An Approach of Top-DOWN Electric Generation Knowledge Graph Construction

Hao Liang¹,²,*, Xiaojun Peng¹,², Ningning Zhao¹, Liguo Duan¹, Pengfei Yu¹ and Da Fu¹

¹Beijing Jingnen Energy Technology Research Co., LTD, Beijing, China
²The Engineering Research Center of Coal Industry Mine Robot, Beijing, China

*Corresponding author email: lianghao@powerbeijing.com

Abstract. Knowledge Graph technology is a key research area of artificial intelligence technology. It has an obviously advantage in semantic processing and content interconnection, and widely used to provide intelligent information services in the open domain such as Web intelligent, NLP, user profile, etc. With the exploration and attempt of Knowledge Graph in vertical fields, it is confirmed that it can accelerated the intelligence application in all the electric generation life cycle, such as operation optimization, technical overhaul, intelligent marketing, material procurement, equipment maintenance. A Top-Down construction method of electric generation Knowledge Graph is proposed under the Kraftwerk-Kennzeichen system (KKS) to integrate electric generation knowledge and the knowledge management formation.

Keywords: Knowledge Graph; Top-Down; Electric generation; knowledge management; Kraftwerk-Kennzeichensystem (KKS).

1. Introduction

Knowledge graph is not a new concept, it originated from semantic network and was proposed by Richard H in 1956, as a kind of semantic network that represents the semantic relation between concepts. From 1980s to the 1990s, semantic network mainly focused on modeling the relationship between concepts and proposed the logic of belonging to logic and description logic. In 2006, Berners-Lee Tim enriched the concept of semantic Web and called for the promotion and improvement of ontology models to formalize the semantics in data. The formalized models of RDF Schema and Web Ontology Language (OWL) were generated based on the purposes [1]. In 2012, knowledge graph was formally proposed by Google [2], with original intention of improving the capability of search engine, enhancing search quality and experience. At present, with the continuous development of intelligent information service application, knowledge graph has been mature application to intelligent search, intelligent Q & A, personalized recommendation and other fields. With the integration of digital twin, big data, industrial Iot, intelligent robot and other electric generation automation technologies, the intelligent level of power enterprises will be further improved, and the value of enterprise data will be further mined. At the same time, with the continuous enrichment of enterprise knowledge management, precondition of constructing electric knowledge graph is qualified.

2. The Framework of Knowledge Graph

The framework of typical vertical knowledge graph is mainly composed of logical structure and architecture.
2.1. Logical Structure
Logically, the knowledge graph is composed of pattern and data layer. The data layer is mainly composed of facts which are on behalf of knowledge. Triple such as <entity, relationship, entity> can be used to describe facts, <entity, attribute, attribute value> can be used to describe instances. Graph database are used to storage facts and instances [3]. The pattern layer is on the top of the data layer, and it is mainly used to regulate a series of fact expressions in the data layer. Ontology is a concept template of structured knowledge. The knowledge base formed by ontology is hierarchy and less redundancy.

2.2. Architecture
In Figure 1, we can see the architecture of the knowledge graph. The construction process of knowledge graph needs to be updated and iterated continuously with expert cognitive ability. At present, the effective method of knowledge graph construction are top-down and bottom-up. Before adding entities to the knowledge base, top-down begins with definition of domain ontology and data schema. This approach takes advantage of some existing structured knowledge bases, such as the freebase project. Bottom-up extracts entities from corpus such as design and equipment documents, adds to the knowledge base, and iteratively form the top-level ontology pattern. At present, most knowledge graph is constructed bottom-up, the typical is Google's Knowledge Vault [4-6].

![Knowledge Graph Architecture](image)

**Figure 1.** Technical architecture of electric generation Knowledge graph.

3. The construction of Knowledge Graph
The iterative process of ontology construction mainly includes information extraction, knowledge processing and fusion. Information extraction extracts concepts, attributes, and entities relationships from various types of data sources, on which knowledge expression is formed. Knowledge fusion deletes contradictions and ambiguities, especially the situation such as a concept corresponds to a number of entities, etc. During the knowledge processing, the qualified knowledge can be added to the preliminary knowledge base after quality assessment (sometimes by manual) to ensure the quality of the knowledge base. By adding new data, the completeness of knowledge is gradually improved, and the ability of reasoning is constantly improved [7-8].

