In Equation (6), Q_α denotes the embedding of entity i, K is the number of attention operations. W is the shared parameter matrix, and σ represents a nonlinear activation function. In summary, based on KG technology, this study introduces both GCN and GAT to construct the GGAT fault decision reasoning model. The detailed process is shown in Fig. 3.

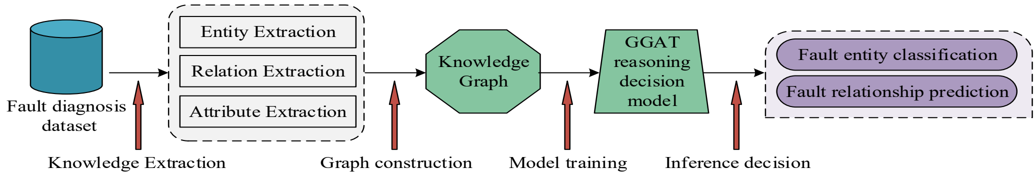


Fig. 3. Flowchart of the GGAT fault decision reasoning model based on KG.

As shown in Fig. 3, the proposed GGAT fault decision reasoning model is constructed based on historical power grid fault data. A KG is built by extracting fault entities, fault relationships, and fault knowledge triples. Then, the GGAT model is trained, and fault data requiring reasoning is input into the model. Finally, the model performs tasks such as classifying fault entity types and predicting fault relationships.

3.2. Construction of the multitemporal data fusion fault reasoning model

Although the GGAT fault decision reasoning model can perform classification and prediction of power grid faults based on the knowledge base, the integration of real-time and historical data often faces a sequence alignment problem due to differences in their time scales. The DTW algorithm automatically scales time series through dynamic programming to evaluate the similarity between two sequences and align them, thereby eliminating phase differences [18]. Therefore, this study applies DTW to integrate real-time data with historical fault data, aiming to improve the accuracy of fault decision-making. Suppose there are two time series, $\tilde{x}_i = (\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)$ and $\tilde{y}_j = (\tilde{y}_1, \tilde{y}_2, ...\tilde{y}_m)$, with lengths denoted by n and m respectively. DTW calculates the cumulative distance matrix D by computing the Euclidean distance between the two series. The equation for each element in the matrix is shown in Equation (7).

$$D\left(\tilde{x}_i, \tilde{y}_j\right) = \left(\tilde{x}_i - \tilde{y}_j\right)^2 \tag{7}$$

The goal of DTW is to find the optimal warping path that minimizes the distance between sequences $\tilde{x}_n$ and $\tilde{y}_m$, as shown in Equation (8).

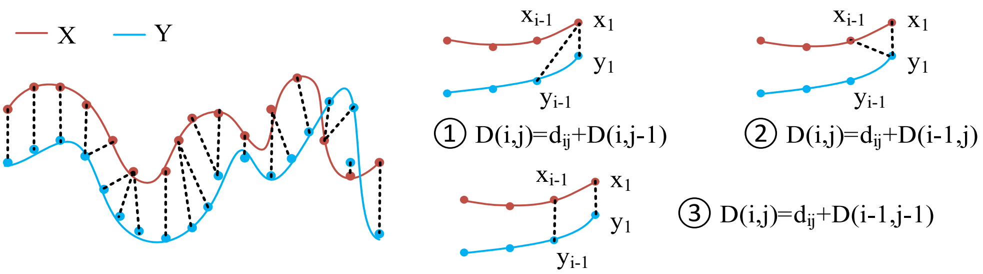
$$DTW\left(\tilde{x}_n, \tilde{y}_m\right) = \min\left(\sum_{o=1}^{O} L\left(\omega_o\right)\right) \tag{8}$$

In Equation (8), $L\left(\omega_o\right)$ represents the local distance matrix, and o represents the elements of the optimal path. The warping path is denoted as T, where the o-th element is expressed in Equation (9).

$$\omega_o = \left(i, j\right)_\omega,\ \omega_o \in \mathbb{R}^2,\ T = \left(\omega_1, \omega 2, \cdots, \omega_o, \cdots, \omega_K\right),$$
$$\max\left(n, m\right) \le O \le n + m - 1 \tag{9}$$

In Equation (9), $\left(p, q\right)$ represents the coordinates in the distance matrix. The principle of how DTW processes two time series is illustrated in Fig. 4.

