Journal on Artificial Intelligence
DOI:10.32604/jai.2021.025175
Article

Construction and Application of Knowledge Graph for Quality and Safety Supervision of Transportation Engineering

Sheng Huang and Chuanle Liu*

Quality and Safety Supervision Bureau of Hunan Transportation Construction, Changsha, 410116, China
*Corresponding Author: Sheng Huang. Email: zjjhuang@126.com
Received: 02 December 2021; Accepted: 22 January 2022

Abstract: Knowledge graph technology play a more and more important role in various fields of industry and academia. This paper firstly introduces the general framework of the knowledge graph construction, which includes three stages: information extraction, knowledge fusion and knowledge processing. In order to improve the efficiency of quality and safety supervision of transportation engineering construction, this paper constructs a knowledge graph by acquiring multi-sources heterogeneous data from supervision of transportation engineering quality and safety. It employs a bottom-up construction strategy and some natural language processing methods to solve the problems of the knowledge extraction for transportation engineering construction. We use the entity relation extraction method to extract the entity triples from the multi-sources heterogeneous data, and then employ knowledge inference to complete the edges in the constructed knowledge graph, finally perform quality evaluation to add the valid triples to the knowledge graph for updating. Subgraph matching technology is also exploited to retrieve the constructed knowledge graph for efficiently acquiring the useful knowledge about the quality and safety of transportation engineering projects. The results show that the constructed knowledge graph provides a practical and valuable tool for the quality and safety supervision of transportation engineering construction.

Keywords: Knowledge graph; transportation engineering quality and safety supervision; information extraction

1 Introduction

With the rapid development of the modern construction industry, the information management of engineering quality supervision is playing an increasingly important role in quality and safety supervision. In China, the government required definitely to improve mechanism of the project quality supervision and innovate the supervision method for effectively supervising and managing the quality and safety of all projects. Under this background, various information systems for engineering management have been established in the various walks of life, which can generate a large amount of data. To mine deep and effective information from these massive multi-sources data from different information systems can provide highly meaningful and valuable information services for administrative departments and various enterprises, thus improving the development of engineering quality and safety supervision. However, there are significant differences in information systems of engineering quality and safety supervision in various regions, as the construction mode, content, and investment funds of different engineering projects differ in most situations. Moreover, the operation flow and information flow of engineering quality and safety supervision have not yet reached deep integration and organic aggregate, which has greatly affected the effectiveness of the construction informatization of engineering quality and safety in a certain extent. The quality standardization management system has not yet been established, which also would restrict the development of engineering construction. Therefore, there is an urgent need to use artificial intelligence and information technologies to promote the management and supervision capability of engineering quality and safety.

Currently, knowledge graph [1], which plays a vital role of artificial intelligence, has been already widely employed in various intelligent applications and fields, such as economy, industry, education, etc. [2,3], as knowledge graph has powerful capability of information extraction, knowledge representation and inference. Therefore, this paper attempts to explore various modern artificial intelligence technologies, such as knowledge graph, natural language processing, etc., to improve the information management capability of the engineering projects and guarantee the safety during the construction process. The information management based on knowledge graph of engineering construction quality and safety can efficiently collect, store, process a large amount of information, and display it intuitively, so as it can significantly improve the quality and efficiency of information management; meanwhile, it can accurately grasp and transmit information in a timely and rapid manner, and realize effective communication and timely management. It also can promote the capability and efficiency of prediction, decision-making, and planning in the process of supervising the engineering quality and safety. Motivated by the above analysis, this paper takes into account of the transportation engineering projects and multi-sources data to construct knowledge graph, and applies the constructed knowledge graph to design a system for querying knowledge of the quality and safety supervision of transportation engineering.

The rest of this paper is structured as follow. In Section 2, the framework and general process of the knowledge graph is introduced. Section 3 gives the details of constructing a knowledge graph for the quality and safety supervision of transportation engineering projects. And then Section 4 describes the example results of query the constructed knowledge graph. Finally, we conclude this work in the Section 5.

2 Construction of the General Knowledge Graph

2.1 The General Framework of the Knowledge Graph

The knowledge graph can be seen as a huge graph, which was firstly and originally presented by Google Knowledge Graph project in 2012. In a knowledge graph, nodes represent entities or concepts, while edges consist of attributes or relations. Knowledge graph has established an entity-centered knowledge system to show the relationships among massive data. At present, it has been widely used in the finance, medical industry, and so on. The development of knowledge graph has dramatically promoted the digitization process in e-commerce and social platforms [4]. Big data processing technology, knowledge search technology, and information visualization technology have also been fully and effectively utilized, as so to make knowledge graph achieve remarkable results.

