

Optimal Matching of Ring and Radial Wiring Modes with Three-Level Feeder Automation Configuration in 10 kV Urban Distribution Networks Based on Graph Attention Networks

Fei Zheng^{1,*}, Jinseng Liu¹, Ning Luo¹, Qingyu Zhao¹ and Mao Miao¹

¹Power Grid Planning & Research Center Guizhou Power Grid Co., Ltd., Guiyang 550002, China

Abstract

The matching between wiring modes and feeder automation (FA) levels in 10 kV urban distribution networks lacks systematic methods, particularly neglecting communication infrastructure constraints. This paper proposes a two-stage matching model based on graph attention networks (GAT) that fuses electrical, topological, and information and communication technology (ICT) features. A 360-scenario dataset is constructed from IEEE 33/69-node test systems, covering four wiring modes (single-radial, multi-section radial, two-supply ring, three-supply ring) and three FA levels (FA-I/II/III). The first stage employs GAT for wiring mode identification, achieving 91.7% accuracy. The second stage recommends optimal FA configurations through feature fusion with embedded ICT feasibility hard constraints, attaining 83.3% Top-1 matching accuracy. Transfer validation on 72 operational feeders from three Nanjing districts yields 87.5% wiring mode accuracy and 79.2% Top-1 FA matching. Communication features contribute a 3.1 percentage-point gain to matching accuracy, and the constraint mechanism effectively prevents infeasible FA recommendations in low-coverage scenarios. Ablation studies confirm the superiority of the two-stage cascaded architecture over end-to-end alternatives.

Keywords: Graph Attention Network, Distribution Network Wiring Mode, Feeder Automation Configuration, ICT Feasibility Constraint

Received on 20 February 2026, accepted on 17 May 2026, published on 12 October 2026

Copyright © 2026 Fei Zheng *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/ew.14640

*Corresponding author. Email: zhengfyawy@163.com

1. Introduction

The urban medium-voltage distribution network at the 10 kV voltage level plays an important role as the backbone of power supply to end users in Chinese cities. With the power system evolution towards a new type of architecture that incorporates distributed generation and flexible loads, the wiring modes of 10 kV distribution networks, such as single radial, multi-section radial, and ring-type configurations, have become a critical issue to balance the reliability of power supply and capital investment. A thorough review of the topology planning of distribution networks revealed that the

final decision between radial and loop configurations plays a critical role in determining fault isolation, voltage regulation, and load transfer capabilities of the network [1]. Moreover, network reconfiguration techniques that utilize sectionalizing and tie switches have further complicated the selection of wiring modes [2].

Equally important to supply reliability is the configuration of feeder automation (FA), which determines the speed at which faults can be detected and isolated. Recent studies on the optimal configuration of feeder terminal units have quantitatively related FA deployment density with supply reliability indices such as SAIDI and SAIFI [3]. The

theoretical foundation of the three-level FA configuration, based on local-logic FA-I (recloser-based with no communication channel), centralized FA-II (using a communication channel between the master station and distribution automation devices via fiber optic cable and wireless private networks), and distributed FA-III (using peer-to-peer GOOSE protocol between intelligent devices), has been codified in Chinese distribution automation technical standards to define the grading basis. The recent investigation on FA configuration optimization in distribution network resilience reinforcement indicated that the grade of FA should be correlated to the topology to ensure that supply investment in FA is neither excessive nor insufficient[4]. While the interdependence between topology and FA configuration is well recognized in theory, existing practice in FA configuration relies on rules based on experience rather than quantification of how topology and communication infrastructure features interact in determining the optimal match.

The development of graph neural networks (GNNs) has been recognized as a significant paradigm for learning from graph-structured data, and their application to power systems has seen rapid growth. A review of GNN applications in power systems summarized the achievements of GNNs in power system fault diagnosis, power flow calculation, and topology identification [5]. Among the GNNs, the graph attention network (GAT) proposed by Velickovic et al. is highlighted for its capability to assign learnable attention weights to neighboring nodes to differentiate the importance of topologically heterogeneous nodes without using predefined aggregation rules [6]. In the power system field, GAT has been applied to differentiate transient rotor angle instability from short-term voltage instability; meanwhile, the framework of graph reinforcement learning has been integrated with attention mechanisms to implement real-time scheduling for active distribution networks [7,8].

Spatial-temporal recurrent GNNs have been developed for fault diagnostics in distribution systems by simultaneously considering the graph structure of the network and the temporal characteristics of the electrical signals[9]. Recently, deep graph neural networks have been specifically designed for fault detection/identification in distribution grid systems, showing highly accurate results for fault section identification [10]. The possibility of using GNNs for power flow analysis in distribution systems has been explored, though there are some issues to be addressed regarding the generalization of results over different topological structures of the network [11]. Feature-enhanced GAT models have been proposed for topology identification of low-voltage distribution networks, showing that enriching node features with domain-specific electrical features enhances identification accuracy [12]. GNNs have been used for operational risk assessment of power systems considering changing schedules of unit commitment at the transmission network level [13]. Reinforcement learning over graphs has been used to address real-time outage management of active distribution networks [14]. Recently, an edge-based GAT method was proposed for power grid topology detection, showing the increasing

maturity of GNN models for topology-related problems in power systems [15].

Despite all these advances, there exists an interesting gap that this study attempts to fill by not only proposing the joint problem of distribution network wiring mode identification and FA configuration matching as a graph classification and multi-label recommendation problem, but also by proposing the first study that incorporates the features of the information and communication technology (ICT) infrastructure into the FA configuration matching model. The readiness of the communication infrastructure plays an important role in imposing hard constraints on the feasibility of deploying FAs. F-A-II requires dedicated fiber channels or wireless private network channels, while F-A-III requires low-latency GOOSE communication links.

To address this gap, this paper proposes a two-stage matching model using GAT that combines the fusion of electrical, topological, and ICT characteristics for data-driven optimal matching between wiring modes and FA levels in 10 kV urban distribution networks. The methodological pipeline of the proposed framework is as follows: (1) constructing a multi-scenario dataset of 10 kV distribution networks using IEEE standard test feeders with parameters adjusted to Chinese urban distribution grid specifications; (2) modeling each scenario as an attributed graph with multi-dimensional node and edge attributes that cover both electrical and communication characteristics; (3) developing a two-stage GAT model for wiring mode identification and FA level recommendation with embedded ICT feasibility constraints; and (4) validating the proposed method using comparative studies with conventional methods and alternative GNN architectures.

2. Materials and Methods

2.1 Data Sources and Distribution Network Scenario Construction

The dataset is built upon two widely used IEEE standard distribution test systems, the 33-node and 69-node IEEE radial feeder models, which have been documented in terms of design philosophy and parameter settings by Schneider et al. [16]. To make the IEEE standard distribution test systems suitable for the Chinese 10 kV urban distribution grid, the voltage levels, conductor materials, and load levels are adjusted according to the technical guidelines for 10 kV urban distribution network planning (DL/T 5729—2016) and the distribution automation technical guideline (Q/GDW 382—2010, publicly accessible version). The line impedances are adjusted to fit the typical overhead and cable conductor materials used in Chinese urban MV grids, and the load levels are set in the range of typical values in Chinese urban 10 kV networks (2 to 4 MW per feeder section).

