

A domain knowledge graph-driven joint recommendation system for wiring mode and automation configuration in differentiated power supply zones

Min Li^{1,*}, Junhua Gong¹, Jing Tan¹, Feipeng Huang¹, Xiaohong Tan¹

¹Guangxi Power Grid Co., Ltd., Nanning 530023, China

Abstract

INTRODUCTION: Distribution network planning requires jointly determining wiring modes and automation configurations, yet existing methods treat them independently without regulatory constraints.

OBJECTIVES: To develop a knowledge graph-driven framework for joint wiring mode and automation configuration recommendation across differentiated supply zones.

METHODS: A BERT-CRF pipeline constructs a knowledge graph from regulatory texts. TransR-initialized graph attention networks and CRF decoders with SWRL constraints are coupled via multi-task learning, evaluated on 1,443 records and 12 field tasks.

RESULTS: Acc@3 of 0.817, F1 of 0.757, and CCR of 0.733, surpassing the best baseline by 4.3% and 10.4% respectively. Expert validation scores 4.50/5.00 with 73.3% planning time reduction.

CONCLUSION: Integrating knowledge graph embedding, GNN aggregation, and constraint-aware joint optimization yields accurate, compliant, and traceable distribution planning recommendations.

Keywords: domain knowledge graph, graph neural network, multi-task joint learning, distribution network planning, knowledge reasoning

Received on 16 January 2026, accepted on 17 May 2026, published on 12 October 2026

Copyright © 2026 Min Li *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.14666

*Corresponding author. Email: 49504072@qq.com

1. Introduction

The rapid advancement of smart grid technology has significantly impacted the operational paradigms of modern power distribution systems. The increasing trend of distributed power generation, the complexity of power loads, and the ever-tightening reliability constraints have cumulatively put unprecedented pressure on the advancement of smart grid-based methodologies for power distribution system planning [1]. Distribution network planning, an essential part of power system operations, involves a sophisticated array of interrelated decisions regarding network topology, wiring mode, and automation configuration, all of which are to be simultaneously

optimized to meet the differentiated reliability and serviceability requirements of heterogeneous power supply zones [2]. Under the existing power distribution grid governance structure, power supply zones are classified into five categories of A+, A, B, C, and D, each of which is subject to specific regulatory requirements regarding wiring mode, automation terminal density, and fault self-healing thresholds. This configuration complexity poses a significant challenge to existing power distribution network planning methodologies [3]. The basic issue here is not merely to select an optimal wiring topology according to the specific power supply zone type. Rather, it involves simultaneously configuring the associated automation configuration, such as

the configuration of feeder terminal units, distribution terminal units, and communication protocols, to be consistent with the specific regulatory requirements of each power supply zone type. This issue has generally been left to the experiential expertise of power distribution network planning engineers rather than any systematic decision support methodology, a situation that becomes increasingly unsustainable with the ever-increasing complexity of power grid operations.

The advent of knowledge graphs as an organized model for knowledge representation provides a promising avenue for intelligent decision support in power system engineering. An increasing number of studies show its potential in representing the intricate entity-relation knowledge typical in the domain [4]. Domain knowledge graphs built from existing knowledge bases, operational data, and engineering standards have proven to be effective in supporting dispatch decisions and other operations-related tasks by modeling heterogeneous relational knowledge in an expressive yet computationally efficient manner [5]. Moreover, topology identification based on graph structures has extended this capability to network topology analysis in uncertain conditions or with missing measurements [6]. Graph neural networks have also proven to be an important tool in modeling the intricate relationships in power system topology with considerable efficacy in tasks ranging from fault prediction to optimal power flow approximation [7]. In recommendation systems, knowledge graph-enhanced neural network architectures have shown that domain knowledge can play an important role in improving both recommendation accuracy and model interpretability in comparison to traditional collaborative filtering methods [8]. Furthermore, multi-task learning frameworks based on knowledge graph representations have shown that optimizing multiple interdependent prediction tasks can significantly improve model generalization compared to single-task optimization methods [9]. Knowledge graph-enhanced configuration recommendations have also been studied in communication network management systems where device relationship knowledge can be used to provide more interpretable configuration recommendations [10]. The integration of deep learning with knowledge graph representations of operational constraints has shown considerable improvement in security-constrained optimization in power systems [11].

Despite such advances, an important methodological void appears to persist within the existing body of literature. Existing methodologies generally treat the selection of wiring mode and the design of automation configuration as relatively independent planning problems, without a principled joint modeling methodology that simultaneously respects zone-specific regulatory constraints, accounts for the semantic relationship between the two decision spaces, and produces recommendation outputs that can be traced to verifiable domain knowledge. This issue appears to be especially relevant to differentiated supply zone planning problems, where the joint decision of wiring mode selection and configuration design should be simultaneously treated under a single regulatory umbrella to ensure grid reliability.

Furthermore, the lack of any structured mechanisms to represent domain knowledge means that critical regulatory constraints embedded within technical specifications are largely inaccessible to data-driven methodologies, potentially resulting in statistically valid yet technically non-compliant or inconsistent results.

To this end, this research proposes a domain knowledge graph-driven framework to jointly recommend wiring modes and automation configurations in the context of differentiated power supply zones. Regulatory knowledge, extracted from technical specifications, is systematically encoded in a structured graph representation, which simultaneously enables the recommendation of wiring modes by graph neural networks and automation configurations by knowledge reasoning [12]. A multi-task joint learning framework is proposed to jointly address the two recommendation objectives by coupling them in a shared knowledge graph embedding space, which enables end-to-end co-optimization under zone-specific regulatory constraints while producing interpretable outputs supported by traceable reasoning paths based on encoded domain knowledge [13]. The proposed framework is an advancement in the existing literature by bridging the existing gap between isolated planning decision support tools and integrated knowledge-driven recommendation systems, which have the potential to jointly address the semantic dependencies and regulatory constraints in differentiated supply zone planning, thus providing a transferable research methodology with important implications for the digital transformation of power distribution grid governance.