The construction of knowledge graph mainly refers two ways: top-down and bottom-up [5]. In the early stage of the development of knowledge graph technology, most methods use the top-down way to extract ontology and pattern information from high-quality data and add them to the knowledge database. The high-quality data are always structed data, which are derived from encyclopedias website, such as Wikipedia. At present, the main researches are adopted the bottom-up way to construct the knowledge graph. The bottom-up construction method employs some technical means to extract information mode or pattern from the publicly collected data, and with the help of manual check, then select the new pattern with high confidence to enrich the knowledge base.

The goal of constructing a knowledge graph is acquiring knowledges from different data, including structured data, semi-structured data and unstructured data. The process of knowledge acquisition mainly contains stages: information extraction, knowledge fusion, and knowledge processing. Fig. 1 shows the general framework of constructing a knowledge graph in a bottom-up way. Seen from the Fig. 1, it can be found that data acquisition is the first step of constructing a knowledge graph. We can collect data from public, or semi-public or private channels, which are structured, semi-structured or unstructured. Data from different sources and structures have different preprocessing methods. And then the collected data are processed by the methods of information extraction, knowledge fusion and knowledge processing to learn knowledge and update the knowledge graph. Next, we will introduce the process of knowledge acquisition in details.

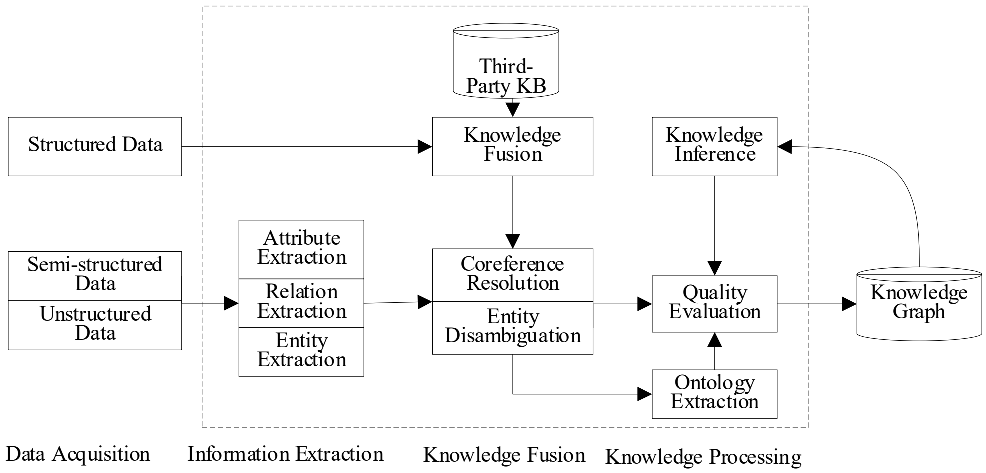

Figure 1: Framework of the knowledge graph construction [5]

2.2 Information Extraction

Information extraction is the key of learning knowledge from semi-structured and unstructured data. It extracts entity concepts, attributes, and relationships between entities from various data sources to form ontology-based knowledge representation. The key technologies involved in information extraction are entity extraction or name entity recognition, relation extraction, and attribute extraction. Entity extraction automatically identifies named entities from text data. The quality of entity extraction greatly influences subsequent relation extraction and is the most basic and critical part of information extraction.

Early entity extraction methods mainly used the rule-based way combining heuristic algorithms and manual rule writing [6]. This kind of methods has obvious limitations of time-consuming and poor scalability, which is challenging to adapt to data changes. Then, people began to try to use statistical machine learning to solve the problem of entity extraction. Liu et al. [7] used the k-Nearest Neighbor and Conditional Random Fields model to realize the recognition of entities contained in Twitter text data. With the development of deep learning, Misawa et al. [8] proposed combining Conditional Random Fields model and Long Short Term Memory Network to automatically learn entity recognition features, avoiding the manual features required by traditional machine learning and achieving better entity recognition results.