With these two base topologies, four kinds of wiring modes are derived by manipulating tie switch connections and adding or deleting branches. The single radial mode (R1) maintains the original tree structure with all loads supplied by

a single source via a trunk line with two sectionalizing switches and no tie switches. The multi-section single tie radial mode (R2) extends R1 with more sectionalizing switches (a total of four) and a normally open tie line to connect the feeder end to an adjacent feeder via a feeder-feeder tie switch in a hand-in-hand structure with an open-loop configuration. The two supply one backup ring mode

(L1) involves connecting two different feeder sources to a ring main unit with a normally open tie switch to enable load transfer in case of a fault. The three supply one backup ring mode (L2) comprises three source points with a dual-ring structure and two normally open tie points for maximum flexibility in load transfer. Figure 1 presents the topological characteristics of these four kinds of wiring modes.

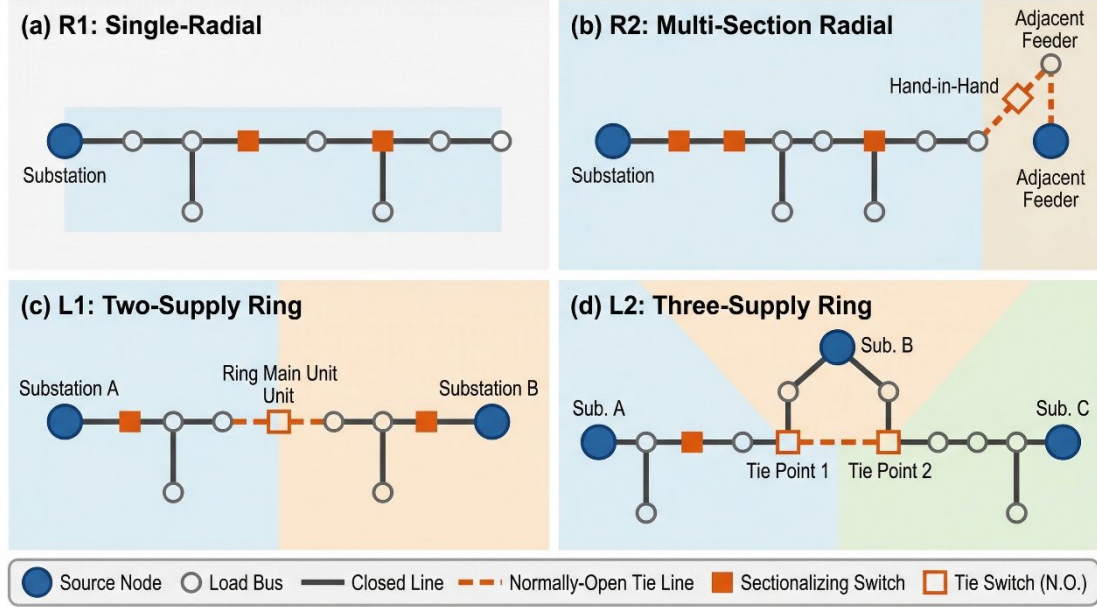


Figure 1. Topological schematics of four 10 kV distribution network wiring modes: (a) single-radial R1; (b) multi-section single-tie radial R2; (c) two-supply one-backup ring L1; (d) three-supply one-backup ring L2

Three FA levels are configured for each of the wiring mode categories: FA-I (local recloser-based decision-making without communication), FA-II (centralized master station decision-making with fiber or 4G wireless private network channels in a star topology), and FA-III (peer-to-peer distribution decision-making with GOOSE fiber links in a ring topology between terminals). This gives 4 categories * 3 FA levels = 12 mode-level combinations. The parameters of the communication infrastructure are synchronized with the three FA levels: for FA-I scenarios, there are no communication links; for FA-II scenarios, there are fiber/wireless channels connecting the terminals to the master station; for FA-III scenarios, there are peer-to-peer GOOSE fiber links between terminals. For each of these 12 mode-level combinations, 30 scenarios are created at random by varying the load levels (by 15% of the base value) and the positions of the switch configurations.

Power flow analysis and fault simulation are carried out using OpenDSS, an open-source simulation tool. A single line to ground fault simulation and a three-phase short circuit fault simulation are performed for each scenario. Supply restoration time, fault isolation success rate, SAIDI, and SAIFI are recorded as reliability labels.

An external sample of 72 operational feeders is collected from three Nanjing districts (Gulou 26, Jiangning 28, Pukou 18) covering 2022–2024, with wiring mode distribution R1/R2/L1/L2 = 20/26/18/8 and FA level distribution FA-I/II/III = 28/38/6. The feeder length distribution and key topological metrics relative to the synthetic dataset are summarized in Figure 2. This sample is reserved exclusively as an external test set and is not used for training.

2.2 Graph Structure Modeling and Multi-Dimensional Feature Extraction

Each distribution network scenario is represented as an attributed graph $G = (V, E, X)$, where V represents the nodes of the distribution network corresponding to buses, load points, and communication terminals; E represents the edges of the distribution network corresponding to lines, switches, and communication links; and X represents the node feature matrix. There are five node attributes: node degree, active power load, reactive power load, voltage deviation, and communication terminal type, represented as a categorical variable (0: absence of terminal; 1: local FTU; 2: communication-type FTU; 3: distributed DTU). For the edges

of the distribution network, there are five attributes: line impedance, tie-switch indicator (Boolean), type of switch (circuit breaker; load switch; sectionalizer), type of communication link (0: absence of communication; 1: wireless private network; 2: fiber optic), and estimated communication latency grade (0: absence of communication; 1: high latency > 500 ms; 2: medium latency 50–500 ms; 3: low latency < 50 ms).