2. Data and methods

2.1. Data description and preprocessing

The empirical basis of the proposed framework is grounded upon a multi-source heterogeneous dataset sourced from a regional power distribution operator in China. The dataset includes geographic information system records of existing distribution network assets, structured regulatory specification texts defining wiring and automation specifications for specific zones, and historical planning case archives recording previously approved configuration choices for supply zones of varying classification. The combination of these different data modalities is grounded upon the understanding that there is no single modality that is capable of fully encapsulating the semantic, structural, and regulatory aspects of the problem space to enable meaningful joint recommendations; knowledge graph construction in complex problem spaces such as this inherently requires the simultaneous processing of structured relational data and unstructured technical corpora [14]. Prior to the construction of the model, standardized preprocessing of the text data is performed to extract regulatory clauses and entity-candidate pairs, with the geographic information system records being standardized to a single classification schema to enable

semantic consistency across the dataset. The distribution of the compiled dataset is reported in Table 1.

Table 1. Multi-source dataset statistical description

Zone Type	Sample Size	Regulation Text Entries	Wiring Mode Categories	Configuration Dimensions	Primary Data Sources
A+	156	172	6	23	GIS Records, Regulation Texts, Historical Cases
A	298	143	5	19	GIS Records, Regulation Texts, Historical Cases
B	412	118	4	16	GIS Records, Regulation Texts, Historical Cases
C	364	87	3	12	GIS Records, Regulation Texts, Historical Cases
D	213	62	2	8	GIS Records, Regulation Texts
Total	1,443	582	6	23	—

As can be seen in Table 1, it is clear that both regulatory text density and configuration dimensionality are monotonically reduced from A+ to D zones in line with the reliability standards defined in China's differentiated supply zone governance structure. Moreover, the relative asymmetry in sample count in A+ zones where the lowest sample count corresponds to the highest regulatory specification density further reinforces the argument that zone classification should be treated as a primary conditioning variable in knowledge graph ontology building and recommendation modeling.

2.2. Automatic construction of the domain knowledge graph via natural language processing

The domain knowledge graph is built through a multi-stage natural language processing pipeline that extracts structured relational knowledge from regulatory specification texts and transforms it into a semantically meaningful graph. The architecture of the pipeline is shown in Fig. 1.

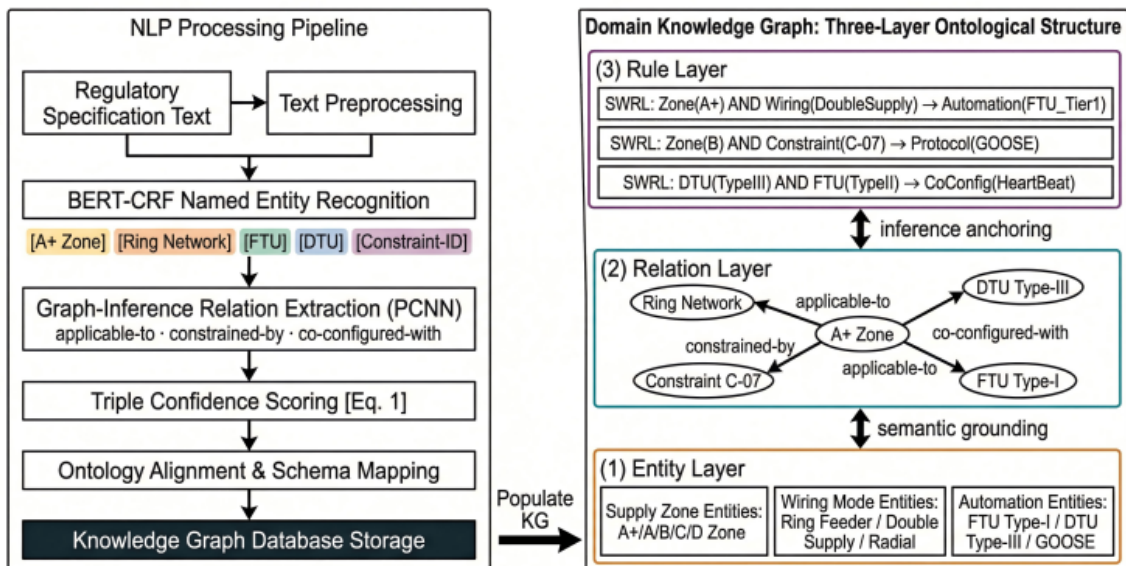


Figure 1. Overall architecture of the domain knowledge graph automatic construction pipeline

As depicted in Figure 1, named entity recognition of power-domain entities, including supply zone type, wiring mode designation, automation terminal category, and regulatory constraint identifiers, is achieved through a BERT-CRF model fine-tuned on a domain-annotated dataset. This model utilizes the power of bidirectional transformer-based representations with conditional random fields, which have been shown to provide substantial precision improvements over traditional taggers, especially for texts with high terminology density, as typically found in power engineering domains [15]. Relation extraction of the identified entities, including predicates such as “applicable-to,” “constrained-by,” and “co-configured-with,” across sentence boundaries, is achieved through a graph inference-enhanced model, which utilizes the power of graph augmentation-based encoding architectures, which have been proven effective for representing cross-sentence interactions between entities [16].

Each extracted candidate triple is assigned a composite confidence score jointly evaluating extraction pipeline reliability and embedding-space plausibility, as formalized in Equation (1):

$$s(h, r, t) = \beta \cdot p_{\text{NER}}(h) \cdot p_{\text{RE}}(r | h, t) \cdot p_{\text{NER}}(t) + (1 - \beta) \cdot \sigma(\gamma - \|\mathbf{h}_r + \mathbf{r} - \mathbf{t}_r\|_2) \quad (1)$$

where $p_{\text{NER}}(h)$ and $p_{\text{NER}}(t)$ denote entity recognition confidence scores, $p_{\text{RE}}(r | h, t)$ denotes the conditional relation extraction probability, $\mathbf{h}_r$ and $\mathbf{t}_r$ are relation-projected entity embeddings in the TransR space, $\mathbf{r}$ is the relation embedding, γ is a margin threshold, and $\beta \in (0, 1)$ balances extraction confidence against embedding-space consistency. Retained triples are organized into a three-layer ontological structure comprising an entity layer encoding zone types and configuration elements, a relation layer encoding regulatory predicates as typed directed edges, and a rule layer encoding SWRL-compliant logical constraints enabling formal compliance inference.