As shown in Fig. 4(a), the two lines represent two different time series, and the connecting points indicate similar data points. DTW constructs an alignment path between the sequences to minimize the cumulative distance between corresponding points, quantifying the similarity between unequal-length sequences [19]. In the process of dynamic time warping, the algorithm selects the optimal path based on the principle of minimum distance. As shown in Fig. 4(b), it chooses the best option among three splitting methods to match similar points. The points on the warping path must follow monotonicity, which is expressed in Equation (10).



(a) Finding similar points of two time series

(b) Optimization of distance between similarities

Fig. 4. Schematic diagram of the DTW algorithm.

$$\omega_1 = \left(1, 1\right), \omega_o = \left(p, q\right) \tag{10}$$

The boundary condition ensures that the first and last elements of the two sequences correspond, as shown in Equation (11).

$$u_1 \le u_2 \le \cdots \le u_p,\ v_1 \le v_2 \le \cdots \le v_q \tag{11}$$

The step size must be set to avoid missing elements, and its mathematical expression is given in Equation (12).

$$\left(\omega_{o+1}-\omega_{o}\right)\in\left\{(1,0),(0,1),(1,1)\right\},\quad for\ o\in\left[1:O-1\right] \tag{12}$$

To meet the above constraints, the cumulative loss matrix as shown in Equation (13).

$$L(u,v)=\min\left\{L(u-1,v-1),L(u-1,v),L(u,v-1)\right\}+D\left(\tilde{x}_{u},\tilde{v}_{v}\right) \tag{13}$$

In Equation (13), D represents the cumulative loss matrix, and $u\in[1:p],v\in[1:q]$ and $D\left(\tilde{x}_{u},\tilde{v}_{v}\right)$ denote the loss values. However, DTW matches two time series based solely on the similarity of amplitude values between individual points, ignoring the similarity in their trends. This often results in many singular points in the outcome. To address this issue, this study further introduces SW. Real-time data streams are first segmented through sliding windows and then aligned with historical data in segments using DTW, achieving sequence alignment [20]. This module takes long real-time data streams and historical reference sequences as inputs. Its core functionality involves segmenting the real-time stream into windows and performing DTW alignment with the historical sequence for each segment. For a time series $\bar{X}$, assuming the SW window length is b and the step size is s, the m-th window is shown in Equation (14).

$$\mathrm{E}_{m}-\left\{\bar{x}_{m\cdot s},\bar{x}_{m\cdot s+1},\cdots,\bar{x}_{m\cdot s+b-1}\right\} \tag{14}$$

The resulting SW-DTW algorithm calculates the similarity loss function between the two sequences, as shown in Equation (15).

$$L\left(\tilde{x}_{u},\tilde{y}_{v}\right)=b(\theta)\cdot\mathrm{E}(\theta) \tag{15}$$

In Equation (15), θ represents the window length, and $\mathrm{E}(\theta)$ is calculated using the equation in Equation (16).

$$\mathrm{E}(\theta)=\alpha\left(\bar{X}_{\left[a-\frac{\theta}{2}:a+\frac{\theta}{2}\right]}-\bar{Y}_{\left[e-\frac{\theta}{2}:e+\frac{\theta}{2}\right]}\right)+(1+\alpha)\left(\Delta\left[\bar{X}\right]_{\left[a-\frac{\theta}{2}:a+\frac{\theta}{2}\right]}-\Delta\bar{Y}_{\left[e-\frac{\theta}{2}:e+\frac{\theta}{2}\right]}\right)^{2} \tag{16}$$

In Equation (16), a and e represent the index positions of the sequences. In summary, this study introduces the SW-DTW algorithm to address the temporal misalignment between real-time and historical data in fault diagnosis and builds the SW-DTW-GGAT fault decision reasoning model. The model structure is shown in Fig. 5.

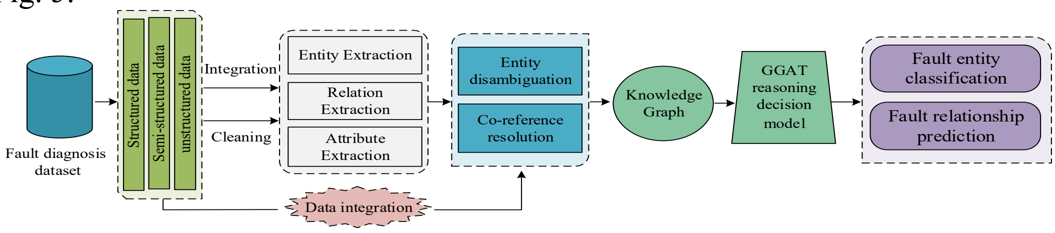


Fig. 5. KG-GGAT fault decision reasoning model flow chart.

