ISSN: 2688-8645

Construction and Dynamic Update Technology of Knowledge Graph for Domain-specific Large Language Models in Electric Power Communication

Xiangcheng He *, Yang Yang [a], Xuedong Sun [b], Maoping Li [c], Rong Lin [d]

State Grid Yibin Electric Power Supply Company, Yibin, Sichuan Province, China

* Corresponding author: Xiangcheng He (Email: 18111150519@163.com),

[a] 13314422295@163.com, [b] 15196998148@163.com, [c] 253803509@qq.com,

[d] 18583116083@163.com

Abstract

Aiming at the problems in knowledge management of Electric Power Communication (EPC), such as insufficient use of unstructured information and lagging update of dynamic knowledge, this article studies the construction and dynamic update technology of knowledge map of EPC-specific Large Language Model (LLM). Firstly, the domain knowledge system is analyzed, and the core types of entities and relationships are defined. A knowledge extraction algorithm based on domain-enhanced prompt learning is proposed, which combines EPC terminology dictionary and prompt template library to improve the accuracy of entity and relationship recognition. An ontology alignment method integrating semantic similarity and rule reasoning is also designed to realize the unified representation of multi-source heterogeneous data. The incremental updating framework of time sequence awareness is constructed, and the inconsistency in the process of knowledge evolution is dealt with through conflict detection and confidence assessment mechanism. The results show that the F1 value of entity recognition is 92.3%, the F1 value of relation extraction is 89.7%, and the accuracy of ontology alignment is 94.6%. In the dynamic update scenario, it takes an average of 12.5 seconds to process 1000 pieces of data at a time, and the accuracy of conflict handling is 90.3%. The research shows that the proposed method can effectively construct a knowledge map adapted to EPC field and realize efficient dynamic update, which provides reliable knowledge support for intelligent operation and maintenance of EPC network.

Keywords

Electric Power Communication; Knowledge Map; LLM; Dynamic Update; Knowledge Extraction.

1. Introduction

With the acceleration of the construction of new Electric Power System (EPS), EPC network, as the "nerve center" of energy Internet, has grown exponentially in scale and complexity. From substation automation system to trans-regional transmission line monitoring, from distribution internet of things to dispatching cloud platform, massive equipment operation data, business interaction information and fault handling records constitute a multi-source heterogeneous knowledge ocean [1]. This knowledge not only includes static parameters such as circuit breaker action logic and optical cable transmission loss, but also involves dynamic information such as communication topology adjustment caused by load fluctuation and link failure caused by extreme weather [2]. Its efficient management and deep utilization have become the key to ensure the safe and stable operation of EPS.

At present, knowledge management in EPC field still faces multiple challenges [3]. On the one hand, it is difficult for traditional database and document management systems to deal with the hidden association in unstructured texts, such as the causal relationship between "sudden increase in error rate" and "aging of optical modules" in fault reports, which often depends on expert experience mining [4]. On the other hand, the dynamic evolution of knowledge intensifies the difficulty of management, and new knowledge is constantly produced in scenarios such as new equipment entering the network and communication protocol upgrading, while the existing manual updating mode is lagging behind and cannot meet the real-time decision-making needs [5]. In this context, knowledge map has become an important carrier to integrate knowledge in EPC field by virtue of its structural representation ability of entity relationship [6].

Traditional knowledge map construction relies on manual definition rules or small sample supervised learning, which is limited in fields such as EPC with dense technical terms and complex relationship types [7]. In recent years, the LLM provides a new way to solve this problem by virtue of its deep understanding of natural language and domain knowledge accumulated in mass text pre-training [8]. The dependence on labeled data can be significantly reduced by prompting engineering to guide the LLM to identify entities and extract relationships. Using its logical reasoning ability to deal with knowledge conflicts can improve the consistency of the atlas [9-10]. However, there are still shortcomings in the existing research on multi-focus general fields, aiming at the exclusive adaptation of EPC: how to accurately identify professional entities such as SDH multiplexing section and OTN optical layer, how to dynamically capture the time-series relationship such as protection switching and bandwidth adjustment, and how to realize automatic incremental updating of knowledge map to adapt to network topology changes, which have not been effectively solved [11].