2.3. Knowledge graph embedding and GNN-based wiring mode recommendation

The wiring mode recommendation module operates over the domain knowledge graph to produce zone-conditioned rankings. The initialization of the entity embeddings is done by the TransR knowledge graph embedding method. This maps the vector representations of entities to relation-specific subspaces by using learnable projection matrices, as shown in Equation (2):

$$\mathbf{e}_r = \mathbf{e}\mathbf{M}_r, \quad f_r(h, t) = \|\mathbf{h}_r + \mathbf{r} - \mathbf{t}_r\|_2^2 \quad (2)$$

where $\mathbf{e} \in \mathbb{R}^k$ is the entity embedding, $\mathbf{M}_r \in \mathbb{R}^{k \times d}$ is

the relation-specific projection matrix, and $f_r(h, t)$ is the translational distance score. This relation-specific projection is particularly suitable for the differentiated supply zone knowledge graph, where the entity is involved in semantically differentiated roles across different categories of zones, and where aggregation of these roles in a shared embedding space would lose sight of planning constraints for individual zones. The research in multi-relational knowledge graph embedding has shown that relation-specific spaces greatly enhance the discriminative capability for recommendation tasks that involve heterogeneous regulatory knowledge [17].

The TransR-initialized embeddings are refined through a relation-aware graph attention network, with attention coefficients computed as in Equation (3):

$$\alpha_{vu}^{(l)} = \frac{\exp(\text{LeakyReLU}(\mathbf{a}^\top [\mathbf{W}^{(l)} \mathbf{h}_v^{(l)} \parallel \mathbf{W}^{(l)} \mathbf{h}_u^{(l)} \parallel \mathbf{r}_{vu}]))}{\sum_{w \in \mathcal{N}(v)} \exp(\text{LeakyReLU}(\mathbf{a}^\top [\mathbf{W}^{(l)} \mathbf{h}_v^{(l)} \parallel \mathbf{W}^{(l)} \mathbf{h}_w^{(l)} \parallel \mathbf{r}_{vw}]))} \quad (3)$$

where $\mathbf{r}_{vu}$ is the relation type embedding connecting nodes v and u , and $\parallel$ denotes vector concatenation. The explicit inclusion of relation-type embeddings ensures that different edge types with different semantics will have different aggregation weights, a property of substantive significance for a regulatory knowledge graph. The node representations are updated according to Equation (4), while the compatibility of zone-conditioned wiring modes is evaluated through the bilinear function as defined in Equation (5):

$$\mathbf{h}_v^{(l+1)} = \sigma(\sum_{u \in \mathcal{N}(v)} \alpha_{vu}^{(l)} \mathbf{W}^{(l)} \mathbf{h}_u^{(l)}) \quad (4)$$

$$f_w(z, m) = \mathbf{z}^\top \mathbf{W}_s \mathbf{m} + \mathbf{b}_s \quad (5)$$

where $\mathbf{z}$ is the zone embedding after L aggregation layers, $\mathbf{m}$ is the candidate wiring mode embedding, and $\mathbf{W}_s$ is a bilinear interaction matrix. The use of such knowledge graph-based techniques can be especially beneficial for low-frequency zone categories like A+, where there are limited historical instances to generalize from without the use of such an alternative representational channel of regulatory knowledge [18].

2.4. Knowledge reasoning-based constraint-aware automation configuration recommendation

The automation configuration recommendation module is designed to jointly determine a coherent set of configuration items, which are feeder terminal unit type, distribution terminal unit deployment scheme, communication protocol selection, and inter-device coordination parameters, all within the regulatory constraint network represented by the knowledge graph rule layer. These configuration items are

not individually determined but are subject to intricate pairwise and higher-order constraints resulting from technical specifications, such that the selection of a specific terminal unit type within a specific zone may require specific communication protocols while at the same time disallowing specific deployment schemes. Treating these configuration items as independent prediction problems results in outputs that may be statistically valid but regulatory non-compliant, a situation that has direct reliability implications on the grid, prompting the formalization of constraints via a knowledge graph [19]. The joint probability of configuration conditioned on the zone and wiring mode is represented by the conditional random field formulation as given by Equation (6):

$$P(\mathbf{c} | z, m) = \frac{1}{Z(z, m)} \exp(\sum_{i=1}^K \phi_i(c_i, z, m) + \sum_{i < j} \psi_{ij}(c_i, c_j)) \quad (6)$$

where $\phi_i(c_i, z, m)$ is the unary potential capturing compatibility between item c_i and context (z, m) , $\psi_{ij}(c_i, c_j)$ is the pairwise potential encoding inter-item dependencies, and $Z(z, m)$ is the partition function. A SWRL-based constraint penalty is incorporated into the decoding objective as defined in Equation (7):

$$P(\mathbf{c}, \mathcal{R}) = \sum_{r_k \in \mathcal{R}}^{\sum_{i=1}^{0,1} [c_i = r_k]} \max \quad (7)$$

where $\mathcal{R}$ is the set of applicable SWRL rules for the given zone-wiring combination and $\mathbf{1}[c_i = r_k]$ indicates constraint satisfaction. The optimal configuration $\mathbf{c}^*$ maximizes Equation (6) subject to minimization of Equation (7), ensuring outputs that are simultaneously statistically consistent with historical patterns and formally compliant with encoded regulatory knowledge.

2.5. Multi-task joint learning end-to-end recommendation framework

The two recommendation modules are contained within a unified multi-task learning paradigm that simultaneously optimizes both objectives using a shared knowledge graph embedding space. This ensures mutual improvement rather than separate optimization. The mutual dependency between wiring mode selection and automation configuration design, as well as the dependency of wiring mode feasibility on compatible automation solutions, provides a principled reason to conduct joint learning. This is consistent with evidence that shared representational capacity across interdependent prediction tasks results in substantial improvements to generalization over separate-task models [20]. The overall architecture is shown in Fig. 2.