As shown in Fig. 5, the KG-GGAT fault decision reasoning model first dynamically segments the real-time data stream using a sliding window to output dynamic window segments. Subsequently, it performs temporal alignment between real-time and historical data via SW-DTW, computing the cumulative loss matrix, optimal smoothing path, and global temporal alignment mapping table. This ultimately generates unified-timeline device status data. This data is then fed into the KG-GGAT fault inference model, where fault knowledge triples undergo vectorization to derive initial feature vectors for entities and relationships. The final decision is made through the GGAT inference model. This process computes and outputs, layer by layer, node feature matrices propagated through the graph structure and attention weight coefficients. It ultimately generates the classification probability distribution for fault entity types and the predicted probability for fault relationships, thereby completing the fault decision task.

4. Validation of the power grid multitemporal data fusion and fault diagnosis model

4.1. Effectiveness verification of the SW-DTW-GGAT algorithm

To verify the effectiveness of the proposed SW-DTW-GGAT fault diagnosis algorithm, this study conducted comparative experiments. In terms of experimental setup, high-performance hardware and software environments were selected. The configuration is shown in Table 1.

Table 1

Experimental environment and configuration

/	Category	Parameters
Hardware	CPU	Intel Xeon E52680 v4
	GPU	NVIDIA A100, 80GB PCIe
	RAM	64GB DDR4, 3200MHz
	Storage	512GB SSD, 2TB SATA SSD
Software	Operating system	Ubuntu 20.04 LTS, 64-bit
	Programming language	Python 3.8

As shown in Table 1, the experiments adopted high-performance computing equipment, with the CPU being Intel Xeon E5-2680 v4 and the GPU being NVIDIA A100. The algorithm was implemented using Python 3.8. The experimental dataset originates from the 2023-2024 operational maintenance records of a 220kV smart substation, encompassing comprehensive monitoring data of primary equipment across two main transformers, five transmission lines, and four bus sections. Following data anonymization and annotation, 17 field-verified typical fault events and their corresponding normal operating conditions were extracted, establishing a multi-source fault diagnosis benchmark dataset featuring temporal, graph-structural, and physically authentic characteristics. This dataset comprises approximately 85,000 valid time-series samples spanning a total duration of roughly 23,000 minutes. It contains synchronized monitoring records for 12 key physical quantities including voltage, current, power, and temperature, fully meeting the training and validation requirements for neural networks. The data has undergone linear normalization and had geographic identifiers removed, fully complying with industrial privacy requirements. A sample data input is shown in Table 2.

Table 2

Data Input Sample

Relative Time (ms)	Measurement Point (Example)	Value (p.u.)	Data Window	Label
-100	Line L3 Current	0.95	Historical	Normal
$0(t_0)$	Line L3 Current	0.97	Real-time Start	Fault a
+50	Line L3 Current	8.62	Real-time	Fault a
-100	Bus A Voltage	1.01	Historical	Normal
$0(t_0)$	Bus A Voltage	1.00	Real-time Start	Fault a
+10	Bus A Voltage	0.65	Real-time	Fault a

As shown in Table 2, the algorithm inputs are based on the fault onset time t0. The historical data window is defined as the steady-state data from [t_0-100ms, t_0] prior to the fault, used to establish a reference baseline. The real-time data window is defined as the transient data from [t_0, t_0+200ms] after the fault, used for capturing fault characteristics. First, the study performed statistical analysis on the loss function and fault diagnosis accuracy of the algorithm. GAT and GGAT were selected as comparison algorithms, and the results are shown in Fig. 6.