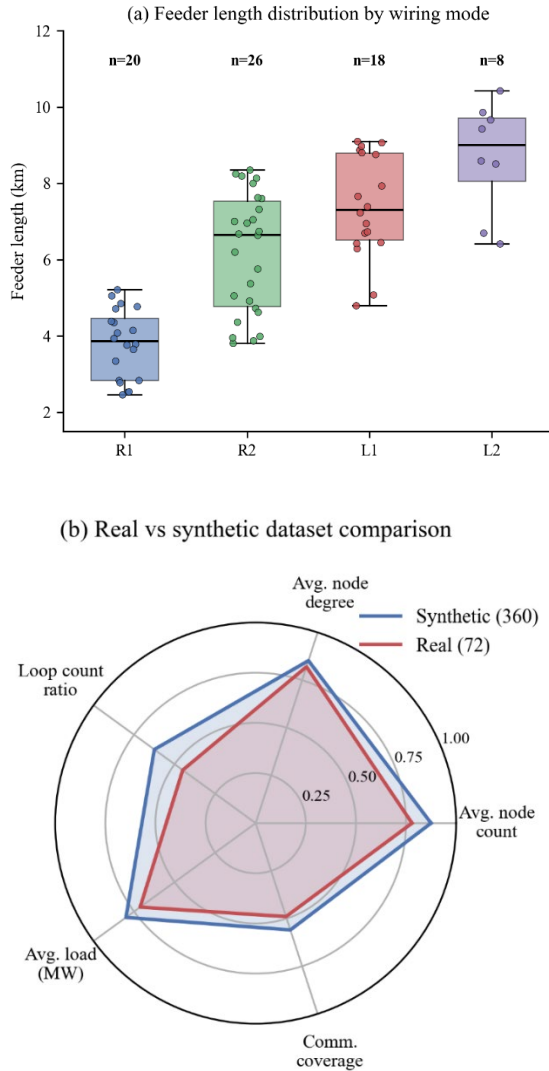


Figure 2. Characteristics of the 72 operational feeders: (a) feeder length distribution by wiring mode; (b) real vs synthetic dataset comparison on key metrics

In addition to these node- and edge-level features, several graph-level global features are computed for each scenario. These features are: total number of nodes, total number of edges, number of loops with ring or radial topology, average degree of nodes, network diameter, tie switch ratio, sectionalizing switch density, communication coverage ratio (defined as the proportion of switch terminals that are equipped with communication capability), and average

communication latency grade. These global features are concatenated with the GAT-learned node-level embeddings after graph pooling to be used as input to the next stage. The wiring mode has a four-class encoding: R1, R2, L1, L2. The FA level has a three-class encoding: FA-I, FA-II, and FA-III. The optimal match label for each scenario is obtained by ranking all possible combinations of mode/level pairs comprehensively according to two criteria: SAIDI performance and reliability gain per unit investment, which is defined as SAIDI improvement divided by FA incremental cost. Here, the FA incremental cost is further divided into primary equipment cost and communication infrastructure cost. The first-ranked combination is defined as the optimal match.

2.3 GAT-Based Matching Model Architecture

The overall model utilizes a two-stage cascaded model structure. This follows recent developments in wavelet-based variants of the GAT model in distribution grid topology detection[17]. The first component in the model is the wiring mode identification module. The input to this module is the attributed distribution network graph. Two GAT layers are employed to aggregate features at the node level. Each employs 4 heads with a hidden dimension size of 64 and utilizes the ELU activation function. The multi-head attention mechanism calculates the coefficients between neighboring nodes. This allows the model to focus on more important nodes in the topology, such as tie points and sectionalizing switch positions, and reduce the importance of less informative nodes, such as the loads. Following the second GAT layer, global mean pooling reduces the dimension to a vector. This vector is then passed through a fully connected layer to output probabilities corresponding to the four classes: R1, R2, L1, and L2.

The second stage is the FA-level recommendation module. The graph-level embedding vector from Stage 1 is concatenated with the one-hot encoded identified wiring mode and global topology/communication feature vector (including network node count, edge count, loop count, average node degree, tie-switch ratio, communication coverage ratio, and average communication latency grade). This concatenated vector is then input into a two-layer fully connected network with hidden units of 128 and 64 to produce scores for all 12 combinations of modes at the level. An ICT feasibility constraint layer is added to the model after the raw score output. If the communication coverage ratio is less than the minimum deployment threshold for FA-II (0.6), then all scores for FA-II and FA-III combinations are set to 0. If the communication latency grade does not meet the low latency requirement for FA-III, then all scores for FA-III combinations are set to 0. The highest score among all combinations is used to determine the recommended optimal FA configuration. The entire model architecture is shown in Figure 3.

The loss function is a weighted joint loss: $L = \alpha \cdot L_{\text{cls}} + (1 - \alpha) \cdot L_{\text{rank}}$, where L_{cls} is the cross-entropy loss for wiring mode classification, and L_{rank} is the margin ranking

loss for the matching recommendation. The weighting factor α is set as 0.4, which means that a little more emphasis is put on the matching task. The Adam optimizer with an initial learning rate of 0.005, a batch size of 32, dropout with a rate of 0.3, and L2 regularization are used during training to

prevent overfitting. Training convergence curves for both stages are presented in Figure 4 and confirm stable optimization without overfitting, with the validation accuracy plateauing within 80 epochs.

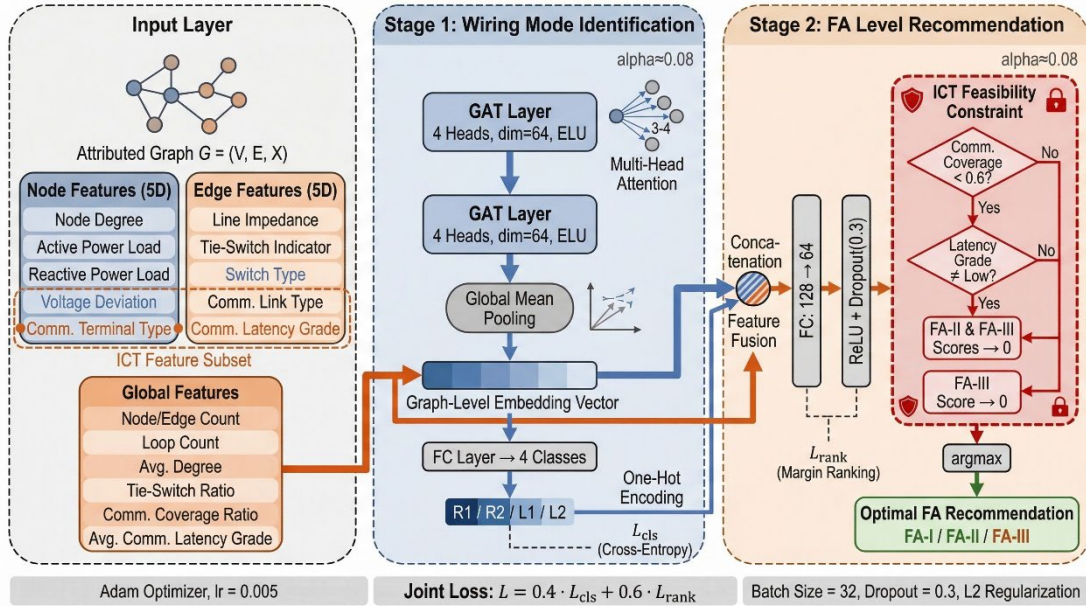


Figure 3. Architecture of the two-stage GAT-based wiring mode–FA level matching model Blue-gray pathways represent electrical–topological processing; orange pathways denote ICT feature flow; the red dashed box marks the ICT feasibility constraint module