This article focuses on the construction and dynamic update of knowledge map of EPC-specific LLM. Firstly, the domain knowledge system is analyzed, and the core types of entities and relationships are defined. Secondly, explore the method of knowledge extraction and ontology construction based on LLM to improve the domain adaptability of atlas; Furthermore, a dynamic updating mechanism is designed to realize the real-time integration of new knowledge and conflict resolution; Finally, the effectiveness of the proposed technology is verified by experiments. The research results are expected to provide knowledge support for intelligent operation and maintenance, fault diagnosis and planning decision of EPC network, and promote the deep collaboration between communication and power in new EPS.

2. Construction Technology of EPC Exclusive Knowledge Map

The construction of EPC exclusive knowledge map needs to be based on domain characteristics, from knowledge extraction, ontology design to multi-source integration to form a complete technical chain [12]. In knowledge extraction, the strategy of combining LLM with domain rules is adopted. Aiming at the core entities such as "optical transmission equipment" and "relay protection device", the model is guided to identify professional terms by constructing prompt words including equipment model and function description. For the relationship types such as "bearer service" and "communication protocol", the extraction rules are equationted in combination with EPC procedures [13]. For example, the relationship pairs such as "220kV substation-carrying-telecontrol signal" and "220kV substation-connection-SDH transmission line" are extracted from "220kV substation carries telecontrol signal through SDH transmission line".

ISSN: 2688-8645

Ontology construction is the basis of ensuring the consistency of the map, which needs to cover multi-dimensional knowledge such as equipment, business and fault. The core elements of the ontology are shown in Table 1. Among them, the hierarchical relationship of entity types is defined, and the entity attributes and relationship types are defined, which provides a clear framework for knowledge organization.

Table 1. Core Elements of EPC Knowledge Graph Ontology

Entity Type	Subtypes	Typical Attributes	Main Relation Types
Communication Equipment	Transmission, Access Equipment	Model, Transmission Rate, Manufacturer	Connect, Carry, Include
Business System	Motion Control, Dispatch System	Service Level, Data Volume	Depend, Access
Fault Event	Link Failure, Equipment Failure	Occurrence Time, Impact Scope	Associate, Cause
Communication Link	Cable, Microwave Link	Length, Attenuation Value	Connect, Belong To

Multi-source fusion technology is used to deal with the differences of heterogeneous data. Structured equipment accounts and configuration tables are transformed into triples through mapping rules [14]. Unstructured fault reports and maintenance records are aligned with structured data after entity relations are extracted by LLM. The semi-structured communication network topology map extracts the information of nodes and links to supplement the spatial relationship. In the process of integration, entity disambiguation algorithm is used to solve the problem of "the same thing with different names", such as associating "synchronous digital hierarchy" and "SDH" as the same entity, and finally forming a logically consistent EPC knowledge map.

