EAI Endorsed Transactions
on Energy Web
Research Article EAI.EU

GNN-Based Method for Data Flow Path Completion in Electric Power Information Networks

Xing Liu[1], Zhengqi Wang[1,*], Wangyong Guo[1], Liang Guo[1], Shuang Chen[2] and Yizhou Chen[1]

[1]Nanjing NARI Information & Communication Technology Co., Ltd., Nanjing 210003, China
[2]State Grid Wuhan Power Supply Company, Wuhan 430077, China

Abstract

The reliability of power information networks is paramount for smart grid security, yet incomplete data transmission paths due to link failures or cyberattacks critically impair topology awareness and operational decision-making. To address the limitations of existing methods in handling the dynamic and structurally complex nature of these networks, this paper proposes a Topology-Enhanced Dynamic Perception Graph Neural Network (TEDP-GNN) for accurate path completion. The core methodology revolves around three integrated innovations. First, we introduce a topological deep learning framework that models high-order node interactions and network connectivity patterns beyond pairwise relationships. This is achieved by employing simplicial complexes to capture multi-node dependencies inherent in power data routing structures. Second, a novel dynamic perception mechanism is designed to continuously monitor link state changes. This mechanism utilizes a temporal gating unit that ingests real-time network alert data, enabling the model to adaptively reweight message-passing pathways in response to topological disturbances such as failures or attacks. Finally, the task is formulated as a topology-constrained optimization problem. Within this framework, we deploy a hybrid attention module that simultaneously computes both node-level and path-level attentions, alongside a multi-dimensional feature aggregation strategy that synthesizes information from topological embeddings, dynamic states, and historical transmission patterns. Preliminary validation on simulated power communication topologies demonstrates that TEDP-GNN significantly outperforms baseline models in predicting missing links, showing marked improvements in precision and recall. The model provides a robust, topology-aware solution for maintaining data path integrity, thereby enhancing the situational awareness and resilience of power information infrastructures.

Keywords: power information network, path completion, Graph Attention Network, persistent homology, topology enhancement

Received on 01 December 2025, accepted on 10 February 2026, published on 11 August 2026

doi: 10.4108/ ew.13904

1. Introduction

With the rapid advancement of smart grids, power information networks have become critical infrastructure supporting power system production, dispatching, operation, and management [1]. Leveraging high-capacity communication technologies such as 5G, these networks carry massive volumes of heterogeneous power-service data whose propagation behavior is strictly constrained by the underlying topology. However, due to increasing cyberattacks and communication failures, incomplete data-transmission paths have emerged as a major threat to the reliability and security of power information systems [2–3]. Missing links distort the perception of the actual network topology and impose multi-layer operational challenges. On one hand, core applications—including state estimation, fault diagnosis, and optimal dispatching—experience degraded performance when reliable data support is absent [4]. On the other hand, missing links fundamentally deteriorate the observability and controllability of modern power systems. Real-time operation relies on uninterrupted data exchanges among control centers, substations, and field devices; topology incompleteness disrupts these data streams, creating

*Corresponding author. Email: zhxj198708@126.com

monitoring blind spots and degrading situational awareness. As observability declines, the accuracy of dispatching, protection coordination, and emergency response is significantly reduced. The weakened controllability further increases the system's vulnerability to disturbances, elevating the probability of misoperations and cascading failures [5–6]. Therefore, developing robust link-prediction methods for completing data-transmission paths is crucial for maintaining secure and stable grid operations.

Traditional statistical and classical machine learning methods [7–9] generally overlook the structural constraints imposed by network topology on data propagation, resulting in predictions that poorly align with the actual network structure. These limitations hinder their applicability in large-scale, topologically complex power systems. By contrast, Graph Neural Networks (GNNs) have demonstrated strong relational-modeling capabilities in structured domains such as communication and power systems [10–11]. Among them, the Graph Attention Network (GAT) enhances node representations by assigning adaptive attention weights to neighbors and has shown promising results in distribution-network topology identification and communication-fault diagnosis [12–13]. Despite these strengths, existing GNN-based approaches face challenges when dealing with the structural complexity of power information networks. Most methods rely primarily on first-order aggregation, whereas data propagation exhibits multi-hop dependencies with significantly varying feature distributions at different topological distances (L-hop). Without an adaptive mechanism to extract high-order relational patterns across multiple L-hop neighborhoods, models fail to adequately capture real data-flow structures, thereby limiting the performance of link-completion tasks. Moreover, high-order interaction patterns beyond simple pairwise relations are common in complex networks, making topological deep learning a promising direction for improving high-order structural modeling [14–15]. Hence, models must flexibly perceive multi-hop structural differences while integrating global properties such as connectivity, branching structures, and critical communication paths.

To address these challenges, this paper proposes a Topology-Enhanced Dynamic Perception Graph Neural Network (TEDP-GNN) for completing data-transmission paths in power information networks. The core idea of the proposed approach lies in learning expressive node representations under global and multi-scale topological constraints, and subsequently predicting the existence of edges between arbitrary node pairs to reconstruct missing paths. To achieve this, a topological deep learning framework is first introduced to capture high-order structural dependencies, enabling node embeddings to more comprehensively reflect the network's overall connectivity patterns. Next, a dynamic perception mechanism is designed to adaptively extract node-feature variations across different topological distances (L-hop), thereby equipping the model with multi-scale structural awareness. Finally, link prediction is formulated as a topology-constrained optimization problem in the learned representation space, where the existence of edges is inferred through similarity metrics or learnable scoring functions. By jointly leveraging high-order topological features, multi-hop structural information, and relational dependencies, the proposed model achieves more robust and accurate reconstruction of critical data-transmission paths in power information networks.