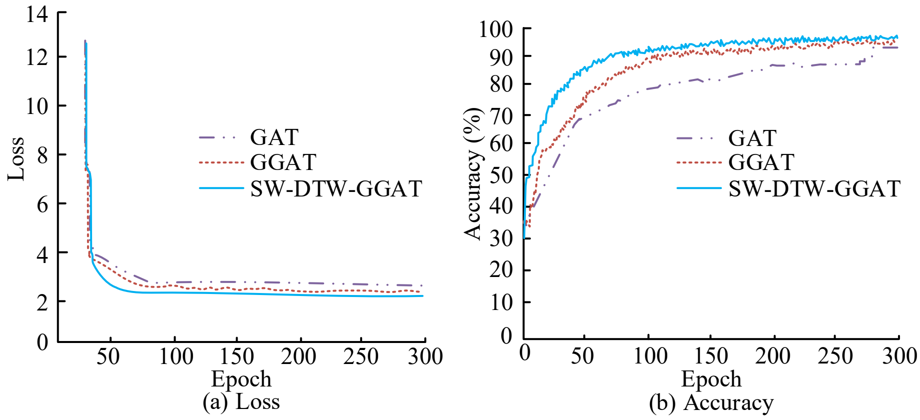
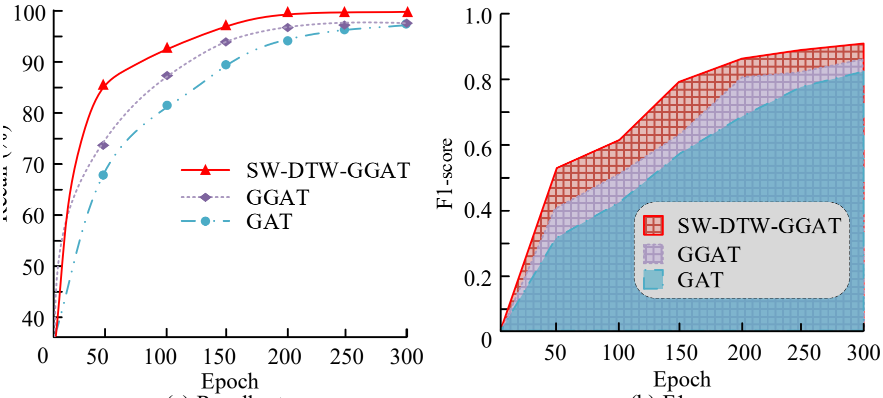


Fig. 6. Statistical results of loss function and fault diagnosis accuracy.

As shown in Fig. 6(a), the loss functions of all three algorithms showed a decreasing trend as the number of training iterations increased. Within the range of 0–50 iterations, the loss values dropped rapidly. When the training iteration reached 75, the loss value of the proposed SW-DTW-GGAT model was 2.3, and it gradually stabilized with more training. At the same time, the loss values for GAT and GGAT were 2.8 and 2.5, respectively, both higher than that of the proposed algorithm. According to Fig. 6(b), the classification accuracy of the three algorithms increased with training iterations. The accuracy improvement of GAT was relatively slow. The accuracy of GGAT began to stabilize around 250 iterations, while the SW-DTW-GGAT algorithm reached its maximum accuracy of 96.3% at 200 iterations. At that point, the accuracies of GAT and GGAT were only 80.6% and 90.4%, respectively. These experimental results indicated that the performance of the SW-DTW-GGAT algorithm surpassed that of the comparison algorithms. To further verify the superiority of the proposed algorithm, this study evaluated the recall rate and F1-score of four algorithms. The results are shown in Fig. 7.

In Fig. 7(a), the recall rate increased with the number of training iterations. At 200 iterations, the SW-DTW-GGAT algorithm reached its highest recall rate. At the same point, the recall rates of GAT and GGAT were 93.7% and 96.3%, respectively, both lower than that of the proposed algorithm. According to Fig. 7(b), throughout the entire training range, the F1-score of GAT remained the lowest, followed by GGAT. The F1-score of the SW-DTW-GGAT algorithm was consistently the highest. These results demonstrated that the SW-DTW-GGAT algorithm performed well and achieved better overall performance than the comparison algorithms.



(a) Recall rate (b) F1-score

Fig. 7. Statistical results of recall rate and F1-score.