3. Dynamic Updating Technology of Knowledge Map

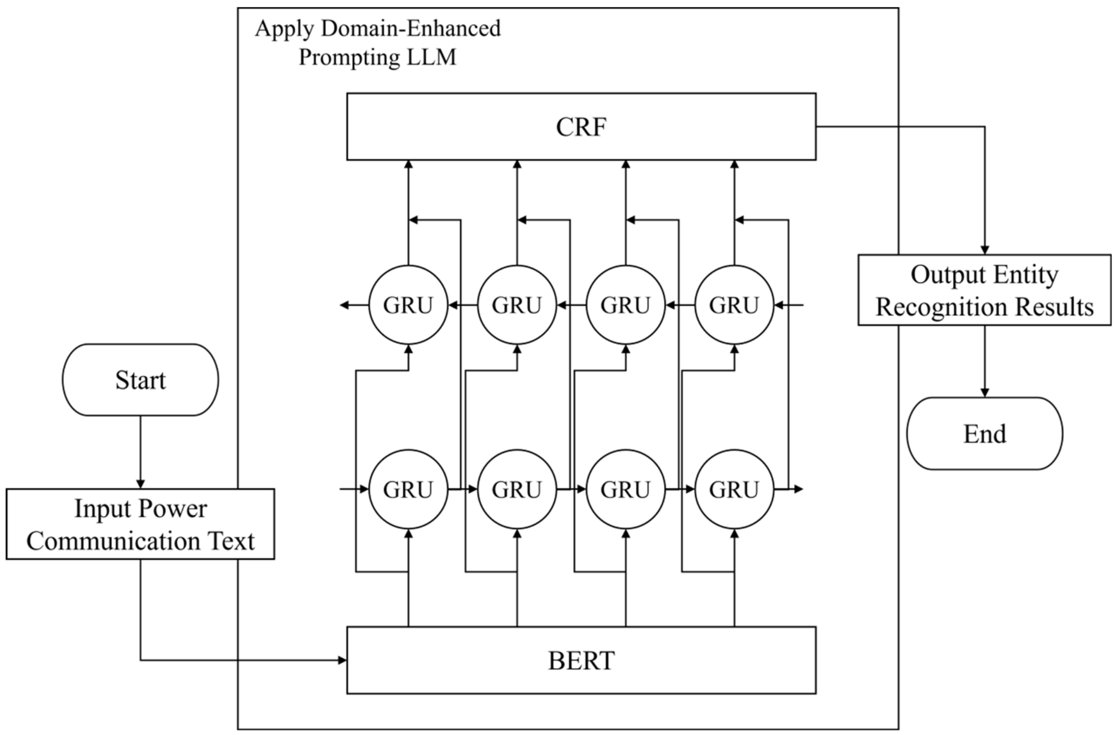

Figure 1. EPC equipment fault diagnosis information model

ISSN: 2688-8645

The real-time running characteristics of EPC network determine that the knowledge map needs to have the ability of dynamic updating to adapt to the continuous evolution of knowledge. This process needs to solve three core problems: knowledge change perception, incremental update implementation and conflict resolution, and form a closed-loop update mechanism. Normalized extraction can obtain a relatively complete tag sequence and improve the accuracy of entity recognition (as shown in Figure 1).

Knowledge change perception depends on real-time monitoring of multi-source data. Through the data acquisition module deployed in the communication network management system, the equipment status message, service configuration instruction and fault alarm information are captured in real time. Different event types correspond to different update requirements, as shown in Table 2. When the event of "new equipment entering the network" is monitored, it is needed to extract attribute information such as equipment model and subnet, and supplement the entity and association; The "protocol upgrade" event requires updating the protocol version attributes and related interaction rules of the corresponding business system.

Table 2. EPC Knowledge Update Triggers and Handling Methods

Trigger Type	Typical Scenario	Knowledge Update Content	Priority
New Equipment Onboarding	5G Base Station Connects to Power Comm Network	Add equipment entity & connections	High
Service Adjustment	Relay Protection Switches Load Paths	Update service-link associations	High
Fault Recovery	Fiber Restored After Interruption	Update fault event status	Medium
Performance Tuning	Adjust Transmission Device Power Threshold	Update device attribute values	Low
Protocol Upgrade	Dispatch Data Network Adopts New Protocol	Update protocol entities & rules	Medium

Incremental update implementation adopts hierarchical processing strategy. For new entities and relationships, the lightweight extraction model is directly written into the atlas to avoid the efficiency loss caused by full reconstruction; For the change of attribute value, such as the adjustment of equipment operating temperature threshold, the version control mechanism is used to record the historical state to ensure traceability. Aiming at the common temporal knowledge in EPC, such as the fluctuation law of link load with time, a timeline marking method is designed to make the map not only reflect the current state, but also trace back the knowledge snapshot at a specific moment.

Knowledge conflict resolution needs to combine domain rules and LLM reasoning. When the newly extracted knowledge is inconsistent with the existing information in the atlas, it is judged according to the pre-defined priority rules, such as the equipment nameplate information takes precedence over the log deduction result; When the rules can't be determined, LLM is called to make semantic analysis of the conflicting contents, and the confidence score is given by combining EPC principle, and the high confidence information is selected as the update basis. This mixed digestion method not only ensures the professional accuracy, but also improves the processing flexibility.