3.1. Information Extraction
The major target of information extraction is to automatically extract information from heterogeneous data sources (semi-structured and unstructured). This kind of information represents as the candidate knowledge, composed of entity, relationship and entity attributes. The candidate knowledge can be used
to eliminate the ambiguity such as entity, relationship and entity attributes. Finally, a preliminary knowledge base is formed, and enriched and expanded by excavating. The key technologies are entity, relationship and attribute extraction [9]. The management process of power plant is generally reflected in work order, operation ticket, work ticket, material application, material purchase, material inventory, material delivery equipment specification, maintenance project, work breakdown structure, Enterprise resource plan (ERP) and so on. The information extraction of electric generation knowledge graph need to deal with a lot of semi-structured data. Through observation, it can be found that the information extraction and entity extraction of power knowledge graph construction process can be greatly simplified on the basis of KKS. Different from the bottom-up approach for open domain knowledge graph construction, the electric generation knowledge graph can construct by top-down, and the process of information extraction, entity extraction and even partial relation extraction can be highly integrated.

3.2. Entity Extraction
Entity extraction is also called entity recognition or entity learning to identify different entities form the Corpus. As entity is the most basic element in knowledge graph, the integrity, precision and recall of extraction is directly affect the quality. The major entity extraction method are statistical machine learning, rule or dictionary based and open domain clustering. [10-11]

In an electric power plant, the main maintenance objects are all kinds of process equipment, such as boiler, gas turbine, steam turbine, generator, fan, water pump, cooling equipment and electrical system, etc. The Kraftwerk-Kennzeichen system (KKS) is used to identify the system, equipment and functional element. KKS is classified step by step from unit, system, equipment, component, etc. The classification cover depth to the level of important parts in order of hierarchy, shown in Figure 2.

| Grading sequence number | Level 0 | Level 1 | Level 2 | Level 3 |
|-------------------------|---------|---------|---------|---------|
| Generation process correlation | Plant | System | Equipment | Component |
| Installation Location | Plant | Installation unit | Installation space |
| Physical Space Location | Plant | Building | Floor & Room |

**Figure 2.** Hierarchy of KKS.

The step of KKS coding is:
- Identify the main system functions, function sub-items, and their codes. The codes are gradually supplemented and refined, and then the codes of the main equipment of the system are determined. Extends to systems and equipment based on the encoding of the primary device. The component level coding is determined according to the equipment level coding.
- The system of the unit is divided into several categories, such as building system, public system, boiler, turbine, chemical water, fuel, electrical system, etc., and then further subdivided.
- Loop the above steps to make sure KKS including all generation process management objects.
- Sorting and checking the hierarchical and description of KKS.

3.3. Relation Extraction
Relation extraction is the main method to form the net of entities by associating. Typically, KKS adopts a 4-level hierarchy of generation process, formed by the plant, system, equipment, and component. Electricity generation is a typical manufacturing process, components comprise equipment, functionally similar equipment constitute a system, and different systems constitute a complete set of the technical unit. The logic relationship of KKS is very intuitive and accurate described method, therefore, the relation extraction of power generation knowledge graph between, can make full use of KKS code hierarchy architecture. The main relation type is “Part of” and “Is A” [12].
3.4. Attribute Extraction
The aim of attribute extraction is to collect detailed attributes from different information sources. Generally, the attribute of an entity is like a relation between entities and attribute. Attribute extraction can be transformed into entity relationship extraction [13]. Different from the Open Domain, the attribute type and attribute value interval of the entity, such as equipment, are determined by manufactory parameters, and the management standards of the electricity generation process stipulate the threshold value of most equipment entity attribute values, while the relationship between the attribute values of equipment entity in the same generation process is relatively fixed.