4.2. Application performance of the SW-DTW-GGAT power grid fault reasoning model

After verifying the effectiveness of the SW-DTW-GGAT algorithm, the study further evaluated the fault diagnosis model based on real power grid data. The study selected actual power grid data from the 2023-2024 operation and maintenance records of a 220kV smart substation. Based on specific primary equipment failure scenarios, five typical fault types were defined: Type “a” (simple fault) refers to a transient metallic ground fault on Phase A of a line. Type “b” (Complete Protection Failure) refers to the failure of the corresponding digital bus protection device to operate due to power loss during a short circuit on the 110kV Bus C. Type “c” (Incomplete Protection Failure) refers to the logical lockout of the longitudinal differential protection during a line fault due to abnormal channel synchronization signals. Type “d” (Switch Failure) refers to a circuit breaker failure to trip due to a jammed operating mechanism during a low-voltage side fault on main transformer T1. Type “e” (Progressive Fault) collectively refers to the gradual process where insulation degradation on the high-voltage side bushing of main transformer T2 evolves from partial discharge to a ground flash-over. Four widely used models were selected for comparison: Graph Sample and Aggregate Network (GraphSAGE), Spatial-Temporal Graph Neural Network (STGNN), Graph Transformer, and Knowledge Graph-enhanced Graph Neural Network (KG-GNN). The study first analyzed the classification results of each model on different fault entities. The results are shown in Fig. 8.

As shown in Fig. 8, the SW-DTW-GGAT model achieved the highest classification accuracy across all five fault entities. For the classification of simple fault entities, GraphSAGE yielded the lowest accuracy at 0.80. STGNN and Graph Transformer followed with accuracies of 0.87 and 0.88, respectively. KG-GNN achieved a slightly higher accuracy of 0.90. The proposed SW-DTW-GGAT model achieved the highest accuracy of 0.91. For the classification of the “partial failure of protection system” fault entity, the SW-DTW-GGAT model reached an accuracy of 0.95, significantly higher than those of the comparison models. To further evaluate model effectiveness, the study assessed the reasoning accuracy of each model for fault causes and plotted confusion matrices. The results are shown in Fig. 9.

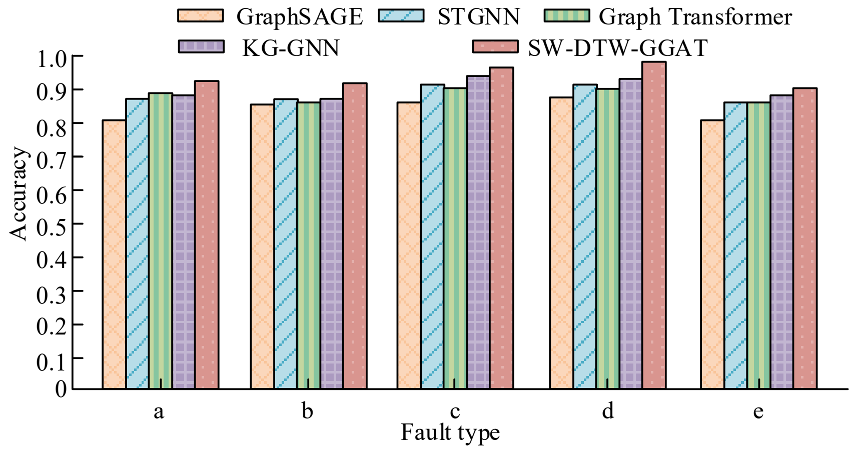


Fig. 8. Comparison of fault identification results across different models.

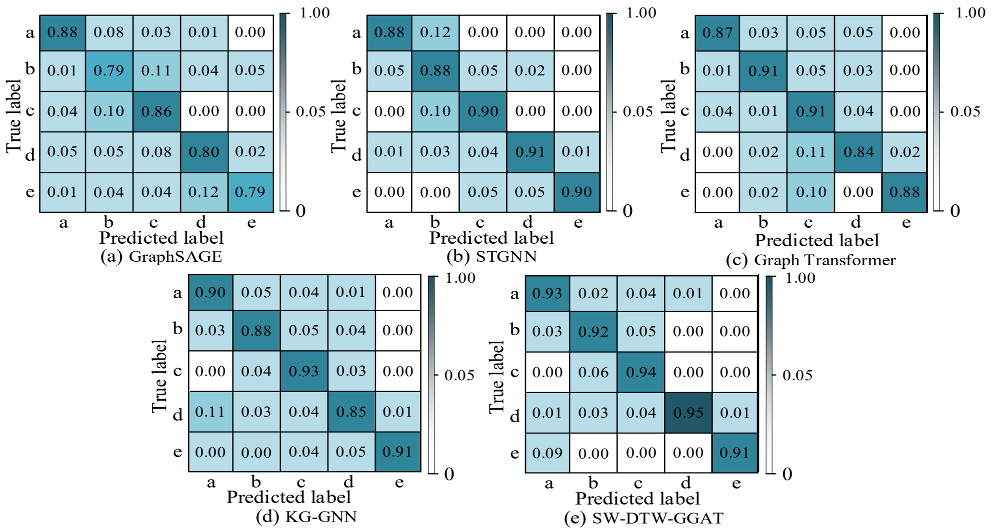


Fig. 9. Accuracy of fault cause reasoning across different models.