ISSN: 2688-8645

4. Algorithm Model and Experimental Verification

4.1. Algorithm Model

In this article, a hybrid algorithm framework integrating LLM is designed for the construction and dynamic updating of EPC exclusive knowledge map. It covers three core modules: knowledge extraction, ontology alignment and dynamic updating. In the knowledge extraction, an entity and relationship extraction algorithm based on domain enhanced prompt learning is proposed. Based on pre-training LLM, this algorithm improves performance by constructing EPC domain prompt template library. The template library contains 8 types of sentence patterns such as equipment description, and each type of sentence pattern is embedded with 30-50 typical examples extracted from industry standards and operation and maintenance manuals. The algorithm adopts a two-stage extraction strategy. In the first stage, the named entity identification submodel is used to identify core entities such as "optical transmission equipment" and "relay protection service", and the model output is constrained by introducing EPC terminology dictionary to calculate the entity boundary identification error:

$$E_b = \frac{1}{N}\sum_{i=1}^{N} | L_i^{pred} - L_i^{true} | \tag{1}$$

Among them, E_b is the average boundary error, N is the total number of entities, and L_i^{pred} and L_i^{true} are the predicted and real entity boundary character positions respectively. In the second stage, the relationship classification sub-model is used to predict the relationship category combined with the entity type information, and the domain relationship weight factor is added to the loss function:

$$L_{rel} = -\sum_{i=1}^{M} w_i \cdot y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \tag{2}$$

In the equation, M is the number of relationship categories, w_i is the weight factor of the i-type relationship, y_i is the label, and p_i is the prediction probability, so as to improve the recognition ability of low-frequency relationships such as "protection switching" and "protocol adaptation".

Ontology alignment module adopts a hybrid matching algorithm based on semantic similarity and rule reasoning. For entity alignment, the semantic vector similarity of entity names to be matched is calculated by BERT model:

$$S_{sem}(e_1, e_2) = \frac{\vec{v}_{e1} \cdot \vec{v}_{e2}}{||\vec{v}_{e1}|| \cdot ||\vec{v}_{e2}||} \tag{3}$$

Among them, $\vec{v}_{e1}$ and $\vec{v}_{e2}$ are the semantic vectors of entities e_1 and e_2, and then combined with attribute rules for secondary verification. For relationship alignment, the domain relationship mapping table is designed to map the relationships with different expressions in different data sources to the standard relationship types defined by ontology. The mapping rules are automatically generated through LLM for semantic parsing of relational contexts.

The dynamic update module proposes an incremental update algorithm based on time sequence awareness. Real-time monitoring of knowledge changes in new data by sliding time window, and the window size is dynamically adjusted according to the data update frequency:

$$W_t = \alpha \cdot W_{t-1} + (1 - \alpha) \cdot \frac{1}{f_t} \tag{4}$$

Where W_t is the current window size, W_{t-1} is the window size at the last moment, α is the smoothing coefficient (take 0.7), and f_t is the current data update frequency.

For the detected new knowledge, the algorithm first compares the conflict detection submodel with the existing knowledge of the atlas. The conflict detection adopts the combination of triple hash index and semantic reasoning. The index is used to quickly locate different attributes of the same entity, and the reasoning is used to identify hidden conflicts. In the conflict resolution stage, a confidence assessment mechanism is introduced, and the reliability of data source, knowledge freshness and the reasoning result of LLM are comprehensively considered to calculate the confidence:

$$C(k) = \beta \cdot C_s(k) + (1-\beta) \cdot C_t(k) \cdot C_{llm}(k) \quad (5)$$

Where $C(k)$ is the comprehensive confidence of knowledge k, $C_s(k)$ is the reliability of data source (0-1), $C_t(k)$ is the time decay factor, $C_{llm}(k)$ is the LLM reasoning confidence (0-1), and it is the weight coefficient (0.4), so as to automatically select the optimal updating scheme.

The algorithm also designs a knowledge attenuation mechanism to mark or delete the dynamic knowledge beyond the preset time limit to ensure the timeliness and simplicity of the atlas.

4.2. Experimental Verification

The experiment is based on the actual operation and maintenance data of the communication network of a provincial power company. The data set includes equipment account (5000+pieces), fault handling report (3000+pieces), communication network topology map (200+pieces) and real-time alarm data (100,000+pieces). The data covers the optical transmission network, dispatching data network and other core systems of 220kV and above substations. The experimental environment is Intel Xeon Gold 6330 processor and 128GB memory server. The software adopts PyTorch framework, and LLM adopts LLaMA-7B model which has been fine-tuned by EPC corpus. The experiment evaluates the knowledge map from three dimensions: construction accuracy, dynamic updating efficiency and conflict resolution effect.