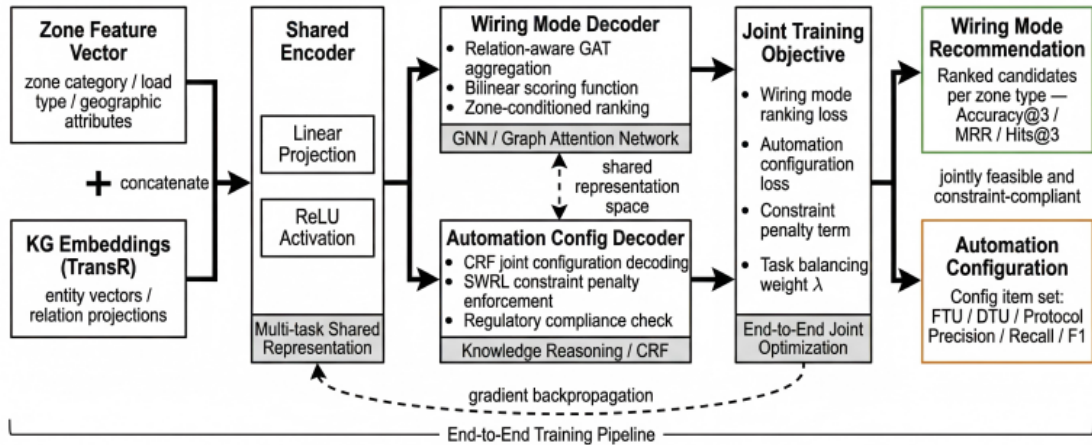


Figure 2. End-to-end multi-task joint learning recommendation framework architecture

As shown in Figure 2, a shared encoder maps zone feature vectors concatenated with the TransR initialized knowledge graph embeddings to a shared space to obtain a unified contextual representation that conditions both decoder heads. The wiring mode decoder uses relation-aware graph attention aggregation and the bilinear scoring function defined in Equation (5), while the automation configuration decoder uses the Conditional Random Field defined in Equation (6) with the constraint penalty defined in Equation (7). The joint training objective is defined in Equation (8):

$$\mathcal{L} = \mathcal{L}_{\text{wiring}} + \lambda, \mathcal{L}_{\text{automation}} \quad (8)$$

where $\mathcal{L}_{\text{wiring}}$ is the cross-entropy loss over wiring mode rankings, $\mathcal{L}_{\text{automation}}$ is the composite automation configuration loss incorporating negative log-likelihood and constraint penalty terms, and λ is a task balancing hyperparameter selected through cross-validation. This optimization process aligns with model-based reinforcement learning paradigms in distribution network operations in

which stable learning under regulatory constraints requires similar frameworks to encode interdependent objectives in a joint manner [21]. Upon convergence, recommendations are provided at a zone level that are jointly feasible and mutually consistent with zone-level regulatory constraints while being traceable to reasoning paths in the knowledge graph.

3. Results

3.1. Experimental setup and evaluation metrics

The experimental evaluation uses a multi-source dataset, which is collected from the regional distribution operator described in Section 2.1, containing 1,443 historical planning case records across all five categories of supply zones. The zone-stratified partitioning is used to split the dataset into 70% for training, 15% for validation, and 15% for testing, ensuring that the distributional properties of each zone category are preserved across all subsets rather than being concentrated in the training set. The partitioning strategy is particularly relevant to A+ zones, which comprise just 10.8% of the total records but have the highest density of regulatory specification.

Five baseline methods are selected with the aim of establishing a structured comparative reference, covering classical as well as contemporary advances in knowledge graph-based recommendation. Collaborative Filtering (CF) is selected as the fundamental statistical baseline, which does not rely on any form of structural knowledge representation, thus establishing a fundamental performance benchmark based purely on co-occurrence pattern matching. TransE, as a fundamental translational embedding method, was the first method to propose the concept of relational knowledge graph embedding without any form of graph aggregation, thus establishing a fundamental benchmark for the evaluation of more sophisticated forms of knowledge embedding. KGCN, as a method proposed by Wu et al. [18], utilizes knowledge graph convolutional neighborhood aggregation for recommendation, thus establishing a fundamental benchmark for the evaluation of graph-based forms of recommendation. KEGNN, as a method proposed by Lyu et al. [8], utilizes a combination of knowledge-enhanced graph neural networks with an emphasis on explainability, thus establishing a fundamental benchmark for the evaluation of GNN-based forms of recommendation. NetCR, as a method proposed by Guo et al. [10], utilizes a form of knowledge graph-based configuration recommendation for communication network management, thus establishing a task-analogous benchmark for the evaluation of the proposed framework, as it also utilizes a form of structured device relationship knowledge for configuration recommendations. The proposed framework extends beyond the five baselines, as it seeks to integrate all four core capability dimensions (knowledge graph embedding, GNN aggregation, joint learning, and constraint reasoning). A systematic comparison of the

capabilities of the proposed framework with respect to the five baselines is presented in Table 2.

Table 2. ICT capability dimensions across baseline methods and proposed framework

Method	KGE	GNN	Joint Learning	Constraint Reasoning
CF	X	X	X	X
TransE	✓	X	X	X
KGCN	✓	✓	X	X
KEGNN	✓	✓	X	X
NetCR	✓	X	X	✓
Proposed	✓	✓	✓	✓

As can be seen from Table 2, none of the baseline methods simultaneously activate more than two ICT dimensions, while our proposed framework uniquely achieves full-dimensional integration. This explains why we can attribute any differences in performance to the cumulative interaction effects among components, rather than individual module contributions. Moving forward to our recommendation results for wiring mode and configuration recommendation tasks, we measure our performance using Accuracy@3, Mean Reciprocal Rank (MRR), and Hits@3 for the former task, where Accuracy@3 refers to the proportion of test cases where the correct wiring mode is among the top 3 predictions. In contrast, we measure our performance for configuration recommendation tasks using Precision, Recall, F1-score, and Configuration Completeness Rate (CCR), where CCR represents the proportion of recommended configuration sets where all items simultaneously meet any applicable zone-specific regulatory constraints. The later metric is more operationally demanding than individual item-level F1-score metrics. Furthermore, we also use a composite metric called Joint Completeness Rate (JCR) in our ablation study to jointly evaluate both wiring feasibility and configuration constraint compliance.