2. Related Work

2.1. Graph Neural Networks

Graph Neural Networks (GNNs) have become a cornerstone in processing graph-structured data, evolving from early models to advanced architectures that handle complex relational information. Comprehensive surveys outline the foundational developments, categorizing GNNs into recurrent, convolutional, and spatio-temporal types, with key contributions like Graph Convolutional Networks (GCN) for spectral filtering and Graph Attention Networks (GAT) for attention-based aggregation [16]. These models leverage message passing to propagate node features across neighborhoods, enabling tasks such as node classification and graph representation learning [17].

In the domain of link prediction, recent advancements focus on enhancing GNNs with hybrid mechanisms and addressing heterophily. Yan et al. [18] propose a hybrid graph neural network framework for link prediction that integrates Graph Convolutional Networks with dual similarity metrics (cosine and dot product) to improve prediction accuracy, demonstrating effectiveness on social recommendation systems. Zhu et al. [19] investigate the impact of feature heterophily on link prediction with GNNs, analyzing how different encoders and decoders adapt to varying homophily levels and introducing designs for improved performance, with empirical validation on synthetic and real-world datasets.

To address limitations of traditional GNNs relying primarily on first-order neighborhoods, higher-order GNN variants incorporate multi-hop dependencies. Besta et al. [20] demystify higher-order graph neural networks by providing a taxonomy and analyzing over 100 related schemes, focusing on expressiveness, time complexities, and applications using higher-order graph data models like hypergraphs and simplicial complexes. Bailie et al. [21] introduce a higher-order graph attention probabilistic walk network that assigns weights to variable-length paths based on feature diversity, reconstructing k-hop neighborhoods with empirical improvements in graph tasks.

2.2. Persistent Homology

Persistent homology has become a central technique in topological data analysis (TDA), offering a principled way to capture multi-scale topological structures arising in complex data. Building on algebraic-topology foundations and progressively evolving toward scalable implementations, it characterizes connected components, loops, and higher-dimensional cavities across filtration scales while maintaining robustness to small perturbations. Advances in computational methods introduce efficient encoding schemes and reduction strategies that support large-scale applications in data science and engineering [22]. In parallel, persistent-homology-based machine learning frameworks incorporate key computational tools such as persistence diagrams, barcodes, and dedicated software libraries, together with feature transformation techniques that enable the integration of topological descriptors into supervised and unsupervised learning tasks [23].

Recent integrations of persistent homology with machine learning and graph neural networks enhance representation power by injecting topological invariants that complement message-passing limitations in standard GNNs. Xin et al. [24] propose TopInG, a topologically interpretable graph learning framework using persistent homology to identify persistent rationale subgraphs via rationale filtration learning and topological discrepancy constraints, improving interpretability and performance in graph tasks. Ying et al. [25] develop a method for boosting graph pooling with persistent homology, integrating PH into GNNs to enrich expressive power and achieve superior results on graph classification benchmarks. Buffelli et al. [26] introduce CliquePH, which leverages persistent homology on clique graphs to incorporate higher-order information into GNNs, demonstrating enhanced ability to capture complex structural patterns and outperforming baselines on various graph learning tasks by enriching representations with clique-based topological features.

3. System Model

The power information network data transmission path completion method proposed in this paper enhances the accuracy of edge existence inference by jointly modeling dynamic node neighborhood features and global network topological characteristics, thereby enabling the reconstruction of a complete network topology. The overall procedure consists of three main stages:

(1) Construction of the Dynamic Perception Graph Neural Network: Using GATv2 as the backbone architecture, a unified attention computation paradigm is applied to the L-hop neighborhood of each node. Through dynamic adjustment of inter-layer attention weights, the model captures the influence of neighborhood feature distributions at different topological distances on node representations, thereby extracting dynamic local features of nodes.

(2) Design of the Topology-Enhanced Module: Based on persistent homology (PH) theory, multiple multilayer perceptrons (MLPs) are designed as filtration functions to screen topological information within the L-hop neighborhoods of nodes. By computing key descriptors such as persistence barcodes and topological invariants, feature vectors that characterize the global topological structure of the power information network are extracted.

(3) Feature Fusion and Edge Classification: A linear layer is employed to align feature dimensions and perform weighted fusion of the dynamic local features and topology-enhanced features, yielding node representations that jointly encode dynamic behavior and topological constraints. Subsequently, node pairs are fed into an MLP classifier to infer edge existence, ultimately completing the power information network.

The proposed model comprises four main modules, and its overall architecture is illustrated in Figure 1. The first is the neighborhood graph sampling module, which constructs L-layer neighborhood graphs for each node under all relational settings. The second is the neighborhood feature extraction module, which utilizes a hierarchical dynamic multi-head attention mechanism to extract features from the neighborhood graphs. The third is the topological feature extraction module, where multiple filtration functions are constructed via MLPs to extract topological features from the neighborhood graphs. Finally, the feature fusion and edge prediction module integrates neighborhood and topological features through a linear layer and employs an MLP classifier to predict the existence of edges between each pair of nodes.