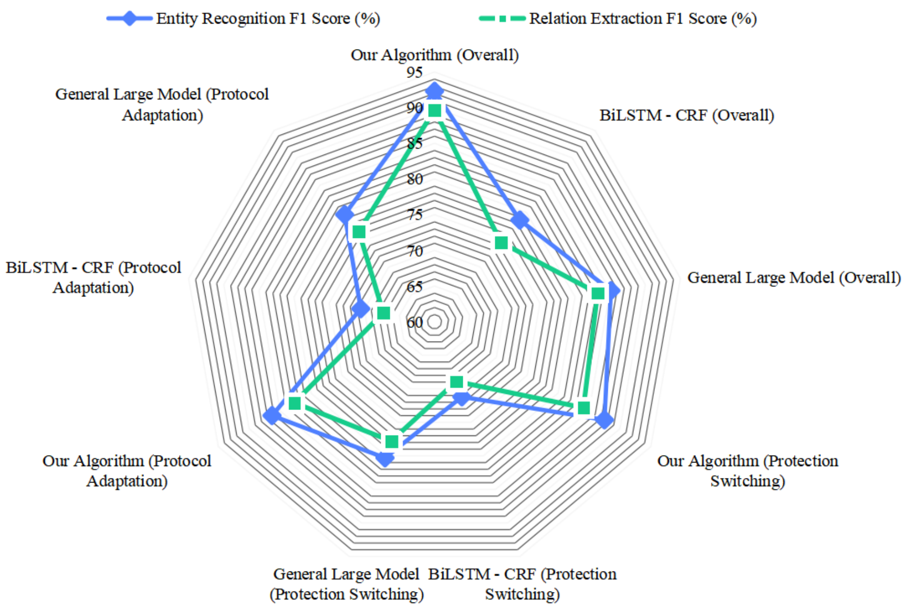

Figure 2. Knowledge extraction performance of different models

In the aspect of knowledge extraction performance, this algorithm is compared with the traditional BiLSTM-CRF model and general LLM (without domain fine-tuning). The assessment indexes include entity recognition accuracy, recall and F1 value, and the relationship extraction adopts the same index. The experimental results are shown in Figure 2. The F1 value of the entity recognition of this algorithm reaches 92.3%, which is 13.7% higher than that of BiLSTM-CRF (78.6%) and 7.2% higher than that of general LLM(85.1%). The F1 value of relation extraction is 89.7%, which is 15.2% and 6.5% higher than that of contrast model. Especially in the low-frequency professional relationships such as "protection switching" and "protocol adaptation", the advantages of this algorithm are more obvious, and the F1 value reaches 87.5% and 86.3% respectively, which is due to the joint action of domain prompt template and terminology dictionary.