3.2. Wiring mode recommendation performance comparison

While the recommendation of wiring mode is the more structurally demanding objective in the jointly optimized pair, this is largely a result of the significant distributional imbalance in the zone categories, in addition to the extent to which historical case records support the learning of generalized patterns. The A+ zones are naturally the most problematic, in which the combination of high specification density and limited case availability results in an inadequate signal upon which data-driven approaches may effectively discriminate between wiring mode options. The results are thus disaggregated by zone category to avoid obscuring sensitivity at this level in aggregated measures, with the full results provided in Table 3.

Table 3. Wiring mode recommendation performance across zone categories

Method	A+	A	B	C	D	Avg Acc@3	Avg MRR	Avg Hits@3
CF	0.598	0.671	0.718	0.743	0.762	0.698	0.629	0.751
TransE	0.634	0.703	0.748	0.771	0.786	0.728	0.659	0.782
KGCN	0.658	0.724	0.763	0.789	0.801	0.747	0.678	0.803
KEGNN	0.681	0.741	0.779	0.806	0.819	0.765	0.697	0.821
NetCR	0.694	0.753	0.792	0.817	0.831	0.777	0.708	0.833
Proposed	0.737	0.791	0.833	0.854	0.869	0.817	0.746	0.871

As can be seen in Table 3, the proposed framework achieves the highest performance among all zone categories, with the largest gap in A+ zones, where our proposed model surpasses NetCR by 4.3% in Acc@3. In comparison with KGCN and KEGNN, which share similar graph-based aggregation mechanisms with our proposed model but lack joint optimization and constraint reasoning, our proposed model achieves 9.4% and 6.8% improvement in average Acc@3 respectively. This implies that the shared encoder does contribute to better performance. The narrowing gap in performance from A+ to D zones is also in line with our expectation that the marginal utility of regulatory knowledge decreases with an increasing number of cases.

3.3. Automation configuration recommendation performance comparison

The recommendation of the automation configuration is subject to a set of evaluation criteria that are more demanding in terms of evaluation output compared to the wiring mode recommendation, since the output is a jointly decoded set of

interdependent configuration items, each of which is subject to empirical plausibility in addition to formal compliance with the constraint encoded in the SWRL constraint network. The presence of even a single non-compliant item in an otherwise correctly decoded set of configurations results in a penalty in CCR, making this metric particularly sensitive to the constraint reasoning capabilities of the underlying architecture. The independence of individual item prediction methodologies to constraint violations means that even relatively low individual item error rates compound to result in substantial degradation in CCR with increasing dimensionality in the configuration space, making this metric a suitable choice to serve as the evaluation metric in determining constraint compliance. The impact is expected to be most pronounced in the case of CF and TransE, which lack any constraint awareness, with the extent to which this is alleviated by the partial constraint integration in NetCR providing indirect evidence for the utility of the SWRL-CRF decoding pipeline employed in the proposed framework. The distribution of the zone-level F1 score across all methodologies is visualized in Fig. 3, followed by the results in Table 4.

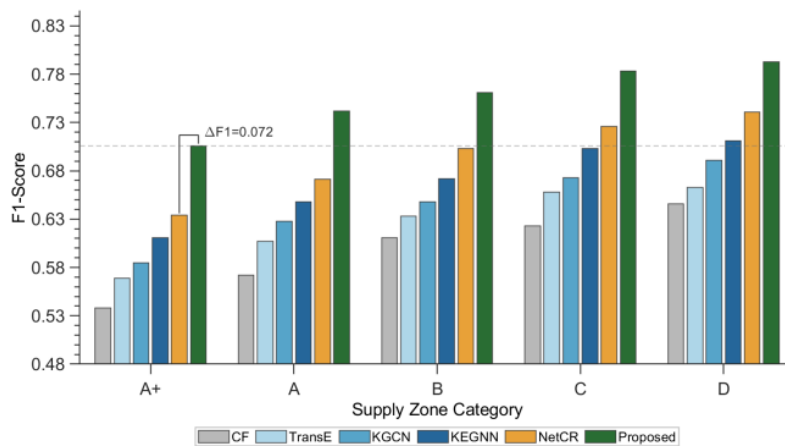


Figure 3. Grouped bar chart of automation configuration F1-score by zone category across methods

As shown in Fig. 3, F1 performance decreases across all methods from D to A+ zones. Meanwhile, the proposed framework always stays at the topmost position within each zone cluster. The gap between the proposed method and

NetCR, which performs best among all baseline methods, increases significantly across A+ and A zones. This shows that there is an additional advantage brought by GNN-based

embedding when constraint reasoning alone is not enough to handle sparse training data.

Table 4. Automation configuration recommendation performance comparison

Method	Precision	Recall	F1	CCR
CF	0.618	0.582	0.599	0.561
TransE	0.643	0.609	0.626	0.594
KGCN [18]	0.664	0.627	0.645	0.617
KEGNN [8]	0.687	0.651	0.669	0.638
NetCR [10]	0.712	0.678	0.695	0.664
Proposed	0.774	0.741	0.757	0.733

From Table 4, the proposed method achieves an F1 of 0.757 and a CCR of 0.733, improving upon NetCR by 8.9% and 10.4% respectively. The proportionally larger CCR margin confirms that constraint-aware joint decoding provides benefits beyond item-level accuracy that neither KGCN nor KEGNN can replicate through graph aggregation without integrated constraint enforcement. The Precision–Recall gap of 3.3 points is consistent with the conservative bias introduced by the SWRL penalty, which suppresses regulatory violations at a modest recall cost acceptable in grid planning practice.

3.4. Ablation study and key module contribution analysis

With the performance improvement of the proposed framework over the external baselines established, the analysis now seeks to identify which of the internal architectural components of the system contributed to the improvement. This is done through the creation of three different ablation variants, each of which removes one of the three ICT components individually: w/o NER replaces the BERT-CRF-based knowledge graph construction pipeline with a rule-based keyword matching process, which compromises both entity recognition quality and relationship extraction accuracy; w/o KGE replaces the TransR initialization of the entity embeddings with random initialization, thus eliminating the relational projection structure from the input representation of the shared encoder; and w/o Joint separately trains each of the wiring and configuration decoders as independent single-task models, thus eliminating the representational regularization effects of the shared encoder. The performance recovery of the system as each of the removed components is added back in is depicted in Fig. 4, with the actual performance figures presented in Table 5.