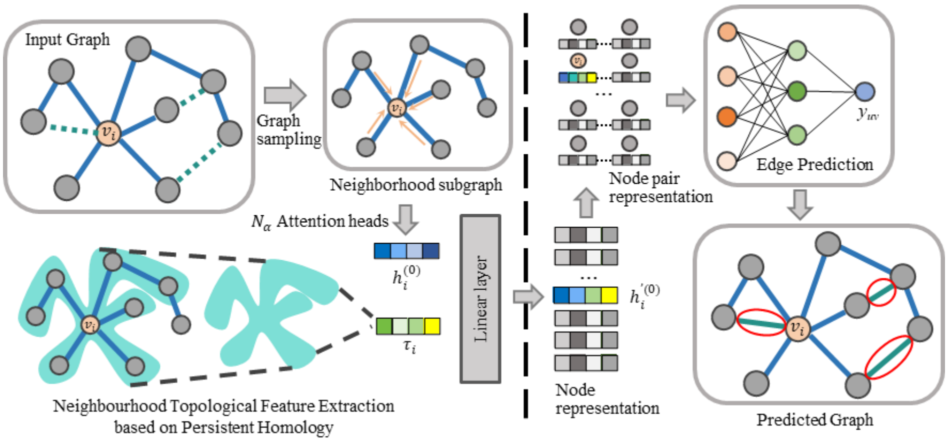


Figure 1. TEDP-GNN Model Architecture

3.1. Neighbourhood Graph Sampling

To address the requirement of capturing multi-scale topological dependencies in the power information network completion task, we consider an input power information network $G = (V, E, X)$, where $V = \{v_1, v_2, \dots, v_N\}$ denotes the set of nodes comprising generators, transformers, load nodes, and other power system components (N is the total number of nodes), $E \subseteq V \times V$represents the set of edges formed by transmission lines, distribution lines, and other physical or logical connections, and $X \in \mathbb{R}^{N\times D}$ is the node feature matrix with D denoting the feature dimension.

In power information networks, the topological scope of an L-hop neighborhood is well aligned with the spatial scale of data transmission. When $L = 1$, the neighborhood subgraph contains only directly connected nodes, corresponding to direct data interaction links. When $L = 2$, the neighborhood expands to indirectly connected nodes, corresponding to indirect data regulation links. For $L \geq 3$, the neighborhood further covers global cooperative links. To systematically capture multi-scale topological dependencies ranging from local direct connections to global indirect associations, and to extract discriminative information embedded in neighborhood feature distributions at different hierarchical levels, an L-hop neighborhood graph $G_i^L = (V_i^L, E_i^L, X_i^L)$ is sampled for each target node $v_i \in V$.

3.2. Neighbourhood Feature Extraction

To address the structural and semantic heterogeneity embedded in different topological levels within a node's L-hop neighborhood, and to enable adaptive modeling of multi-scale features in power information networks, this paper proposes a hierarchical dynamic multi-head attention mechanism for extracting layered feature representations from node neighborhood graphs. This mechanism introduces structurally consistent dynamic multi-head attention units at different topological levels, where neighbor features are aggregated independently for each level, and attention coefficients are computed separately along the hierarchical dimension. Specifically, all levels follow a unified attention computation paradigm; however, independently parameterized attention weights are employed to learn level-specific node interaction patterns corresponding to different topological distances, thereby characterizing the differentiated influence of neighborhoods at various levels on the central node. By assigning exclusive dynamic attention distributions to each level, the proposed mechanism effectively captures multi-level structural characteristics and relational patterns across multi-hop neighborhoods in power information networks.

The neighborhood feature extraction process is illustrated in Figure 2 and adopts a bottom-up hierarchical aggregation strategy. First, nodes at the penultimate level (level 1) aggregate feature representations h_j^2 from their directly connected neighbors (level 2 nodes) through the dynamic multi-head attention mechanism, yielding node features h_j^1 at level 1. Subsequently, the level-1 features are further aggregated to obtain the final neighborhood representation of the level-0 central node v_i. Specifically, during the processing of an L-layer neighborhood, for each node $v_{i'}$ at level $l(l = L-1, \dots, 0)$, N_α attention heads are employed to dynamically learn neighborhood features $h_{i'}^l$from its neighboring nodes v_j. The computation of each attention head is given in Eq. (1):

$$h_{i'}^{(l)} = \sigma\left(\sum_{v_j \in \varkappa_{i'}} \alpha_{i'j}^{(l)} P^{(l)} h_j^{(l)}\right) \tag{1}$$

where $P_l \in \mathbb{R}^{(d'/N_\alpha)\times d'}$, h_j^{l+1} denotes the feature representation of node v_j, and $h_j^L \in \mathcal{X}$. Here, $\kappa_{i'j} = (v_j, v_{i'}) \in E_i^L$, and the attention coefficient $\alpha_{i'jl}$ is computed as shown in Eq. (2):

$$\alpha_{i'j}^{(l)} = \frac{exp\left(a^{(l)}\sigma\left(W^{(l)}\left[h_{i'}^{(l)}||h_j^{(l)}\right]\right)\right)}{\sum_{v_{j'}\in\varkappa_{i'}} exp\left(a^{(l)}\sigma\left(W^{(l)}\left[h_{i'}^{(l)}||h_{j'}^{(l)}\right]\right)\right)} \tag{2}$$

where $\|$ denotes vector concatenation, $W_l \in \mathbb{R}^{d'\times 2d'}$ and $a_l \in \mathbb{R}^{d'}$ are the learnable linear transformation matrix and attention weight vector in the level-l dynamic multi-head attention unit, respectively, $\sigma(\cdot)$ represents the ReLU activation function, and d' denotes the dimensionality of the transformed features.

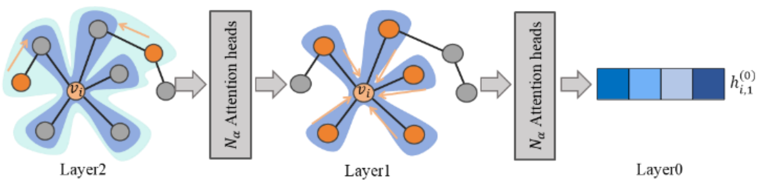


Figure 2. Neighbourhood Feature Extraction

To integrate neighborhood information from multiple perspectives, the outputs of the N_α attention heads are concatenated, yielding the hierarchical neighborhood feature representation $h_{i'}^l \in \mathbb{R}^{d'}$.

By incorporating the proposed hierarchical dynamic multi-head attention mechanism, the model can adaptively adjust neighborhood aggregation strategies across different topological levels, thereby capturing differentiated influence patterns of nodes within multi-hop neighborhoods in power information networks and forming more fine-grained and discriminative neighborhood feature representations. Such hierarchical feature modeling provides stable and structurally informative feature support for subsequent data transmission path completion tasks.