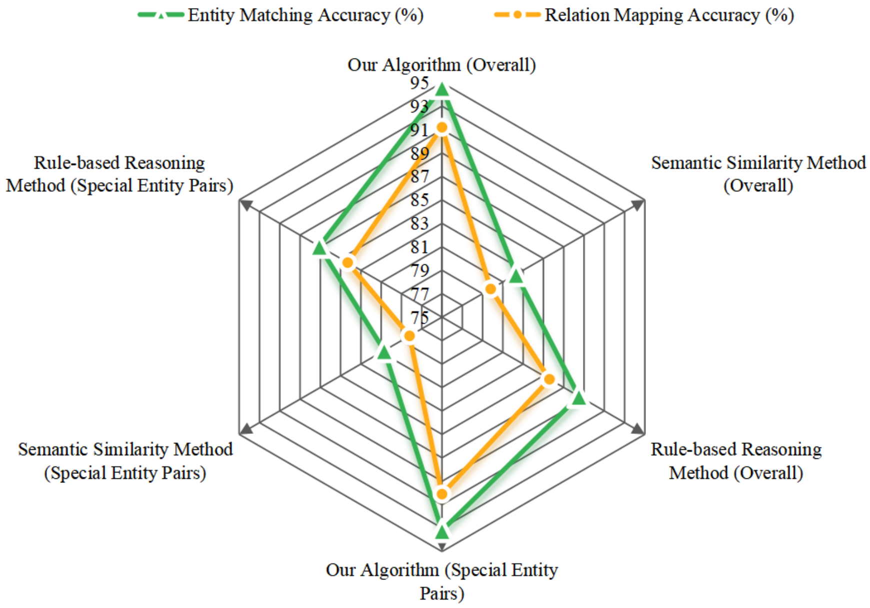

Figure 3. Accuracy of ontology alignment using different methods

The effect of ontology alignment is evaluated by entity matching accuracy and relationship mapping accuracy. In the experiment, 1000 pairs of entity pairs and relationship pairs from different data sources were manually labeled as a test set. The accuracy of entity matching and relationship mapping of this algorithm is 94.6% and 91.2%, as shown in Figure 3. In the comparative experiment, the accuracy of entity matching by only using semantic similarity is 82.3%, and that by using rule-based reasoning is 88.5%, which shows that the hybrid algorithm in this article can effectively combine semantic information with domain rules and reduce the influence of the phenomenon of "synonyms with different names" and "synonyms with different names".

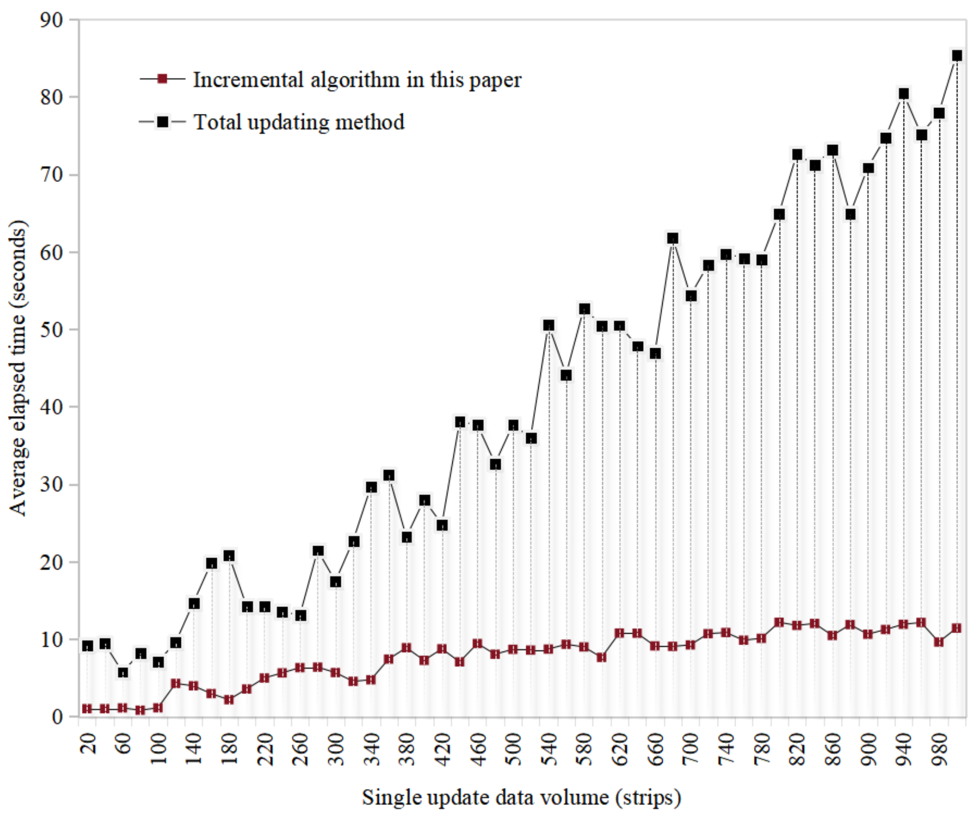

Figure 4. Time-consuming comparison of different updating methods

The dynamic update efficiency test simulates 1000 knowledge update scenarios and records the average update time under different data volumes. Results As shown in Figure 4, when the data volume of a single update is increased from 10 to 1000, the average time consumption of the algorithm in this article is increased from 0.8 seconds to 12.5 seconds, while the time consumption of the full update method is increased from 5.2 seconds to 89.3 seconds. In the aspect of update delay, the processing delay of emergency alarm information in this algorithm is controlled within 1.5 seconds, which meets the real-time requirements of EPC.