As shown in Fig. 9(a), the reasoning accuracy of GraphSAGE for all five fault causes remained below 0.90, with an average accuracy of only 0.82. Fig. 9(b) showed that the STGNN model achieved accuracies of 0.90 for both “partial failure of protection system” and “progressive fault,” and 0.91 for “failure of switching device.” Fig. 9(c) indicated that the Graph Transformer model performed slightly worse than STGNN, with an average accuracy of 0.88. As shown in Fig. 9(d), the KG-GNN model achieved its highest reasoning accuracy of 0.93 for the "partial failure of protection system" cause, followed by the "progressive fault." Fig. 9(e) demonstrated that the SW-DTW-GGAT model outperformed all other models in reasoning accuracy for all five fault causes. Notably, it achieved an accuracy of 0.95 for the cause "failure of switching device," which was 0.15, 0.04, 0.11, and 0.10 higher than the four comparison models, respectively. The average accuracy of the proposed model was 0.93, the highest among all models. These results confirmed that the proposed model effectively predicted the causes of the five types of power grid faults and exhibited strong overall reasoning performance.

5. Conclusion

In recent years, the scale of power grids expanded continuously, and the volume of system data increased exponentially. Traditional fault reasoning methods could no longer meet the demands of analyzing such massive data, which made it urgent to adopt a more efficient and accurate fault reasoning model. Based on this, the study first constructed a power grid fault KG. Then, it integrated the graph reasoning networks GCN and GAT to form the KG-based GGAT fault decision reasoning model. To effectively fuse real-time data streams with historical fault data, the study introduced SW to improve the DTW algorithm, thereby developing the SW-DTW-GGAT fault diagnosis model. According to the results of comparative experiments, the proposed algorithm achieved the lowest loss function value and the highest classification accuracy when the number of training iterations reached 250. At that point, the recall rate and F1-score of the algorithm were 0.99 and 0.90, respectively, both significantly higher than those of the comparison algorithms. The average classification accuracy of the SW-DTW-GGAT model for five fault entities reached 0.91, and the average reasoning accuracy for five fault causes reached 0.93, both notably higher than the comparison models. These experimental results demonstrated that the proposed SW-DTW-GGAT fault reasoning model could accurately identify fault entities and infer fault causes in power grids, showing strong reliability and practical value. Although the model proposed in this study achieved good reasoning performance, the current dataset mainly covered high-frequency fault types, and the identification and reasoning of complex faults have not been deeply explored. Future research needs to consider the complexity and diversity of fault scenarios and optimize the model design to enhance its generalization ability across a wider range.

Acknowledgments

This work was supported by the Research and Application of Power Grid Abnormal Scheduling Decision Technology Driven by Data and Knowledge

Collaboration, with the project number 000005KC24010014 (Science and Technology Project number: ZDKJXM20240015).

REFERENCES

[1] Strejoiu C V, Osman M G, Cernat A C. (2023) Towards sustainable transportation: modeling and simulation of pv panel implementation on national highways for charging electric vehicles which leads to mitigate carbon emission. UPB Sci. Bull., Series. C 85: 321-334. https://doi.org/10.36869/upb/scibullc/2023-85-321-334

[2] Rezk N G, El-Sehiemy R A, Attia A F, et al. (2025) Accurate multi-phase unsupervised and supervised approach to fault detection in power transmission networks. Neural Comput Appl 37: 19751-19772. https://doi.org/10.1007/s00521-025-11422-z

[3] LI J, LU X, ZHAO S, et al. (2023) Design of real-time monitoring system for power device status based on optical fiber sensor network. UPB Sci. Bull., Series. C 85: 235-244. https://doi.org/10.36869/upb/scibullc/2023-85-235-244