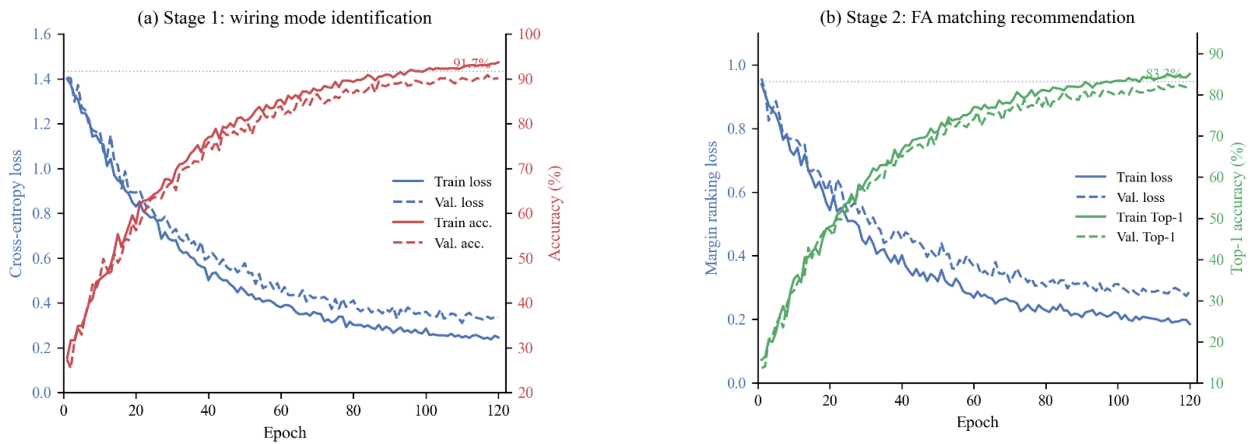


Figure 4. Training convergence of the two-stage GAT model: (a) Stage 1 wiring mode identification; (b) Stage 2 FA matching recommendation

2.4 Experimental Design and Evaluation Metrics

The 360 sample dataset is split into training, validation, and test sets at a 7:1.5:1.5 ratio. This means that 252 samples are used for training, 54 for validation, and 54 for testing. Stratified sampling is used to ensure that all four wiring modes are represented evenly across all sets. Five-fold cross-

validation is used. Average performance metrics are reported. Accuracy, weighted F1-score, and recall are used as evaluation metrics for the wiring mode identification task. Top-1 matching accuracy, Top-2 hit rate, and MRR are used as evaluation metrics for the FA matching recommendation task.

Four baseline methods are defined to serve as a reference point for comparison: a traditional rule-based matching approach based on the DL/T 5729 standard lookup table, a random forest classifier with hand-crafted topological features, a GCN model with the same two-stage structure but using standard graph convolutional layers instead of GAT layers, and a GraphSAGE model with the same structure but again using standard graph convolutional layers instead of GAT layers. The ablation experiments involve the following variants of the model: removal of the attention mechanism to reduce to uniform aggregation, removal of the concatenation of global topology communication features, removal of communication features and keeping only electrical-topological features, removal of Stage 2 to reduce to end-to-end single-task classification on 12 classes, and removal of edge features and keeping only node features.

3. Results

3.1 Dataset Statistical Characteristics

The 360 scenario samples are evenly distributed across the four wiring modes, with 90 scenario samples for each wiring mode. Table 1 shows a summary of the key statistical characteristics. The number of nodes varies between 33 (for simple R1 topologies based on the IEEE 33-node system) and 72 (for complex L2 topologies based on the IEEE 69-node system with added ring connections). The number of edges varies between 32 and 78. The number of loops is zero for all radial scenarios R1 and R2, and up to 3 for the three-supply-one-backup ring mode L2. The average active power load across all scenarios is 3.2 MW with a standard deviation of 0.48 MW, which reflects the $\pm 15\%$ perturbation.

Based on the statistics on the communication infrastructure, differentiation is observed in terms of FA level, with FA-I scenarios having zero communication coverage by definition, while FA-II scenarios have an average coverage ratio of 0.72, with L2 topologies having a higher value of 0.76 due to the higher number of switch terminals, and FA-III scenarios having an average coverage of 0.95. The investment cost estimates, which include the breakdown of costs into primary equipment and communication infrastructure, indicate that the cost of communication is 34% of the total cost in FA-II scenarios and 47% in FA-III scenarios. The SAIDI values vary from 1.8 h/(customer $\cdot$ year) in the best-performing L2 + FA-III scenarios to 8.5 h/(customer $\cdot$ year) in the R1 + FA-I scenarios, with ring topologies performing better than radial ones and with higher FA levels having lower values, although with diminishing improvement.

Table 1. Statistical characteristics of the distribution network scenario dataset

Statistical Item	R1	R2	L1	L2	Overall
Number of samples	90	90	90	90	360
Node count range	33–36	33–38	35–68	42–72	33–72
Edge count range	32–35	35–42	38–65	48–78	32–78
Loop count	0	0	1–2	2–3	0–3
Average node degree	1.94	2.12	2.17	2.29	2.13
Mean active load (MW)	2.8	3.1	3.3	3.6	3.2
Std. active load (MW)	0.42	0.45	0.50	0.54	0.48
FA-I cost (10^3 CNY/circuit)	150	150	150	150	150
Primary equipment	150	150	150	150	150
Communication infra.	0	0	0	0	0
FA-II cost (10^3 CNY/circuit)	350	350	350	350	350
Primary equipment	230	230	230	230	230
Communication infra.	120	120	120	120	120
FA-III cost (10^3 CNY/circuit)	600	600	600	600	600
Primary equipment	320	320	320	320	320
Communication infra.	280	280	280	280	280

FA-I comm. coverage	0	0	0	0	0
FA-II avg. comm. coverage	0.68	0.70	0.73	0.76	0.72
FA-III avg. comm. coverage	0.93	0.94	0.95	0.97	0.95
SAIDI range (h/(cust.·yr))	5.2–8.5	3.8–7.1	2.6–5.4	1.8–4.2	1.8–8.5

3.2 Wiring Mode Identification Performance

The overall four-class wiring mode identification accuracy of the GAT model reaches 91.7% with a weighted F1-score of 0.908. As depicted in Table 2, the recall rates of each class show that the classes are not equally difficult to distinguish. L2, i.e., the three-supply one-backup ring topology, achieves the highest recall of 96.3%, mainly because of the unique dual-ring topology with various source points. R1, i.e., the single radial topology, ranks second with a recall of 93.1%, mainly because of the tree topology without any loops. L1,

i.e., the two-supply one-backup ring topology, reaches a recall of 89.5%, while R2, i.e., the multi-section single-tie radial topology, achieves the lowest recall of 87.8%.