3.5. The Preliminary Construction
The proposed method is carried out in a 780MW turbine power station, the number of power plant KKS is 55,116, composed of 640 independent process system, profile section of KKS is shown in Table 1. The number of concept is 15,566, with different relation 33,652. The number of instances is enormous, all the instances are stored and used for equipment monitoring, prediction, overhaul and replacement. The more valuable use of electric knowledge graph is operation optimization, economical operation instruction.

| System      | Equipment Parts | System      | Equipment Parts | System      | Equipment Parts |
|-------------|-----------------|-------------|-----------------|-------------|-----------------|
| Turbine     | Stator          | Boiler      | Furnace         | Generator   | Stator          |
|             | cylinders       |             | water wall      |             | stator core     |
|             | baffles         |             | superheater     |             | stator coil     |
|             | nozzles         |             | economizer      |             | frame           |
|             | seals           |             | Flue            |             | end cover       |
| fasteners   | Soda system     |             | duct            |             |                 |
| bearings    | Heating surface |             | Rotor           |             |                 |
| Rotor       | Drum            |             | rotor body      |             |                 |
|             | spindles        |             | guard ring      |             |                 |
| impeller    | Pipe            |             | core ring       |             |                 |
| drum        | Furnace wall    |             | rotor coil      |             |                 |
| blades      | Frame           |             | slip ring       |             |                 |
|             |                 |             |                 |             |                 |

4. Conclusion
During the construction, there are different method for the entity linking, entity disambiguation, entity matching, and entity matching and entity synonyms. Under the framework of KKS and description of process system, the entity and attribute extraction is easier than open domain. The preliminary constructed knowledge graph is already used for fault expert system and Q&A. At the same time, combined with the professional correction of technicians, the knowledge description is becoming more and more accurate. Knowledge fusion, knowledge processing is still the future research work, we also need to realize implementation method of reuse and migration the existing knowledge graph. The size of instances in knowledge graph is nearly 40TB per day, we need to research and design the data cleaning, governance method to compress the data.
Acknowledgments

The research of electric generation knowledge graph is a major part of Smart Power Station Construction plan, which is designed and implemented by Beijing Jingnen Energy Technology Research Co., LTD. Relevant research work is supported by the Engineering Research Center of Coal Industry Mine Robot.

References

[1] Berners- Lee T, Hendler J, Lassila O. The semantic Web [J]. Scientific American Magazine, 2008, 23(1): 1-4.
[2] Jianqiang Guo, Shu Bin Liu, Xian Liu. Construction of visual cognitive computation model for sports psychology based on knowledge atlas, Cognitive Systems Research, 2018
[3] Hongyu Wang, Zhao hong Yang, Yu wei Zhou. A review of military knowledge models, IOPConference Series: Materials Science and Engineering, 2020
[4] AMIT S. Introducing the knowledge graph[R]. America: Official Blog of Google, 2012.
[5] Li rong Qiu, Hui li Zhang. Review of Development and Construction of Uyghur Knowledge Graph, 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), 2017
[6] You D, Pin J, Liiqiu Jin, Anming Zou, Jiaoyun Yang, Hong Xie, Ning An. Chapter 3 A Knowledge Graph for Eldercare: Constructing a Domain Entity Graph with Guidelines, Springer Science and Business Media LLC, 2018
[8] Zenglin Xu, Yongpan Sheng, Lirong He, Yafang Wang. Review on Knowledge Graph Techniques, Journal of University of Electronic and Technology of China, 001-0548.2016.04.012
[10] Lin Y, Liu Z, Sun M, et al. Learning entity and relation embedding for knowledge graph completion[C]//Proc of AAAI. Menlo Park, CA: AAAI, 2015.
[11] Xiao H, Huang M, Hao Y, et al. TransG: a generative mixture model for knowledge graph embedding [J]. ArxivPreprint ArXiv, 2015, 1509: 05488.
[12] Geng Xia, ZHANG Ji-jun, LI Wei-yu. Knowledge graph construction techniques [J]. Computer Science, 2014, 41(7):148-152.
[13] ConceptNet[EB/OL]. [2016]. http://www.oschina.net/p/conceptnet.
[14] Constructing a vertical knowledge graph for non-traditional machining industry, 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), 2018