3.3. Topological Feature Extraction

To mine deep structural information beyond local connectivity patterns from power information network data, it is necessary to systematically model the underlying topological characteristics. To this end, this study introduces persistent homology (PH), a fundamental tool in topological data analysis, to capture topological features that persist stably across multi-scale filtration processes. As a core technique in algebraic topology, persistent homology enables precise characterization of the global network topology by quantifying topological invariants such as network holes and connected components. When integrated with machine learning, PH has been successfully applied to tasks including community detection in social networks and traffic flow prediction in transportation networks. However, when transferring persistent homology to the link prediction task in power information networks, two critical challenges remain. The first lies in how to design filtration functions that are well aligned with the data transmission characteristics of power information networks, so as to effectively screen and extract topological information from node L-hop neighborhoods, for example, by distinguishing meaningful topological associations from redundant connections. The second challenge concerns how to construct an effective fusion mechanism between topological features and node operational features, in order to generate node representation vectors that simultaneously encode global topological constraints and local information perception, thereby providing robust support for edge existence inference.

To address these challenges, this paper introduces an MLP-based learnable filtration mechanism, in which k independently parameterized filtration functions are constructed to perform multi-view filtering on node neighborhood subgraphs. Each filtration function consists of a multilayer fully connected network that models high-order interaction features between node pairs through layer-wise nonlinear mappings, thereby adaptively generating filtration values for persistent homology computation. Specifically, for a node $v_j \in V_i^L$, each filtration function is formulated as shown in Eq. (3):

$$F_{v_j} = tanh\big(W_2 ReLU\big(W_1 x_j + b_1\big) + b_2\big) \tag{3}$$

where $W_1 \in \mathbb{R}^{h\times d}$ and $W_2 \in \mathbb{R}^{D\times h}$ denote the weight matrices of the hidden and output layers, respectively; $x_j \in X$ is the node feature vector; b_1 and b_2 are bias terms; h represents the hidden layer dimension; D denotes the output dimension of each filtration function; and $F_{v_j} \in \mathbb{R}^D$ is the resulting filtration value.

Each filtration function captures distinct topological attributes of node v_j within the neighborhood subgraph and determines its activation time during the filtration process, thereby characterizing the importance of the node across multi-scale topological structures. By replacing traditional non-differentiable filtration strategies based on sorting or thresholding with parameterized filtration functions, topological feature computation can be seamlessly embedded into an end-to-end training framework. Meanwhile, while preserving the sensitivity of persistent homology to structural variations, the proposed approach significantly reduces computational complexity from $O(n^3)$ to linear order, enabling efficient scalability to large-scale graph data.

Subsequently, the maximum filtration value among the neighbors of node v_j is computed as shown in Eq. (4):

$$F_{v_j}^{max} = MAX_{u\in \varkappa_j}(F_u) \tag{4}$$

If $F_{v_j}^{\max} > F_{v_j}$, node v_i is considered to have been assigned to another connected component. The persistence difference is then computed according to Eq. (5):

$$d_{v_j} = F_{v_j}^{max} - F_{v_j} \tag{5}$$

This difference approximates the birth–death interval in classical persistent homology, where $d_{v_j} \in \mathbb{R}^D$ reflects the stability of the connected component.

To construct more discriminative topological features, four categories of embedding functions ϕ are employed to map d_{v_j} into a tensor of size $M = d'/4$: linear transformation, triangular transformation, Gaussian transformation, and rational hat transformation [27]. The combination of these four embedding functions equips the model with multi-level topological representation capabilities.

(1) Linear transformation, serving as a basic operator, ensures efficient propagation of simple topological features and is computed as shown in Eq. (6):

$$\phi_{line}(F_{v_j}^{max}, d_{v_j}) = W[F_{v_j}^{max}, d_{v_j}]^{\top} + b \in \mathbb{R}^{D\times M} \tag{6}$$

where $W \in \mathbb{R}^{M\times 2}$ and $b \in \mathbb{R}^M$ are the parameters of the linear layer.

(2) Triangular transformation captures salient band structures in persistence diagrams through learnable peak position parameters, making it particularly effective for identifying critical topological structures. Its computation is given in Eq. (7):

$$\phi_{tri}\left(F_{v_j}^{max}, d_{v_j}\right) = ReLU(d_{v_j} - \left|F_{v_j}^{max} - \tau\right|) \in \mathbb{R}^{D\times M} \tag{7}$$

where $\tau \in \mathbb{R}^M$ denotes the learnable peak location parameters.

(3) Gaussian transformation leverages the smoothness of radial basis functions to robustly model multi-scale topological densities, as shown in Eq. (8):

$$\phi_{gauss}\left(F_{v_j}^{max}, d_{v_j}\right) = exp\left(-\frac{(F_{v_j}^{max} - \mu)^2 + (d_{v_j} - \nu)^2}{2\sigma^2}\right) \in \mathbb{R}^{D\times M} \tag{8}$$

where $\mu, \nu \in \mathbb{R}^M$ are center parameters and $\sigma \in \mathbb{R}^+$ denotes the bandwidth.

(4) Rational hat transformation enhances sensitivity to complex topological critical points through rational function compositions, as defined in Eq. (9):

$$\phi_{hat}\left(F_{v_j}^{max}, d_{v_j}\right) = \left(1 + \left\| \left(F_{v_j}^{max}, d_{v_j}\right) - c \right\|_1\right)^{-1} - \left(1 + \left\| \left(F_{v_j}^{max}, d_{v_j}\right) - c \right\|_1 - r \right\|\right)^{-1} \in \mathbb{R}^{D\times M} \tag{9}$$

where $c \in \mathbb{R}^{2\times M}$ represents the center and $r \in \mathbb{R}^{M}$ denotes the radius.