As the results of entity recognition, a series of discrete named entities are extracted from semi-structured and unstructured data. In order to obtain semantic information, it is also necessary to extract the indirect correlation from the relevant corpus and entities (concepts) through the relation to form a web of knowledge, which will be represented in the form of graph. The early researches of relation extraction mainly used grammatical and semantic rules and pattern matching methods to identify the relations between entities. These methods are challenging to adapt to rich language expression styles and is difficult to expand to different fields. Subsequently, statistical machine learning is introduced to model the relation between entities. For example, Kambhatla et al. [9] used syntactic and semantic features to model entity relations and successfully extracted entity relations through the maximum entropy method. Zhao et al. [10] classified these syntactic and semantic features according to entity attributes, binary attributes, and dependency paths for relation extraction. Zhou et al. [11] systematically studied how to combine various features, including basic phrase blocks (Chunk), to explore the contribution of various linguistic features to the performance of relationship extraction, especially the influence of semantic information such as WordNet and Name List. Currently, deep learning is also contributed to relation extraction. Zeng et al. [12] employed convolutional neural networks for relation classification to automatically learn multiple levels of features. Moreover, Zeng et al. [13] proposed a Piecewise Pooling Convolutional Neural Network and successfully applied it to extract weakly supervised entity relations.

In order to enhance the construction of a knowledge graph, attribute extraction is utilized to find the attributes for the instances of a given semantic class [14,15], for example, “capital city” is an attribute of any “Country”. While the relation extraction is extracting the distinctive relationship between entities. A semantic network may compose of millions of entities, so a knowledge graph construction in some cases always focuses on the semantic attributes among a semantic class and its instances.

By entity extraction, relation extraction, and attribute extraction, the useful and structured information are extracted from unstructured and semi-structured data. However, the extracted data generally are redundant and have errors. Moreover, the relationships among different data are shallow, lacking hierarchy and logic. Thus, it is necessary to perform knowledge fusion and processing to clean up and integrate the extracted information.

2.3 Knowledge Fusion

In order to integrate the knowledges from different sources, such as the third-party knowledge graph, knowledge learned from structured, semi-structured and unstructured data, knowledge fusion technologies are utilized, which mainly includes: coreference resolution [16] and entity disambiguation [17]. Coreference resolution can identify all mentions that involve the same entity in the real world, so as some the massive entities derived from entity extraction can be clustered. Entity disambiguation is to disambiguate the ambiguity among the entities, for example, different people share the same name entity. A general way of solving the ambiguity among entities is to link the extracted entities to an existing knowledge base, which is called Entity linking [18]. Entity linking is a task that identifies the words represented entities in unstructured data (i.e., the so-called mention, a reference of an entity) and finds the corresponding entity represented by the mention from the knowledge base. Namely, it refers to link the entity recognized from the text data to the correct entity corresponding to the existing knowledge base [19]. Through knowledge fusion, the ambiguity, redundancy and errors in the extracted entities, relations and attributes can be eliminated to promote the quality of the acquired knowledge.

2.4 Knowledge Processing

Through above information extraction and knowledge fusion, knowledge elements including entities, relations and attributes are extracted from multi-sources original data, and they are clustered and eliminated ambiguity. Nevertheless, the knowledge elements are still scattered and not treated as the knowledges. To finally obtain a structured, networked knowledge system, knowledge processing techniques should be done, which refer to ontology extraction, quality evaluation and knowledge inference.

Ontology can be manually constructed by manual editing (with the help of ontology editing software) or in a data-driven automated manner. Due to the enormous workload in manual way and the difficulty of finding qualified experts, the current mainstream of ontology extraction expands from some existing ontology libraries oriented to specific fields using automated construction technology. The process of automatic ontology construction includes three stages: 1) similarity calculation of entities in coordination relation; 2) hyponymy relation extraction of entities; 3) ontology generation.

When the ontology is successfully constructed, most of the relations between knowledges are incomplete, and the missing values are severe. At this time, knowledge inference is employed to complete and update the knowledge graph. There are three categories: logic-based inference, graph-based inference, and deep learning-based inference. Quality evaluation is also an essential part of knowledge graph construction. It can quantify the confidence of knowledges and then abandon the ones with low confidence values to ensure the quality of the knowledge base.

3 Knowledge Graph Construction for Quality and Safety Supervision of Transportation Engineering

3.1 Overview of Knowledge Graph Construction

At present, using the technologies introduced above, general domain knowledge graphs have been constructed, such as Freebase, DBpedia, CN-DBpedia, and applications based on encyclopedia knowledge graphs are also emerging. However, there are few research and knowledge graph applications in quality and safety management of transportation engineering construction, primarily due to the vast engineering construction system, massive sources, and unstructured data. Therefore, this paper uses modern information technology such as ontology construction technology, big data processing technology, and natural language processing technology to process the data coming from the engineering construction system, and then constructs the ontology model to construct a special knowledge graph for quality and safety supervision of transportation engineering construction.