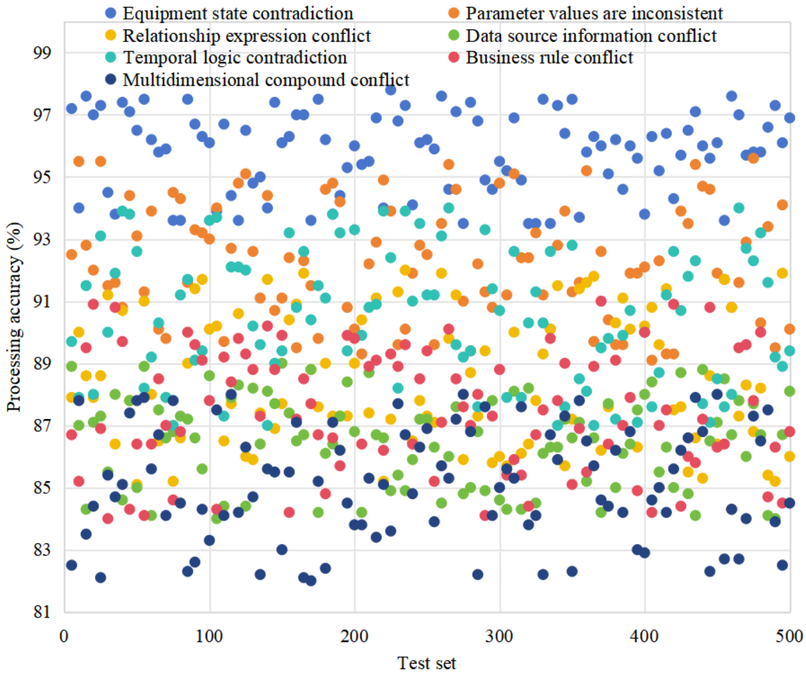

Figure 5. Relationship between conflict handling accuracy and conflict types

The effect of conflict resolution is evaluated by conflict handling accuracy and user satisfaction. The experiment builds a test set containing 500 conflict cases, and the accuracy of conflict handling of this algorithm is about 90.3%, as shown in Figure 5. Among them, for high-priority conflicts, the processing accuracy is 96.7%. 100 conflict cases processed by the algorithm are randomly selected for expert review, and the results show that the conflict resolution results of the algorithm conform to the professional judgment logic in EPC field.

The comprehensive experimental results show that the algorithm model proposed in this article is excellent in the construction accuracy, dynamic updating efficiency and conflict resolution ability of EPC knowledge map, and can meet the actual needs of EPC network knowledge management.

5. Conclusion

In this article, the construction and dynamic update of EPC-specific LLM knowledge map are studied, and a complete technical scheme is formed and its effectiveness is verified by experiments. In the construction of knowledge map, the proposed domain-enhanced cue learning algorithm significantly improves the extraction accuracy of professional entities and relationships, with the F1 value of entity recognition reaching 92.3% and the F1 value of relationship extraction 89.7%, which has obvious advantages over traditional methods and general models. The hybrid matching strategy adopted in ontology alignment improves the accuracy of entity matching to 94.6% and the accuracy of relationship mapping to 91.2%, effectively eliminating the semantic differences in multi-source data.

The results of dynamic updating technology show that the incremental updating framework of time sequence awareness can deal with knowledge evolution efficiently. It takes only 12.5 seconds in the updating scene of 1000 pieces of data, which is nearly 7 times more efficient than the full updating method, and the delay of emergency information processing is controlled within 1.5 seconds, which meets the real-time requirements of EPC. The conflict resolution mechanism passed the multi-dimensional confidence assessment, and the processing accuracy reached 90.3%, and the processing accuracy of high-priority conflicts was as high as 96.7%, which ensured the consistency of the map.