For a node v_j in the neighborhood subgraph, the outputs of the four transformations are concatenated to form $\phi_j \in \mathbb{R}^{d'}$. Finally, a pooling layer is applied to compute the topological representation of the relational subgraph, as shown in Eq. (10):

$$\tau_i = \frac{1}{|v_j|}\sum_{j=1}^{|v_j|}\phi_j \tag{10}$$

Through the above process, the proposed method preserves the overall topological patterns of neighborhood subgraphs while effectively suppressing local noise and maintaining computational efficiency, which is consistent with the theoretical essence of persistent homology in emphasizing global structural stability.

3.4. Feature Fusion and Edge Prediction

In this paper, a linear layer is employed to linearly transform the topological feature τ_i into $\tau_i' = W_k\tau_i + b$, which is then concatenated with the relational features to obtain the node representation, as shown in Eq. (11):

$$h_i'^{(0)} = h_i^{(0)} + (W_k\tau_i + b) \tag{11}$$

After obtaining the representations $h_{i'}^{0}$ for all nodes, a candidate set containing all possible node pairs is constructed as $\mathcal{P} = \{(u, v) \mid \forall u, v \in V, u \neq v\}$. For each candidate node pair (u, v), their feature representations are fused via concatenation, as defined in Eq. (12):

$$\mathbf{z}_{uv} = \left[h_u'^{(0)} \parallel h_v'^{(0)}\right] \tag{12}$$

The fused features are then fed into an MLP classifier for decision making. The output layer adopts a Sigmoid activation function to produce a prediction probability $y_{uv} \in [0,1]$, which indicates the likelihood of a connection existing between the node pair. When $y_{uv} > 0.5$, an edge is inferred to exist between the nodes and the label is set to 1; otherwise, no connection is predicted and the label is set to 0.

4. Experiments and Analysis of Results

4.1. Experimental Setup

All experiments were conducted on a Linux platform using the PyTorch framework, with an NVIDIA A40 GPU (48 GB memory) and 25 vCPUs based on the Intel(R) Xeon(R) Platinum 8481C processor. To evaluate the effectiveness of the proposed method, two publicly available datasets of different scales were employed in the experiments to simulate power information networks of varying sizes, including a small-scale network with 300 communication nodes and 878 links, and a medium-to-large-scale network with 5,715 communication nodes and 19,794 links.

4.1.1. Baseline Model

To evaluate the effectiveness of TEDP-GNN in the network completion task, it is compared with four baseline models in the experiments, as described below:

MLP: A multilayer perceptron composed of two stacked linear layers and nonlinear activation functions.

GAT [28]: A graph attention network that aggregates neighborhood information using an attention mechanism.

CN_NN: The number of common neighbors is a classical metric in link prediction, based on the core assumption that the more common neighbors two nodes share, the more likely a link exists between them [29–30]. In this study, this idea is transferred to the network completion task to construct the baseline method CN_NN.

PA_NN: The preferential attachment (PA) link prediction method based on node similarity [31–32] is adapted to the network completion task to form the baseline algorithm PA_NN. The underlying rationale of this method is that nodes in a network are more likely to have missing connections with nodes that are topologically closer and have higher degrees.

4.1.2. Evaluation Indicators

In this paper, Precision, Recall, and F1-score are adopted as evaluation metrics, and their definitions are given in Eqs. (13)–(15):

$$Precision = \frac{n_1}{n_1 + n_2} \tag{13}$$

$$Recall = \frac{n_1}{n_1 + n_3} \tag{14}$$

$$F1 = \frac{2n_1}{2n_1 + n_2 + n_3} \tag{15}$$

where n_1 denotes the number of missing edges that are correctly predicted, n_2 represents the number of non-missing edges that are incorrectly predicted as missing, and n_3 denotes the number of missing edges that are incorrectly predicted.

4.2. Experiment

4.2.1. Comparative Experiment

First, in the first group of experiments, the performance of TEDP-GNN and the baseline models is compared on two power information network datasets of different scales, and the experimental results are reported in Tables 1 and 2.

Table 1. The performance of a prefectural-level electric power communication network

Model	Precision	Recall	F1
MLP	0.130	0.200	0.158
GAT	0.309	0.257	0.281
CN_NN	0.082	0.116	0.096

PA_NN	0.113	0.132	0.122
TEDP-GNN	0.612	0.306	0.412

Table 2. The performance of a provincial-level electric power communication network

Model	Precision	Recall	F1
MLP	0.110	0.125	0.116
GAT	0.266	0.222	0.242
CN_NN	0.014	0.064	0.023
PA_NN	0.125	0.069	0.089
TEDP-GNN	0.562	0.267	0.362

The experimental results indicate that traditional statistical methods (CN_NN and PA_NN) exhibit clear performance limitations. In small-scale networks with relatively few nodes and edges, CN_NN, which is based on the assumption of common neighbors, performs reasonably well in simple network structures; however, under complex topologies, its performance degrades significantly, achieving a Precision of only 0.096, a Recall of 0.116, and an F1-score of 0.082. These results demonstrate its inability to effectively capture the complex connection patterns present in real-world networks. PA_NN incorporates node degree and distance information and shows a slight performance improvement, but its underlying assumptions are overly simplified and fail to adequately account for the dynamic characteristics of data transmission networks, resulting in limited generalization capability.

In contrast, deep learning–based methods (MLP and GAT) demonstrate stronger adaptability. By extracting features through multiple nonlinear transformations, MLP achieves an F1-score of 0.158; however, it does not effectively exploit network topological information. GAT further improves performance by aggregating neighborhood information via an attention mechanism, increasing the F1-score to 0.281. Nevertheless, its attention mechanism struggles to adaptively capture fine-grained features across different topological levels in dynamically evolving network scenarios, leading to evident performance bottlenecks.