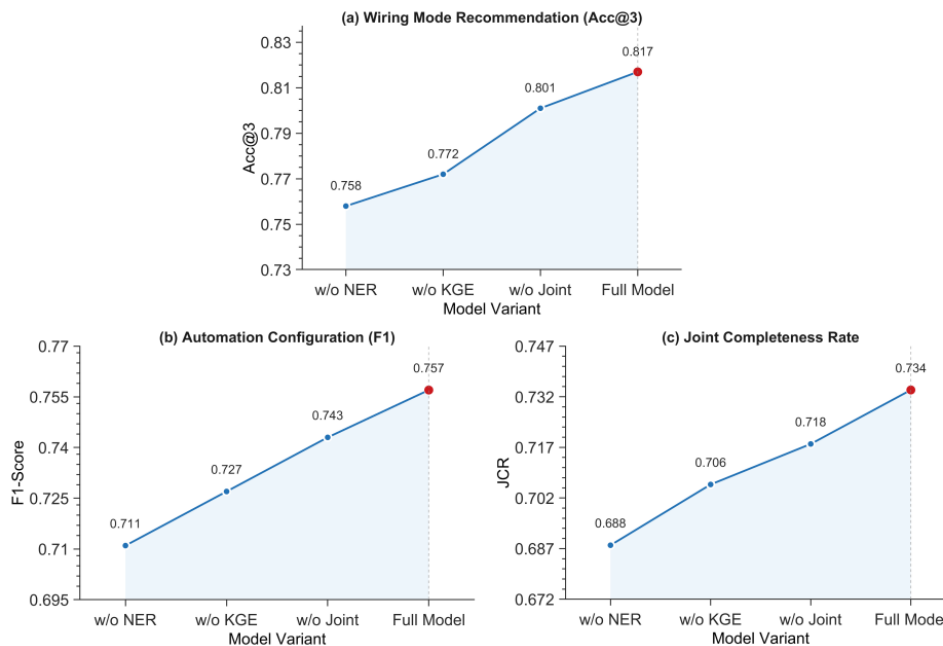


Figure 4. Line chart of ablation variant performance on Acc@3, F1, and joint completeness rate

As depicted in Fig. 4, all three measures show monotonic improvement as the components are incrementally restored from the without NER setting to the full model, with the most significant single-step gain between the without NER and without KGE configurations for Joint Completeness Rate. The without Joint setting has competitive performance for the item-level measures but a much larger gap for Joint Completeness Rate, suggesting that the shared encoder provides constraint coherence benefits that cannot be matched with single-task training.

Table 5. Ablation study quantitative results

Model Variant	Acc@3 (Wiring)	F1 (Config)	Joint Completeness
w/o NER construction	0.758	0.711	0.688
w/o KGE embedding	0.772	0.727	0.706
w/o Joint learning	0.801	0.743	0.718
Full Model	0.817	0.757	0.734

From Table 5, the removal of the NER construction pipeline has the most prominent effect of any individual module on the Joint Completeness Rate, which decreases by 4.6 percentage points compared to the full model. This shows the chain of effects of entity recognition quality on triple confidence scoring and constraint coverage. The KGE component has the most distinctive effect on wiring mode accuracy, consistent with the functionality of TransR, which captures zone-specific relational semantics not recoverable through random initialization. The inclusion of joint learning from the w/o Joint baseline increases Acc@3 by 1.6 and Joint Completeness Rate by 1.6 percentage points, demonstrating the value of cross-task representational regularization, which outperforms each individual model.

3.5. Practical application validation

To further assess the operational applicability of the proposed framework beyond historical dataset evaluation, a practical application validation was conducted in collaboration with the regional distribution network operator that provided the original dataset. The validation involved 12 newly initiated planning tasks across three supply zone categories (3 A+ zones, 4 A zones, and 5 B zones) that had not been included in the training or evaluation datasets. For each planning task, the proposed framework was deployed to generate joint recommendations for wiring mode selection and automation configuration, which were subsequently submitted to a panel of five senior distribution network planning engineers with an average of 14 years of domain experience for independent expert evaluation.

The expert evaluation was conducted along four criteria: technical accuracy of the recommended wiring mode, regulatory compliance of the automation configuration, inter-decision consistency between the jointly recommended wiring mode and automation configuration, and overall planning efficiency gain relative to conventional manual planning workflows. Each criterion was rated on a five-point Likert scale (1 = inadequate, 5 = excellent) by each evaluator independently. The averaged evaluation results are presented in Table 6.

Table 6. Expert evaluation results of practical application validation

Zone Type	Tasks	Technical Accuracy	Regulatory Compliance	Inter-Decision Consistency	Avg. Score
A+	3	4.27	4.53	4.13	4.31
A	4	4.45	4.60	4.35	4.47
B	5	4.68	4.72	4.56	4.65
Overall	12	4.49	4.63	4.38	4.50

As shown in Table 6, the proposed framework received consistently favorable evaluations across all zone categories, with overall average scores exceeding 4.3 on all criteria. Regulatory compliance received the highest average score of 4.63, which corroborates the effectiveness of the SWRL-based constraint enforcement mechanism in producing outputs that conform to zone-specific regulatory requirements in practical settings. The relatively lower scores for inter-decision consistency in A+ zones (4.13) are consistent with the findings in Section 3.2, where A+ zones exhibited the most challenging recommendation conditions due to high regulatory complexity and limited historical precedent. The expert panel further noted that the knowledge graph reasoning paths accompanying each recommendation substantially facilitated the review process by enabling engineers to trace and verify the regulatory basis of each recommended configuration item.

In terms of planning efficiency, the average time required for the expert panel to complete the review and approval of framework-generated recommendations was 2.3 hours per planning task, compared to a reported average of 8.6 hours for conventional manual planning workflows for equivalent zone categories. This represents a reduction of approximately 73.3% in planning cycle time, suggesting that the proposed framework can serve as an effective decision support tool that substantially accelerates distribution network planning processes while maintaining planning quality within acceptable professional standards.