The confusion matrix in Figure 5 also sheds more light on the misclassification issue. The main confusion occurs between the R2 and L1 classes, with a 7% misclassification error from the R2 to the L1 class and a 6% error from the L1 to the R2 class. This is due to the topological similarity between multi-section radial networks operating with an open tie switch and two-supply ring networks in their open-loop state; both have a similar branching topology when the tie switch is open. All other confusion rates are below 4%.

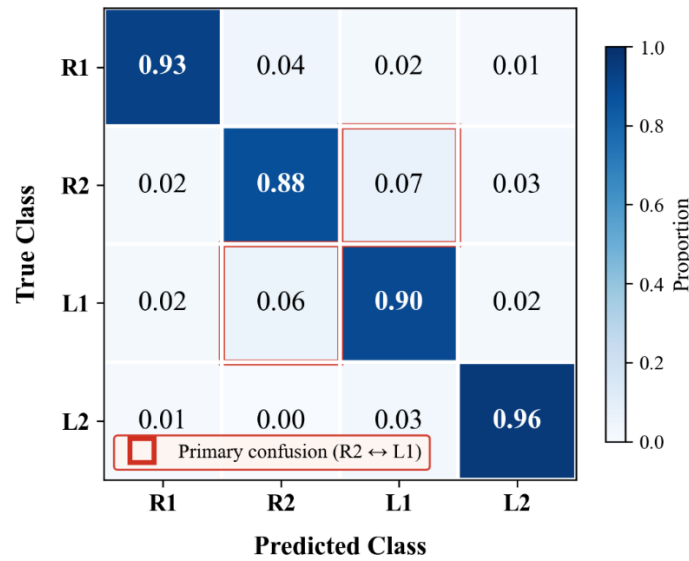


Figure 5. Normalized confusion matrix for GAT-based wiring mode identification
The primary confusion occurs between R2 (multi-section radial) and L1 (two-supply ring) due to their topological similarity in open-loop states

Table 2. Comparison of wiring mode identification performance across methods

Method	Acc. (%)	W-F1	R1 Rec. (%)	R2 Rec. (%)	L1 Rec. (%)	L2 Rec. (%)
Random Forest	82.4	0.819	85.2	78.6	80.1	85.7
GCN (same arch.)	86.3	0.856	88.9	83.3	84.8	88.5
GraphSAGE (same arch.)	88.0	0.873	90.7	85.2	86.4	89.6
GAT (proposed)	91.7	0.908	93.1	87.8	89.5	96.3

Among the basic models, the accuracy of the GCN model with the same structure is 86.3%, the GraphSAGE model reaches 88.0%, and the Random Forest model reaches 82.4%. The accuracy of the GAT model exceeds the GCN model by 5.4%, which again proves the effectiveness of the attention mechanism in adaptively weighing topologically important nodes like tie points and sectionalizing switches in the topology classification of the distribution network.

3.3 FA Configuration Optimal Matching Recommendation Performance

On the premise that the correct wiring mode is recognized, the accuracy of Top-1 matching is 83.3%, the Top-2 hit rate is 94.4%, and MRR is 0.893, as shown in Table 3. The constraint on the communication feasibility has a clear

filtering effect in the recommendation phase. Under conditions in which the communication coverage ratio is less than 0.6, the constraint mechanism filters out the FA-III in the recommendation process, avoiding 4 cases of infeasible deployment suggestions.

If the analysis is done on a per-mode basis, the highest Top-1 accuracy is observed in R1 scenarios, with 88.9%, since the selection decision in single-radial networks is relatively simple, with the model largely preferring either FA-I or FA-II depending on the communication readiness. The L2 scenarios have the lowest Top-1 accuracy, at 77.8%, since the reliability improvement from FA-II to FA-III in three-supply ring networks is relatively low (0.6 h SAIDI) while the investment in the communication infrastructure is substantial, with boundary cases where the score difference is small (0.03 to 0.08).

Table 3. Comparison of the FA configuration optimal matching recommendation performance

Method	Top-1(%)	Top-2(%)	MRR	R1(%)	R2(%)	L1(%)	L2(%)
Rule-based (lookup)	68.5	79.6	0.762	77.8	66.7	66.7	63.0
Random Forest	74.1	85.2	0.812	81.5	70.4	74.1	70.4
GCN (same arch.)	79.6	90.7	0.858	85.2	77.8	81.5	74.1
GraphSAGE (same arch.)	81.5	92.6	0.875	85.2	81.5	81.5	77.8
GAT (proposed)	83.3	94.4	0.893	88.9	81.5	85.2	77.8

In comparison with the baselines, the rule-based method achieves 68.5% Top-1 accuracy using the rule-based method, 74.1% using the Random Forest method, and 79.6% using the GCN method. The GAT model's 14.8% gain over the rule-based method is mainly due to the fine-grained discrimination of FA levels for the R2 and L1 wiring modes and the effectiveness of the communication constraint mechanism in eliminating infeasible recommendations.

3.4 Ablation Study Results

Table 4 shows the results of the ablation study. The accuracy of wiring mode drops to 85.9%, and Top-1 matching accuracy drops to 78.7% when the attention mechanism is removed, returning to uniform aggregation as used in GCN. The accuracy of the wiring mode drops by 1.7% when global topology communication feature concatenation is removed, but Top-1 matching accuracy drops by a larger margin of 3.7%. This suggests that global features, especially the communication coverage ratio and latency grade, play a more important role in the FA recommendation stage compared with the wiring mode identification stage. Eliminating all communication features (keeping only electrical-topological

features) results in a negligible impact on the accuracy of wiring modes (91.5%, a 0.2% decrease) but decreases Top-1 matching accuracy by 3.1% to 80.2%. This suggests that communication features are not important in topology identification but are important in FA-level matching. Eliminating Stage 2 (i.e., using an end-to-end 12-class classifier instead of our two-stage architecture) results in a 9.1% decrease in matching accuracy to 76.9%, confirming our intuition that our cascaded architecture is better than an end-to-end architecture. Eliminating edge features decreases accuracy in both wiring modes by 3.2% and in matching accuracy by 2.7%.

The distribution of the weight of the attention over various wiring modes is depicted in Figure 6. For ring-type topologies, the tie switch nodes and ring switching points have an elevated weight of attention that varies between 2.1 and 2.8 times compared to the average weight of the nodes. For radial-type topologies, the sectionalizing switch nodes and end of feeder nodes have relatively elevated weights of attention. The nodes that have the facility of communication terminals have approximately 35% higher weights of attention compared to other nodes at the same position during the recommendation of the FA.