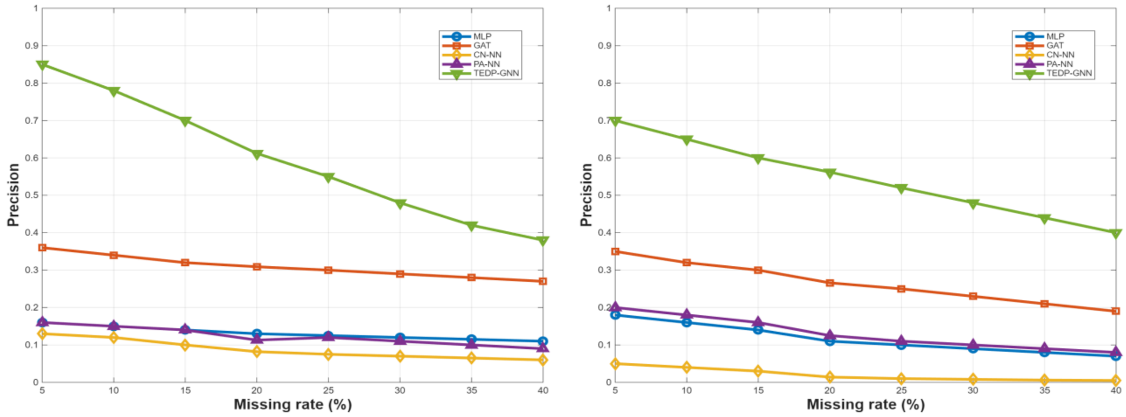
By integrating the dynamic attention mechanism of GATv2 with a persistent homology–based topological feature extraction module, TEDP-GNN is able to efficiently model complex network topologies. On the small-scale dataset, TEDP-GNN achieves an F1-score of 0.400, with a Precision of 0.612 and a Recall of 0.297, significantly outperforming both traditional statistical methods and deep learning–based approaches. The dynamic perception capability enables the model to capture fine-grained features at different hierarchical levels, while the topology-enhanced module effectively extracts high-order topological patterns, such as cyclic structures and critical paths, thereby compensating for the deficiencies of conventional methods. These results provide a reliable solution for link prediction in complex networks and highlight the importance of dynamic perception and topological enhancement in such tasks.

Furthermore, the same experiments are conducted on a larger-scale network dataset. The results show that as the network scale increases, the performance of all methods declines to varying degrees. The F1-scores of the traditional statistical methods (CN_NN and PA_NN) drop to 24% and 72% of their respective values on the small-scale dataset, indicating high sensitivity to data scale. The performance of the deep learning methods (MLP and GAT) also decreases, with F1-scores reduced to 73% and 86%, respectively, which can be attributed to overfitting caused by increased model complexity. In contrast, TEDP-GNN maintains relatively strong performance on the large-scale dataset, with its F1-score decreasing only to 91% of that obtained on the small-scale dataset. These results clearly demonstrate the superior scalability and robustness of the proposed TEDP-GNN model on large-scale graphs.

4.2.2. Reliability Analysis Experiment

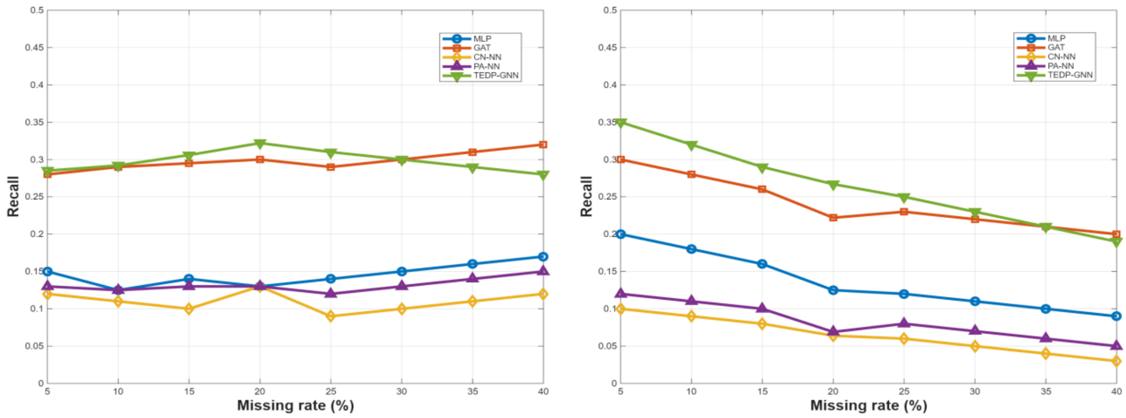
The proportion of missing edges in a network is one of the key factors affecting the reliability of the TEDP-GNN model. To verify the robustness of the proposed model, this section systematically adjusts the number of missing edges in the dataset, thereby controlling the missing-edge ratio in the network, and evaluates the performance of TEDP-GNN on the network completion task under different missing-edge settings. This experimental design aims to comprehensively assess the model's performance across varying levels of edge sparsity and to validate its reliability. On the small-scale network, the experimental results of TEDP-GNN under different missing-edge ratios are illustrated in Figures 3, 4 and 5. As shown in the figures, with an increasing proportion of missing edges, Precision decreases significantly from 0.85 to 0.38. This degradation is mainly attributed to the loss of local structural information caused by missing edges, which reduces the amount of effective features that the model can learn and consequently lowers prediction confidence. Notably, however, the F1-score only declines gradually from 0.422 to 0.348, demonstrating strong stability.

This differentiated performance trend can be primarily attributed to two key design components of the proposed model. First, the neighborhood feature extraction module adaptively adjusts the attention weights of neighborhood features at different layers, enabling the model to intelligently reallocate feature importance when information is missing, thereby effectively mitigating the impact of incomplete local structures. Second, the high-order topological features extracted by the topological feature module exhibit stronger robustness and provide stable structural prior knowledge for the model, further enhancing its resilience under high missing-edge ratios.