[4] Mahendravarman I, Ragavendiran A, Chithradevi S A. (2024) Improved power quality for photovoltaic grid integration power system using an intelligent controller fed SL-SC boost converter supplied reduced switch cascade multilevel inverter. Electr Eng 106: 419-441. https://doi.org/10.1007/s00202-023-01981-6

[5] Liu W, Zhou Z, Cai K, et al. (2023) Application of knowledge graph in smart grid fault diagnosis. Appl Math Nonlin Sci 8: 349-360. https://doi.org/10.2478/amns.2021.2.00258

[6] Kandasamy L M, Jaganathan K. (2023) Intelligent fault diagnosis using deep learning for a microgrid with high penetration of renewable energy sources. Electr Pow Compo Sys 51: 332-350. https://doi.org/10.1080/15325008.2023.2168091

[7] Biswal T, Parida S K, Mishra S. (2023) A DT-CWT and Data mining based approach for High Impedance Fault Diagnosis in Micro-grid System. Procedia Comput Sci 217: 1570-1578. https://doi.org/10.1016/j.procs.2022.12.357

[8] Zhu Y, Su R, Wu J, et al. (2025) Multiple fault diagnosis method for regional power grid based on DTS simulation system. Int J Energy Technol Policy 20: 89-94. https://doi.org/10.1504/IJETP.2025.144302

[9] Zhong Y, Yuan T, Chen B, et al. (2024) Research on power transformer maintenance strategy based on case-based reasoning. J Phys Conf Ser 2788: 012045. https://doi.org/10.1088/1742-6596/2788/1/012045

[10] Kuras A, Brell M, Liland K H, et al. (2023) Multitemporal feature-level fusion on hyperspectral and LiDAR data in the urban environment. Remote Sens-Basel 15: 632-653. https://doi.org/10.3390/rs15030632

[11] Huang X, Jiang Y, Tang J. (2023) MAPredRNN: multi-attention predictive RNN for traffic flow prediction by dynamic spatio-temporal data fusion. Appl Intell 53: 19372-19383. https://doi.org/10.1007/s10489-023-04494-8

[12] Zhang R, Zhao C, Liang Y, et al. (2025) Edge-Based Dynamic Spatiotemporal Data Fusion on Smart Buoys for Intelligent Surveillance of Inland Waterways. J Mar Sci Eng 13: 220-251. https://doi.org/10.3390/jmse13020220

[13] Gulshin I, Kuzina O. (2024) Optimization of Wastewater Treatment Through Machine Learning-Enhanced Supervisory Control and Data Acquisition: A Case Study of Granular Sludge Process Stability and Predictive Control. Automation 6: 2-22. https://doi.org/10.3390/automation6010002

[14] Melgar-García L, Gutiérrez-Avilés D, Rubio-Escudero C, et al. (2023) A novel distributed forecasting method based on information fusion and incremental learning for streaming time series. Inform Fusion 95: 163-173. https://doi.org/10.1016/j.inffus.2023.02.023

[15] Pu T, Tan Y, Ma Z, et al. (2025) Graph Computing Based Knowledge Reasoning in Electric Power System Considering Knowledge Graph Sparsity. CSEE J POWER ENERGY 5: 2083-2093. https://doi.org/10.17775/CSEEJPES.2022.04120

[16] Yasunaga M, Bosselut A, Ren H, et al. (2022) Deep bidirectional language-knowledge graph pretraining. Adv Neural Inform Process Syst 35: 37309-37323. https://doi.org/10.48550/arXiv.2210.09338

[17] Yu J, Tongxin W, Zhiwei Y, et al. (2024) Power equipment defect prediction based on time series knowledge graph. SIDR 1: 105-114. https://doi.org/10.59782/sidr.v2i1.65

[18] Teng Y, Wang G, He C, et al. (2025) Optimization of dynamic time warping algorithm for abnormal signal detection. Int J Data Sci Anal 19: 115-127. https://doi.org/10.1007/s41060-023-00446-0

[19] MEI B X, LYUBIN M A, YIN Jie, et al. (2025) A short-term photovoltaic power prediction method based on improved spectral clustering-DTW and Stacking fusion. HTL 3: 288-299. https://doi.org/10.3772/j.issn.1006-6748.2025.03.009

[20] Cong Giao B, Tuan Anh D. (2024) Subsequence Join in Streaming Time Series under Dynamic Time Warping. Vietnam J Comput Sci 11: 241-274. https://doi.org/10.1142/S2196888823500203