4. Discussion

4.1. Explainability of recommendations via knowledge reasoning paths

The performance improvements outlined in Section 3 are coupled with a practically significant capability that is not reflected in aggregate performance measures: the capacity to trace each recommendation result to a verifiable chain of encoded regulatory knowledge rather than statistical correlation. More specifically, the proposed framework provides a justification path for wiring mode rankings and outputs of the automation configuration step in terms of explicit paths through the knowledge graph. Where applicable, these paths will comprise SWRL rules that encode zone tier regulatory constraints, GAT-derived attention weights that reflect the salience of neighboring entities in the knowledge graph, and TransR-projected embedding distances that provide a knowledge-graph-anchored understanding of the entities involved. The significance of this design principle is underlined by the fact that similar knowledge-graph architectures have been shown to provide substantial improvements in trustworthiness of power system diagnoses in industrial contexts, as outlined in the power engineering literature [22]. The significance of knowledge-graph-anchored explainability is underlined further by the fact that similar design principles of knowledge fusion frameworks for power system applications, such as those applied to the predetermination of generator tripping schemes, also emphasize knowledge-grounded decision-making structures to provide enhanced interpretability of operational outputs in addition to improvements in overall accuracy [23].

4.2. Impact of ICT integration on planning decision efficiency

Apart from the specific accuracy improvements documented in Tables 3 and 4, some broader implications of the results on understanding the efficiency benefits that accrue to intelligent support tools within the context of multi-dimensional ICT integration are worthy of mention. The consistent performance superiority of the full framework over individual ICT dimension baselines, as documented across all test cases, but most pronounced within the A+ zone category where the Acc@3 superiority over the highest-performing baseline exceeds 4.3 percentage points, points to a concentration of efficiency benefits within planning zones that are subject to the highest cognitive loads on human engineers. This is because A+ zones control some of the most reliability-constrained urban-centric electric distribution infrastructure. The overall value of graph-structured knowledge representations to distribution network operational support has been independently validated within reinforcement learning-based architectures for real-

time fault management, where graph-structured relational representations have been shown to support more reliable decision-making under conditions of uncertainty [24]. This indicates that the representational efficiency benefits leveraged by the proposed framework are not limited to recommendation tasks but are instead a more general property of graph-structured representations that are applicable to a broader range of knowledge integration scenarios. The value of regulatory and physical constraint representations to enhancing the reliability of intelligent energy management has been documented within multi-energy microgrid management architectures, where explicit consideration of operational constraints within management architectures has been shown to produce measurably safer and more compliant management results compared to unconstrained architectures [25]. This places the CCR superiority of the proposed framework within a broader context as a property of the constrained management approach rather than an artifact of data or evaluation protocol.

4.3. Limitations and future research directions

Despite the evident performance benefits across all evaluation criteria, certain limitations of the current framework must be clearly acknowledged. The knowledge graph created from regulatory specification texts is static in nature at the time of inference; hence, regulatory changes or reclassification of zones require manual re-execution of the pipeline to incorporate these changes into the embedding and reasoning layers. This may impact operational efficiency in jurisdictions that experience regular changes to standard revisions. Additionally, the stratified case records across zones may impact performance for newly created zones or grid configurations without historical precedent; hence, performance may degrade to that of A+ zones under the current evaluation criteria, where case records were limited to being a primary factor for data-driven generalization. The extension of neural collaborative filtering techniques to knowledge graph embedding has been recognized as a significant research direction to further enhance the quality of knowledge-driven recommendations in knowledge-intensive domains [26]. Extending the framework to dynamically update knowledge graphs is a technically viable extension of the current framework to address the static knowledge limitation without impacting the transparency of constraint reasoning. The extension of knowledge graph construction using large language model techniques to potentially eliminate the need for manually annotated training datasets for the NER pipeline, as well as extending the multi-task architecture to accommodate additional planning objectives such as economic cost optimization and carbon emissions constraint satisfaction, represent research directions whose pursuit would extend the application of the framework beyond the wiring mode

and automation configuration recommendation problem that is the focus of the current study.

5. Conclusion

The present study suggests a framework of domain knowledge graph-informed joint recommendation for wiring mode selection and automation configuration across differentiated power supply zones. It fills the methodological gap of distribution network planning in that these two interdependent decision problems have traditionally been considered independently as optimization problems without a mechanism to jointly address these problems under a common regulatory constraint structure. The framework is comprised of a BERT-CRF pipeline for automatic construction of knowledge graphs from regulatory specification texts; TransR-initialized knowledge graph embedding with relation-aware graph attention networks for wiring mode recommendation conditioned upon power zones; a conditional random field decoder with SWRL constraint penalty for regulatory-compliant configuration recommendation; and a shared encoder architecture for multi-task joint learning of these two decision problems. Experimental evaluation of the framework using a multi-source dataset of 1,443 historical distribution network planning case records demonstrates that the suggested framework achieves Acc@3 of 0.817 and F1 of 0.757 for five differentiated power zones, surpassing the best baseline performance by 4.3 percent in Acc@3 for A+ zones and 6.9 percent for Configuration Completeness Rate. Ablation study of the suggested framework demonstrates that each of the three ICT components of the framework provides non-redundant performance gain to the overall framework performance. Practical application validation on 12 newly initiated planning tasks further confirms the operational applicability of the proposed framework, with an average expert evaluation score of 4.50 out of 5.00 and a 73.3% reduction in planning cycle time compared to conventional manual workflows, demonstrating that the framework can serve as an effective decision support tool in real-world distribution network planning practice.

Collectively, the experimental results demonstrate that the integration of structured knowledge graph representation, graph neural network aggregation, constraint-aware reasoning, and multi-task joint optimization achieves a quality of recommendation that is not possible with any of these components alone. The performance improvement of the proposed framework is most apparent when dealing with high complexity, data-scarce planning scenarios, as intelligent decision support systems have the most value in these scenarios. The proposed framework advances the state of the art by bridging the gap between isolated data-driven planning tools and integrated knowledge-driven recommendation systems that are able to simultaneously manage the complex relationships, constraints, and representational

coupling inherent with differentiated supply zone planning, thus offering a framework for the digital transformation of power distribution grid governance. While the proposed framework has several benefits, there are some limitations, including the static nature of the proposed knowledge graph during inference, as well as the need for annotated training corpora for the construction pipeline of named entity recognition. This has created a need for future extensions of the proposed framework, including dynamic knowledge graph updating, as well as large language model-based construction pipelines, which will further enhance the scalability of the proposed framework.