Table 4. Ablation study results

Model Variant	Topo. Acc.(%)	Δ Acc.	Top-1 Match(%)	Δ Match
GAT full model	91.7	—	83.3	—
w/o attention (uniform agg.)	85.9	−5.8	78.7	−4.6
w/o global topo-comm features	90.0	−1.7	79.6	−3.7
w/o communication features	91.5	−0.2	80.2	−3.1
w/o Stage 2 (end-to-end 12-cls)	89.3	−2.4	76.9	−6.4
w/o edge features	88.5	−3.2	80.6	−2.7

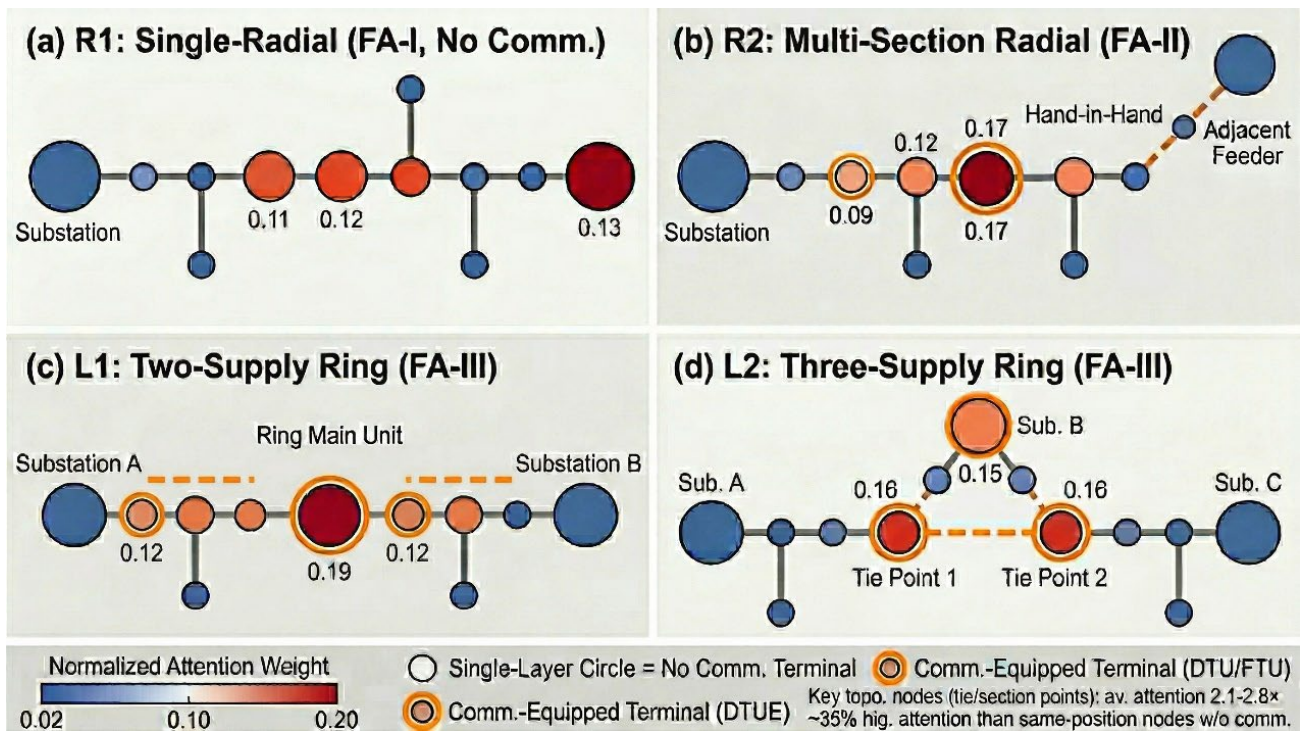


Figure 6. Visualization of GAT attention weight distribution across four wiring modes:

(a) R1; (b) R2; (c) L1; (d) L2

Node size is proportional to attention weight; color ranges from blue (low) to red (high)

Double-ring orange markers denote communication-equipped terminals

3.5 Transfer Validation on Operational Feeders

The model trained on the synthetic dataset is applied directly to the 72 Nanjing feeders without fine-tuning. The composition of this external sample across the three districts is shown in Figure 7. Wiring mode accuracy reaches 87.5%, with the 4.2-percentage-point gap relative to synthetic performance attributable to non-standard branch extensions and long-term maintenance states that blur the R2/L1 boundary. Top-1 FA matching accuracy is 79.2%, and the

ICT constraint filters out 7 infeasible suggestions in feeders with coverage below 0.6. Per-district results are reported in Table 5; performance is highest in Gulou (dense fiber coverage) and lowest in Pukou (hybrid overhead-cable sections under-represented in training).

Table 5. Performance on operational feeders by district

District	Feeders	Mode Acc. (%)	FA Top-1 (%)	FA Top-2 (%)
Gulou	26	92.3	84.6	92.3
Jiangning	28	85.7	78.6	89.3
Pukou	18	83.3	72.2	88.9
Overall	72	87.5	79.2	90.3

The per-class identification results and the operation of the ICT feasibility constraint are visualized in Figure 8. Figure 8(a) confirms that the dominant confusion occurs between R2 and L1, consistent with the synthetic-data pattern reported in Figure 3, but with higher off-diagonal proportions due to the real-world ambiguities discussed above. Figure 8(b) shows the distribution of raw FA-III scores against communication coverage, where seven feeders with coverage below 0.6 received high pre-constraint FA-III scores and were correctly filtered by the constraint mechanism.

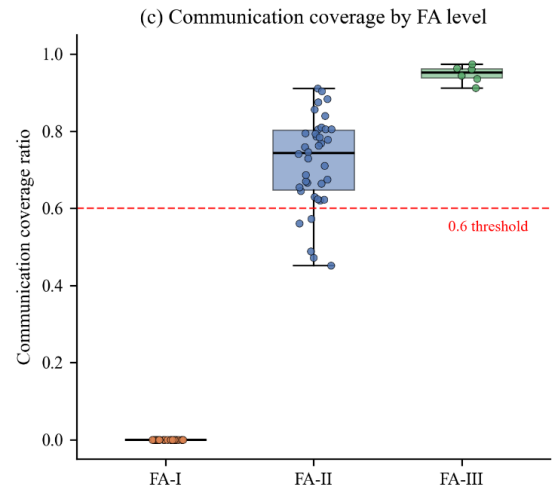


Figure 7. Composition of the 72 operational feeders:
(a) wiring mode distribution by district;
(b) FA level distribution by district;
(c) communication coverage by FA level