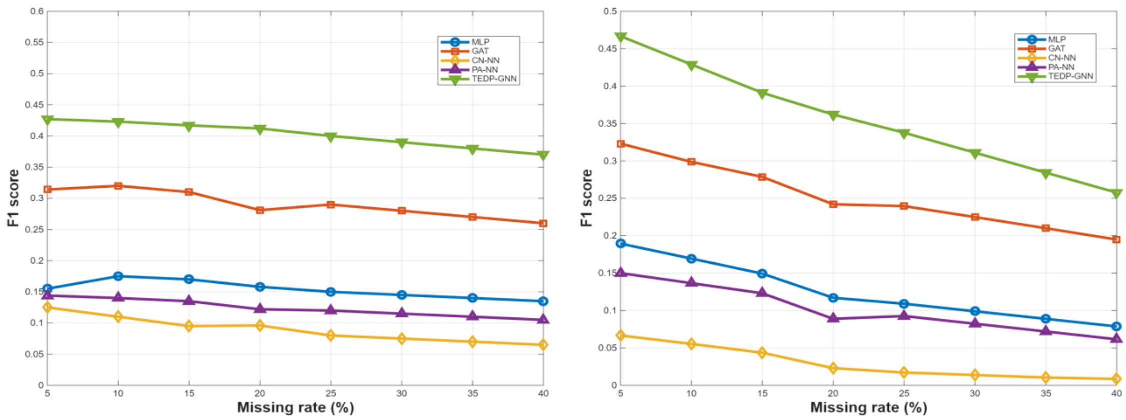
(a) (b)

Figure 3. The Impact of Edge Loss Rate on Precision



(a) (b)

Figure 4. The Impact of Edge Loss Rate on Recall



(a) (b)

Figure 5. The Impact of Edge Loss Rate on F1 score

4.2.3. Ablation Study

To verify the effectiveness of topological features and neighborhood features in node representation learning, this section conducts ablation experiments on a small-scale data flow network by separately removing the neighborhood feature module and the topological feature module. TE-GNN and DP-GNN denote the variants obtained by removing the neighborhood feature module and the topological feature module, respectively, and the corresponding performance variations are summarized in Table 3.

Table 3. The influence of neighbourhood features and node features

Model	Precision	Recall	F1
TEDP-GNN	0.612	0.297	0.400
TE-GNN	0.230	0.205	0.217
DP-GNN	0.312	0.269	0.289

The experimental results demonstrate that both the topological enhancement module and the dynamic perception module make indispensable contributions to the performance improvement of the TEDP-GNN model. When the topological enhancement module is removed, the F1-score decreases significantly from 0.400 to 0.289, with a noticeably larger drop in Precision than in Recall. This phenomenon indicates that this module primarily improves prediction accuracy by enhancing the extraction of high-order topological features. In contrast, removing the dynamic perception module leads to more severe performance degradation, with the F1-score decreasing by 0.183, Precision dropping by 0.382, and Recall decreasing by 0.092. These results clearly validate the critical role of this module as a core component of the model in capturing the dynamic characteristics of the network.

Further analysis reveals that the two modules exhibit strong functional complementarity. The topological enhancement module extracts topological invariants—such as those derived from persistent homology—that effectively identify key structural patterns in the network, thereby significantly improving prediction precision. Meanwhile, the dynamic perception module leverages the attention mechanism of GATv2 to dynamically capture variations in inter-node association strengths, simultaneously optimizing both Precision and Recall. This multi-level collaborative mechanism not only explains the superior performance of TEDP-GNN but also provides valuable insights for the design of graph neural networks targeting dynamic and complex networks. In particular, for application scenarios that require simultaneous consideration of topological structure and dynamic evolution, this dual-module collaborative design offers an important and practical reference.

5. Conclusions

This paper addresses the problem of incomplete topological awareness caused by missing link association data in power information networks and proposes a graph neural network–based network completion method that integrates dynamic perception and topological enhancement. The proposed approach leverages GATv2 to capture feature correlations among nodes across multiple topological distances, while global topological information is extracted using persistent homology. Link prediction is then performed through feature fusion and MLP-based classification. Experimental results on two network datasets of different scales demonstrate that the proposed method achieves significantly higher completion accuracy and robustness than traditional approaches, effectively overcoming the limitations of existing techniques in adapting to the topology of power data flow networks.

Nevertheless, several limitations remain. The proposed model does not fully account for the heterogeneous topological characteristics of multi-energy coupled networks, and its computational efficiency in ultra-large-scale power grid scenarios still requires improvement. Future work will focus on enhancing the adaptability of the topological enhancement module, introducing lightweight network architectures to improve inference efficiency, and exploring privacy-preserving completion schemes under federated learning frameworks, in order to better meet the complex and security-critical operational requirements of next-generation power systems.

Author Contribution

Conceptualization, X.L., Z.W., and W.G.; methodology, X.L., Z.W., and W.G.; software, L.G. and S.C.; validation, L.G., S.C., and Y.C.; formal analysis, X.L., W.G., and S.C.; investigation, X.L. and Z.W.; resources, L.G.; data curation, S.C.; writing—original draft preparation, X.L., Z.W. and W.G.; writing—review and editing, L.G., S.C., and Y.C.; visualization, S.C.; supervision, W.G. and L.G.; project administration, W.G.; funding acquisition, S.C. All authors have read and agreed to the published version of this manuscript.

Acknowledgements

This research was funded by the Science and Technology Project of State Grid Corporation of China (52150025001H-146-ZN).

Data Availability Statement

Data will be made available on request.

Ethical Approval

Not applicable

Competing Interest

The authors have no relevant financial or non-financial interests to disclose.