The research results of this article provide a feasible path for the structured management and dynamic maintenance of knowledge in EPC field, and the constructed knowledge map can support practical applications such as equipment condition monitoring and fault diagnosis. Subsequent research can further optimize LLM's domain fine-tuning strategy, improve the robustness of knowledge extraction in extreme scenarios, and explore the integrated application of knowledge map and digital twinning technology.

Acknowledgments

Sponsored by Science and Technology Project of State Grid Sichuan Electric Power.

Company Project: Research on Key Technologies of Knowledge Base for Large Language Models in Backbone Power Communication Networks(ERP Code: 521911250004).

References

[1] Liu Z, Yang C, Jiang W. Power Allocation Technology for DC Microgrid Energy Storage Systems Based on Consensus Algorithm[J]. Automation of Electric Power Systems, 2020, 44(07): 61-69. DOI: 10. 7500/AEPS20190515005.

[2] Tan G, Chen Y, Peng Y Z. A Power Grid Customer Service Q&A System Integrating Domain Feature Knowledge Graph[J]. Computer Engineering and Applications, 2020, 56(03): 232-239. DOI:10. 3778/ j.issn.1002-8331.1907-0385.

[3] Li T, Sun Z, Wang Y. Undersea Simultaneous Wireless Power and Data Transfer System with Extended Communication Distance and High Rate[J]. IEEE Transactions on Power Electronics, 2023, 39(3): 2917-2921. DOI:10.1109/tpel.2023.3337801.

[4] Zhou S W, Huang H, Li R Q. Knowledge Graph and Correlation Analysis of Urban Interdependent Infrastructure Research Based on Literature[J]. Journal of Catastrophology, 2023, 38(2): 150-154.

[5] Zhu P Y, Cai X Z, Xu S Y. Research on Defect Diagnosis for Edge Computing in Electric Power Communication Networks[J]. Application of Electronic Technique, 2021, 47(04): 30-35+45. DOI:10.16157/j.issn.0258-7998.201347.

[6] Zhang K, Wu X G, Lyu Y T. An Automated Generation Method for Typical Accident Plans Based on Knowledge Graph and Progressive Learning[J]. Electrical Automation, 2024, 46(5): 85-87.

[7] Xu T, Zou Z H, Hu R Y. A Method for Constructing a Knowledge Graph of Hydropower Station Dispatching and Operation Based on BiGRU+CRF Deep Learning Model[J]. Journal of Electric Power Science and Technology, 2025, 40(1): 180-189.

[8] Li Y C, Luo H, Wang X Y. Construction of an APT Attack Detection Model Based on Provenance Graph and Attention Mechanism[J]. Journal on Communications, 2024, 45(3): 117-130. DOI:10.11959/ j.issn. 1000-436x.2024039.

[9] Meng M, Ma C N, Xue W C. A Hierarchical Distributed Energy Coordination Method for Integrated Energy Systems Based on Multi-Agent[J]. Modern Electric Power, 2021, 38(02): 129-139. DOI: 10. 19725/j.cnki.1007-2322.2020.0256.

[10] Zhang Y, Xu H F, Duan H S. Design of a Java-Based Visualization Communication System for Power Dispatching Information[J]. Electronic Design Engineering, 2021, 29(08): 43-47. DOI:10. 14022/j. issn1674-6236.2021.08.010.

[11] Jiang W C, Wang T, Zheng G N. A Critical Link Identification Algorithm for Power Communication Networks Based on Primary-Backup Routing[J]. Automation of Electric Power Systems, 2024, 48(10): 203-215.

[12] Li B, Jia B C, Sun Y. A Cross-Layer Protection Algorithm for Power Communication Networks Based on Network Coding[J]. Automation of Electric Power Systems, 2020, 44(03): 132-138. DOI:CNKI: SUN: DLXT.0.2020-03-016.

[13] Xue S P, Gao D Q, Zhao Z Y. A Routing Algorithm for Power Communication Networks Based on Differentiated Service Transmission Requirements[J]. Electric Power, 2024, 57(11): 183-190.

[14] Sendin A, Losada-Sanisidro P, García-Rodríguez J. Broadband Over Power Line Communication Prototype Development for Next Generation Smart Meters: Validation in Access Electric Power Distribution Networks[J]. IEEE Access, 2024, 12: 30191-30208. DOI:10.1109/ ACCESS. 2024. 3367 987.