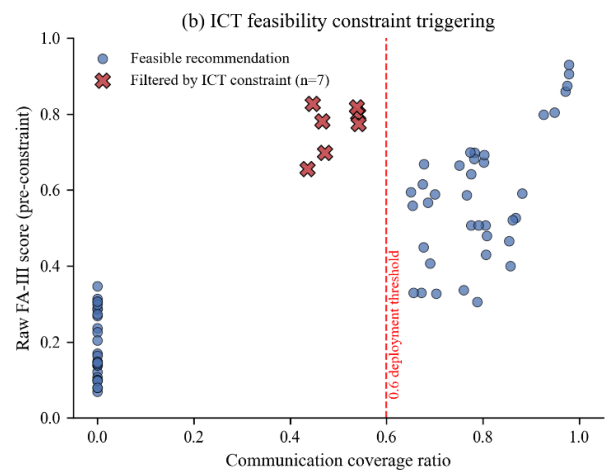
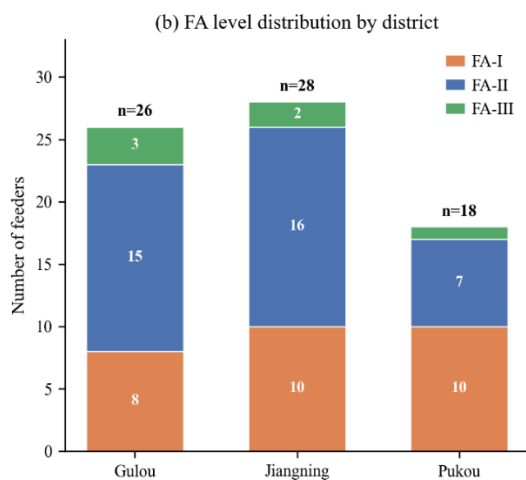
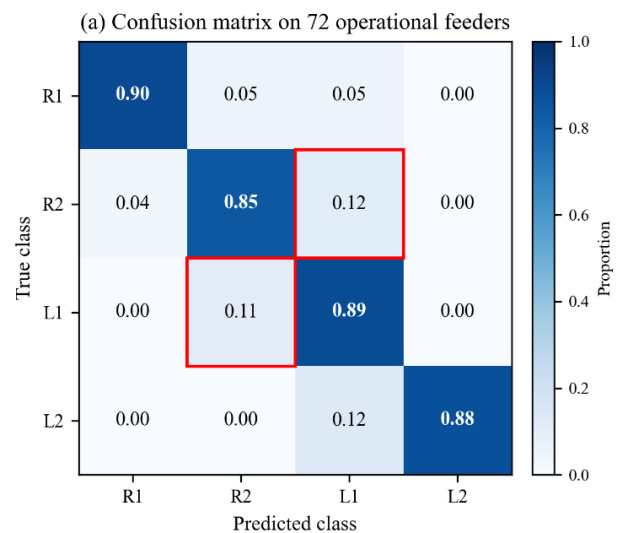
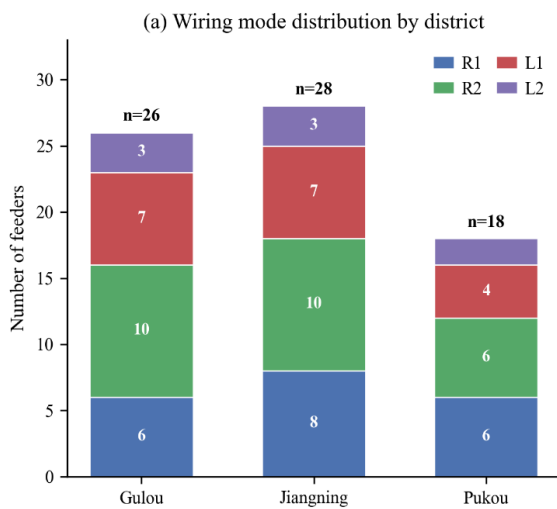


Figure 8. Error analysis on the operational sample:
(a) wiring mode confusion matrix;
(b) ICT constraint triggering against communication coverage

4. Discussion

4.1 Engineering Interpretability of Attention Weight Distributions

The attention weight patterns learned by the GAT model are in line with existing knowledge in distribution engineering. In ring topology structures, tie switch nodes and ring switching points tend to have the highest attention weights because they are the nodes that define whether a network is a ring or a radial structure. They are where the difference between L1/L2 and R1/R2 is defined. This type of interpretability was also identified in vulnerability assessment frameworks for MG based on the GAT model, where attention weights were used to identify critical nodes with high topological centrality [18]. The attention-based interpretability in optimal power flow studies has also shown that attention can be used to locate nodes that are critical to decisions in a network [19].

The fact that nodes with communication-terminal-equipped switch nodes receive 35% more attention weights in the recommendation stage of FA selection can be interpreted as an indication that the model implicitly took into account the impact of ICT readiness in selecting FAs. This is because, in engineering practices, the communication ability of a switch terminal directly influences whether a centralized or decentralized FA can be implemented in that switch node. However, it should be noted that attention weights are not a rigorous way to perform causal inference and that similar concerns are also present in interpretable wind power prediction models based on deep graph attention methods [20].

4.2 Impact of ICT Infrastructure on Matching Decisions

The ablation study verifies that communication features indeed decrease Top-1 matching accuracy by 3.1 percentage points, especially at the FA-II/FA-III decision boundary. In R2 wiring modes, the model always favors FA-II over FA-III, not only because FA-III results in only a limited improvement in SAIDI by approximately 0.6 h but also because the cost of constructing a GOOSE P2P communication link for FA-III is approximately 280,000 CNY/circuit, accounting for 47% of the total cost. This cost is higher than that of primary equipment, which makes FA-III not only less reliable but also economically unjustifiable. This analysis echoes the results of robust distribution network reconfiguration considering investment cost/economic reliability improvement trade-off [21], as well as some

comparative studies on data-driven distribution network reconfiguration decision-making methods [22].

In the case of L2 ring topologies, the difference in scores between FA-II and FA-III reduces to 0.03-0.08 in situations in which communication coverage exceeds 0.9. This is because in situations in which the ICT infrastructure is already well-developed, the question of which configuration to prefer is indeed ambiguous. The ICT feasibility hard constraint mechanism's utility is demonstrated in low-coverage situations in which the model would otherwise recommend FA-III configurations that cannot physically be deployed. This is a failure mode to which data-driven models would indeed be unable to respond. The importance of standardized frameworks in evaluating such graph-based power system models has been underscored by the PowerGraph benchmark initiative [23]. In terms of actual deployment in the hands of planners, the Top-2 recommendations of the model could be presented in conjunction with investment budget constraints and actual communication infrastructure status.

4.3 Limitations and Future Directions

There are some limitations that should be noted. First, although 72 Nanjing feeders provide external validation, this single-city sample cannot fully represent the topological diversity of Chinese urban distribution networks. Therefore, the complexity of the topology of the distribution networks in the dataset is still relatively ideal. The combined sample of 360 synthetic scenarios and 72 operational feeders remains modest for deep learning standards. The generalization of the model to larger networks still needs to be verified. The communication features are represented as discrete categories of grades rather than continuous physical parameters such as bandwidth and packet loss rate. The investment estimates of FA do not take into account the variations of the communication infrastructure costs in different regions and the aging of equipment. The impact of the penetration of DG on the wiring mode is not considered.

There are several avenues in which future improvements could be made. The addition of real SCADA and communication network management system data would greatly increase the accuracy and size of the training set. Physics-informed GNN has shown promising results in terms of improvement in the overall generalization capabilities, particularly in dynamic reconfiguration problems [24]. The addition of graph temporal attention mechanisms would allow the model to address dynamic changes in the topology and time-varying states in the communication links. The addition of 5G network slicing and edge computing to the FA deployment models is a new avenue of research, considering the rapid deployment of 5G infrastructure in Chinese cities. Physics-informed graph neural networks in collaborative dynamic reconfiguration and voltage regulation in unbalanced distribution systems [25] is another avenue in which future improvements could be made to the model.