References

[1] Krause T, Ernst R, Klaer B, Hacker I, Henze M. Cybersecurity in power grids: Challenges and opportunities. Sensors. 2021; 21(18):6225.
[2] Rajkumar V, Subramaniam V, Ştefanov A, Presekal A, Palensky P, Torres JLR. Cyber attacks on power grids: Causes and propagation of cascading failures. IEEE Access. 2023; 11:103154–103176.
[3] Islam MZ, Lin Y, Vokkarane VM, Venkataramanan V. Cyber-physical cascading failure and resilience of power grid: A comprehensive review. Front. Energy Res. 2023; 11:1095303.
[4] Deng R, Xiao G, Lu R, Liang H, Vasilakos AV. False data injection on state estimation in power systems—Attacks, impacts, and defense: A survey. IEEE Trans. Ind. Inf. 2016; 13(2):411–423.
[5] Li Y, Ge Y, Xu T, Zhu M, He Z. Controllability evaluation of complex networks in cyber–physical power systems via critical nodes and edges. Int. J. Electr. Power Energy Syst. 2024; 155:109625.
[6] Jia H, Li J, He B, Li S, Cheng Z, Zhao C. Enhancing the robustness of cyber-physical power systems against cross-domain cascading failures: Cyber-physical dynamic reconfiguration. Reliab. Eng. Syst. Saf. 2026; 112329.
[7] Wu H, Song C, Ge Y, Ge T. Link prediction on complex networks: An experimental survey. Data Sci. Eng. 2022; 7(3):253–278.
[8] Jin ZX, Yi JF, Shang KK. Learning-based link prediction methods integrating network topological features and embedding representations. arXiv preprint arXiv:2512.06677. 2025.
[9] Zare G, Navimipour NJ, Hosseinzadeh M, Sahafi A. Network link prediction via deep learning method: A comparative analysis with traditional methods. Eng. Sci. Technol. Int. J. 2024; 56:101782.
[10] Liao W, Bak-Jensen B, Pillai JR, Wang Y, Wang Y. A review of graph neural networks and their applications in power systems. J. Mod. Power Syst. Clean Energy. 2021; 10(2):345–360.
[11] Dolatyabi P, Khodayar M. Graph neural networks and their applications in power systems: A review. In 2025 IEEE Int. Conf. Electro Information Technology (eIT), 2025: 1–10.
[12] Zhao C, Li X, Cai Y. A power grid topology detection method based on edge graph attention neural network. Electr. Power Syst. Res. 2025; 239:111219.
[13] Karabulut B, Manna C, Develder C. Generalization of graph neural network models for distribution grid fault detection. In 2025 IEEE Int. Conf. Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), 2025: 1–7.
[14] Millán AP, Sun H, Giambagli L, et al. Topology shapes dynamics of higher-order networks. Nat. Phys. 2025; 21: 353–361.
[15] Papamarkou T, Birdal T, Bronstein M, Carlsson G, Curry J, Gao Y, Hajij M, et al. Position: Topological deep learning is the new frontier for relational learning. Proc. Mach. Learn. Res. 2024; 235:39529.
[16] Wu Z, Pan S, Chen F, Long G, Zhang C, Yu PS. A comprehensive survey on graph neural networks. IEEE Trans. Neural Netw. Learn. Syst. 2020; 32(1):4–24.
[17] Zhou J, Cui G, Hu S, Zhang Z, Yang C, Liu Z, Wang L, Li C, Sun M. Graph neural networks: A review of methods and applications. AI Open. 2020; 1:57–81.
[18] Gharsallah S, Yahia S, Bouchelligua W, Bouchrika T. Advancing link prediction with a hybrid graph neural network approach. Mathematics. 2025; 13(22):3594.
[19] Zhu J, Li G, Yang YA, Zhu J, Cui X, Koutra D. On the impact of feature heterophily on link prediction with graph neural networks. Adv. Neural Inf. Process. Syst. 2024; 37:65823–65851.
[20] Besta M, Scheidl F, Gianinazzi L, Kwasniewski G, Klaiman S, Müller J, Hoefler T. Demystifying higher-order graph neural networks. IEEE Trans. Pattern Anal. Mach. Intell. 2026; 48(3):2544-2565.
[21] Bailie T, Koh YS, Mukkavilli K. Higher order graph attention probabilistic walk networks. arXiv preprint arXiv:2411.12052. 2024.
[22] Malott NO, Chen S, Wilsey PA. A survey on the high-performance computation of persistent homology. IEEE Trans. Knowl. Data Eng. 2022; 35(5):4466–4484.
[23] Pun CS, Xia K, Lee SX. Persistent-homology-based machine learning and its applications—A survey. arXiv preprint arXiv:1811.00252. 2018.
[24] Xin C, Xu F, Ding X, Gao J, Ding J. Toping: Topologically interpretable graph learning via persistent rationale filtration. arXiv preprint arXiv:2510.05102. 2025.
[25] Ying C, Zhao X, Yu T. Boosting graph pooling with persistent homology. Adv. Neural Inf. Process. Syst. 2024; 37:19087–19113.
[26] Buffelli D, Soleymani F, Rieck B. CliquePH: Higher-order information for graph neural networks through persistent homology on clique graphs. arXiv preprint arXiv:2409.08217. 2024.
[27] Hofer CD, Kwitt R, Niethammer M. Learning representations of persistence barcodes. J. Mach. Learn. Res. 2019; 20(126):1–45.
[28] Veličković P, Cucurull G, Casanova A, et al. Graph attention networks. arXiv preprint arXiv:1710.10903. 2017.
[29] Lü L, Zhou T. Link prediction in complex networks: A survey. Physica A. 2011; 390(6):1150–1170.
[30] Wang M, Lou X, et al. A degree-related and link clustering coefficient approach for link prediction in complex networks. Eur. Phys. J. B. 2021; 94(1):33.
[31] Scellato S, Noulas A, et al. Exploiting place features in link prediction on location-based social networks. In Proc. 17th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining. 2011:1046–1054.
[32] Duan Y, Guan Q. Predicting potential knowledge convergence of solar energy: Bibliometric analysis based on link prediction model. Scientometrics. 2021; 126(5):3749–3773.