References

- [1] Matijašević T, Antić T, Capuder T. A systematic review of machine learning applications in the operation of smart distribution systems. *Energy reports*. 2022;8:12379-12407.
- [2] Pham TN, Shah R, Dao MN, Sultanova N, Islam S. Low and medium voltage distribution network planning with distributed energy resources: a survey. *Electrical Engineering*. 2025;107(2):1797-1828.
- [3] Mahdavi M, Alhelou HH, Hatziaargyriou ND, Jurado F. Reconfiguration of electric power distribution systems: Comprehensive review and classification. *IEEE Access*. 2021;9:118502-118527.
- [4] Ji S, Pan S, Cambria E, Marttinen P, Yu PS. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE transactions on neural networks and learning systems*. 2021;33(2):494-514.
- [5] Chen J, Lu G, Pan Z, Yu T, Ding M, Yang H. Research review of the knowledge graph and its application in power system dispatching and operation. *Frontiers in Energy Research*. 2022;10:896836.
- [6] Wang C, An J, Mu G. Power system network topology identification based on knowledge graph and graph neural network. *Frontiers in Energy Research*. 2021;8:613331.
- [7] Liao W, Bak-Jensen B, Pillai JR, Wang Y, Wang Y. A review of graph neural networks and their applications in power systems. *Journal of Modern Power Systems and Clean Energy*. 2021;10(2):345-360.
- [8] Lyu Z, Wu Y, Lai J, Yang M, Li C, Zhou W. Knowledge enhanced graph neural networks for explainable recommendation. *IEEE Transactions on Knowledge and Data Engineering*. 2022;35(5):4954-4968.
- [9] Gao M, Li JY, Chen CH, Li Y, Zhang J, Zhan ZH. Enhanced multi-task learning and knowledge graph-based recommender system. *IEEE Transactions on Knowledge and Data Engineering*. 2023;35(10):10281-10294.
- [10] Guo Z, Li F, Shen J, Wang X. Netcr: Knowledge-graph-based recommendation framework for manual network configuration. *IEEE Internet of Things Journal*. 2023;11(7):12941-12952.
- [11] Liu S, Guo Y, Tang W, Sun H, Huang W, Hou J. Varying condition SCOPF based on deep learning and knowledge graph. *IEEE Transactions on Power Systems*. 2022;38(4):3189-3200.
- [12] Pu T, Tan Y, Ma Z, et al. Graph computing based knowledge reasoning in electric power system considering knowledge graph sparsity. *CSEE Journal of Power and Energy Systems*. 2025;11:2083-2093.
- [13] Wang Y, Mo L, Wang W, Wei J, Yang J. Research and application of automatic mapping method of distribution

- network protection power supply based on knowledge graph and graph convolution network. *International Journal of Low-Carbon Technologies*. 2024;19:964-971.
- [14] Han H, Wang J, Wang X, Chen S. Construction and evolution of fault diagnosis knowledge graph in industrial process. *IEEE Transactions on Instrumentation and Measurement*. 2022;71:1-12.
 - [15] Meng F, Yang S, Wang J, Xia L, Liu H. Creating knowledge graph of electric power equipment faults based on BERT–BiLSTM–CRF model. *Journal of electrical engineering & technology*. 2022;17(4):2507-2516.
 - [16] Han Z, Wang J. Knowledge enhanced graph inference network based entity-relation extraction and knowledge graph construction for industrial domain. *Frontiers of Engineering Management*. 2024;11(1):143-158.
 - [17] Zhu X, Li Z, Wang X, Jiang X, Sun P, Wang X, et al. Multi-modal knowledge graph construction and application: A survey. *IEEE Transactions on Knowledge and Data Engineering*. 2022;36(2):715-735.
 - [18] Wu C, Liu S, Zeng Z, Chen M, Alhudhaif A, Tang X, et al. Knowledge graph-based multi-context-aware recommendation algorithm. *Information Sciences*. 2022;595:179-194.
 - [19] Li C, Wang B. A knowledge graph method towards power system fault diagnosis and classification. *Electronics*. 2023;12(23):4808.
 - [20] Dong Y, Liu Y, Dong Y, Wang Y, Chen M. Multi-knowledge enhanced graph convolution for learning resource recommendation. *Knowledge-Based Systems*. 2024;291:111521.
 - [21] Yan D, Shi Z, Wang X, Gao Y, Pu T, Wang J. Efficient and Stable Learning for Distribution Network Operation: A Model-Based Reinforcement Learning Approach. *CSEE Journal of Power and Energy Systems*. 2025;11(3):1080-1092.
 - [22] Wu Y, Liu F, Wan L, Wang Z. Intelligent fault diagnostic model for industrial equipment based on multimodal knowledge graph. *IEEE Sensors Journal*. 2023;23(21):26269-26278.
 - [23] Zeng L, Yao W, Hu Z, Shuai H, Li Z, Wen J, et al. Intelligent predetermination of generator tripping scheme: Knowledge fusion-based deep reinforcement learning framework. *CSEE Journal of Power and Energy Systems*. 2023;10(1):66-75.
 - [24] Jacob RA, Paul S, Chowdhury S, Gel YR, Zhang J. Real-time outage management in active distribution networks using reinforcement learning over graphs. *Nature Communications*. 2024;15(1):4766.
 - [25] Wang Y, Qiu D, Sun M, Strbac G, Gao Z. Secure energy management of multi-energy microgrid: A physical-informed safe reinforcement learning approach. *Applied Energy*. 2023;335:120759.
 - [26] Shokrzadeh Z, Feizi-Derakhshi MR, Balafar MA, Mohasefi JB. Knowledge graph-based recommendation system enhanced by neural collaborative filtering and knowledge graph embedding. *Ain Shams Engineering Journal*. 2024;15(1):102263.