5. Conclusion

This paper attempts to address the problem of a lack of systematic matching between 10 kV urban distribution network wiring modes and three-level feeder automation configurations, especially in the context of the often-neglected constraint related to the readiness of the communication infrastructure. A two-stage matching model using graph attention networks is proposed, which includes a unified framework combining electrical, topological, and ICT-related characteristics. The first stage employs the topological awareness of GATs to automatically identify four different types of wiring modes, which include single radial R1, multi-section radial R2, two supply ring L1, and three supply ring L2. The second stage proposes the optimal level of FA configuration, which includes FA-I, FA-II, and FA-III, using a unified framework combining graph-level embeddings and global topology-communication features, as well as ICT feasibility hard constraints.

Experimental evaluation on a 360-scenario dataset created based on IEEE 33/69-node test systems with MDFA feature augmentation leads to the following significant findings. The GAT model attains 91.7% accuracy in identifying the wiring mode four-class problem and 83.3% accuracy in FA configuration Top-1 matching accuracy. This performance is superior to the performance of GCN (86.3% and 79.6%), GraphSAGE (88.0% and 81.5%), and the Random Forest model (82.4% and 74.1%). The addition of communication features to the model contributes 3.1% to FA matching accuracy. The ICT feasibility constraint ensures that infeasible FA-III configurations are not recommended in low-communication-coverage scenarios. Attention weights show that tie points, sectionalizing switches, and communication-equipped terminals are automatically identified by the model and assigned 2.1-2.8 times more weights than regular load nodes. The two-stage cascaded model outperforms the end-to-end single-task model by 6.4% in matching accuracy. External validation on 72 Nanjing feeders attains 87.5% wiring mode accuracy and 79.2% Top-1 FA matching, quantifying the residual domain shift.

The above results offer an ICT constraint awareness, data-based decision reference for distribution network planners in choosing wiring modes and designing automation configuration schemes. The validity of the above method is still tested on typical test systems, while its applicability to real-world complex distribution networks with higher scales, higher variability in loads, and diversity in ICT infrastructures is to be researched.

References

- [1] El Berkaoui, A., et al. Distribution network topology planning and optimization: A brief review. in E3S Web of Conferences. 2023. EDP Sciences.
- [2] Lotfi, H., M.E. Hajiabadi, and H. Parsadust, Power distribution network reconfiguration techniques: A thorough review. Sustainability, 2024. 16(23): p. 10307.
- [3] Wang, H., et al., Optimal Configuration of Feeder Terminal Units in Power Distribution Networks Considering Distributed Generation. Electronics, 2025. 14(11): p. 2117.
- [4] Sun, Y. and W. Wu, Feeder automation deployment optimization for resilience reinforcement of distribution networks. IEEE Transactions on Power Delivery, 2025.
- [5] Liao, W., et al., A review of graph neural networks and their applications in power systems. Journal of Modern Power Systems and Clean Energy, 2021. 10(2): p. 345-360.
- [6] Veličković, P., et al., Graph attention networks. arXiv preprint arXiv:1710.10903, 2017.
- [7] Zhang, R., et al., A graph attention networks-based model to distinguish the transient rotor angle instability and short-term voltage instability in power systems. International Journal of Electrical Power & Energy Systems, 2022. 137: p. 107783.
- [8] Xing, Q., et al., Real-time optimal scheduling for active distribution networks: A graph reinforcement learning method. International Journal of Electrical Power & Energy Systems, 2023. 145: p. 108637.
- [9] Nguyen, B.L., et al., Spatial-temporal recurrent graph neural networks for fault diagnostics in power distribution systems. IEEE Access, 2023. 11: p. 46039-46050.
- [10] Ngo, Q.-H., et al., Deep graph neural network for fault detection and identification in distribution systems. Electric Power Systems Research, 2025. 247: p. 111721.
- [11] Yaniv, A., P. Kumar, and Y. Beck, Towards adoption of GNNs for power flow applications in distribution systems. Electric Power Systems Research, 2023. 216: p. 109005.
- [12] Lei, Y., et al., A topology identification strategy of low-voltage distribution grids based on feature-enhanced graph attention network. Energies, 2025. 18(11): p. 2821.
- [13] Zhang, Y., P.M. Karve, and S. Mahadevan, Graph neural networks for power grid operational risk assessment under evolving unit commitment. Applied Energy, 2025. 380: p. 124793.
- [14] Jacob, R.A., et al., Real-time outage management in active distribution networks using reinforcement learning over graphs. Nature Communications, 2024. 15(1): p. 4766.
- [15] Zhao, C., X. Li, and Y. Cai, A power grid topology detection method based on edge graph attention neural network. Electric Power Systems Research, 2025. 239: p. 111219.
- [16] Schneider, K.P., et al., Analytic considerations and design basis for the IEEE distribution test feeders. IEEE Transactions on power systems, 2017. 33(3): p. 3181-3188.
- [17] Zhang, H., et al., Topology detection of distribution grid using wavelet-based Graph Attention Network. Electric Power Systems Research, 2025. 248: p. 111848.
- [18] Liu, W., et al., Graph Attention Networks Unleashed: A Fast and Explainable Vulnerability Assessment Framework for Microgrids. Expert Systems with Applications, 2026: p. 131837.
- [19] Li, C., et al., Optimal Power Flow in a highly renewable power system based on attention neural networks. Applied Energy, 2024. 359: p. 122779.
- [20] Zhang, J., et al., Interpretable wind power short-term power prediction model using deep graph attention network. Energies, 2024. 17(2): p. 384.
- [21] Choobdari, M., et al., Robust distribution networks reconfiguration considering the improvement of network resilience considering renewable energy resources. Scientific reports, 2024. 14(1): p. 23041.
- [22] Gholizadeh, N., N. Kazemi, and P. Musilek, A comparative study of reinforcement learning algorithms for distribution network reconfiguration with deep Q-learning-based action sampling. Ieee Access, 2023. 11: p. 13714-13723.

- [23] Varbella, A., et al., PowerGraph: A power grid benchmark dataset for graph neural networks. *Advances in Neural Information Processing Systems*, 2024. 37: p. 110784-110804.
- [24] Authier, J., et al., Physics-informed graph neural network for dynamic reconfiguration of power systems. *Electric Power Systems Research*, 2024. 235: p. 110817.
- [25] Qin, J., R. Yang, and N. Yu, Physics-informed graph neural networks for collaborative dynamic reconfiguration and voltage regulation in unbalanced distribution systems. *IEEE Transactions on Industry Applications*, 2025. 61(2): p. 2538-2548.