To construct this special knowledge graph is different from the general one. It is challenging of representing the professional knowledge from the limited semi-structured and unstructured data, which are always from various sources. Therefore, we first finish the ontology construction to determine the entities, relations and attributes in the field of transportation engineering management, and then annotate the collected multi-sources data for information extraction, finally to construct a knowledge graph for quality and safety supervision of transportation engineering projects.

3.2 Ontology Construction

Ontology in knowledge graph is mainly used to describe the relationship among concepts (entities) in a particular domain, and make them have a standard, clear and unique definition within the scope of sharing. The knowledges corresponding to each application domain are different. The construction method of general knowledge graph is not necessarily applicable to the construction of knowledge graph in a particular domain. In order to construct a domain special knowledge graph, the domain ontology must first be determined.

When establishing the knowledge graph in the domain of transportation engineering construction for quality and safety supervision and management, we first determine the relations between the concepts (entities). By studying the relevant regulations and processes of quality and safety supervision and summarizing the management experience of quality and safety supervision, the data generated by multi-sources, such as the construction personal management, quality management, safety management, engineering data, and other systems, are systematically sorted and classified. The critical risk points of the construction site are analyzed. The quality and safety supervision requirements are analyzed. As a result, the knowledges of project persons related to the quality and safety supervision of the transportation engineering are determined. On the other hands, we studied the persons of the engineering construction projects, engineering safety, engineering quality, and engineering data, and classified, analyzed, and sorted out dangerous operation license management, hidden danger investigation and management, hazard source management, and mechanical equipment management. Finally, we learned the knowledges about the quality and safety information on site.

The entities in the knowledges of project personal management for the quality and safety supervision mainly involve the persons, companies and projects, and their attributes link the corresponding entities. The entities in the knowledges of the quality and safety information of construction site are the name of each project and company. The attributes of each entity include construction content, scheme, materials, and so on. The triple “entity-relation-entity” is used to store knowledge, and finally, a knowledge graph in a form of net is formed.

Fig. 2 is an example diagram of part of the knowledge graph for the project personal management of the quality and safety supervision of the transportation engineering. The entities including persons, companies, projects, and so on, which are denoted as nodes in the knowledge graph. The entities are linked by their attributes or relations. For example, seen from the Fig. 2, the native place of ‘Zhang San’ and ‘Li Si’ is ‘Liu Yang’, so entity ‘Li Si’ and ‘Liu Yang’ is linked by their relation “Hometown”. And entity “Hunan Road and Bridge Co.” is a company, who employs “Li Si” and “Wang Wu”. So in the knowledge graph, there exist edges between entity “Hunan Road and Bridge Co.” and “Li Si” with attribute “Employer”.

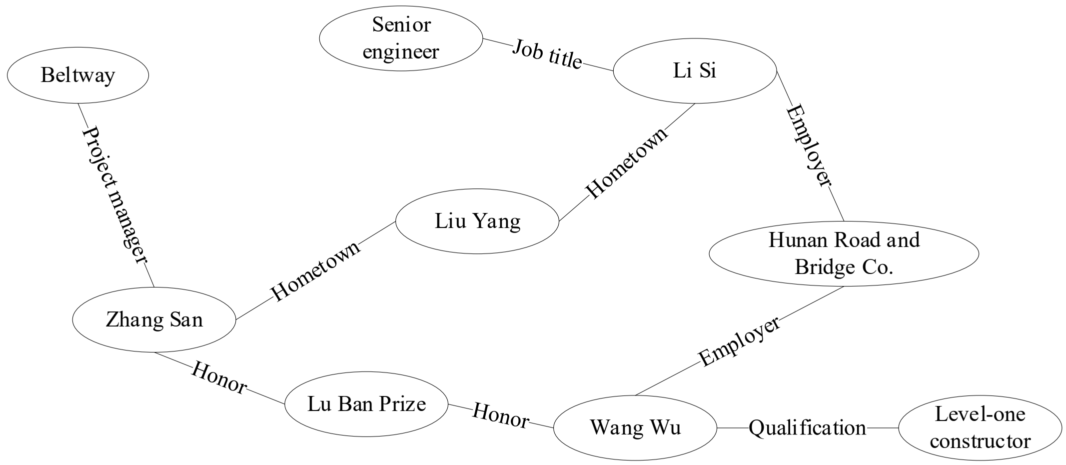

Figure 2: An example diagram of the constructed knowledge graph

The knowledge graph constructed for quality and safety supervision of transportation engineering refers the projects, companies, and persons. The information of project experiences, technical achievements, and work dynamics of the prominent participants in various fields of engineering construction will be collected from the management platforms such as ministries and provinces' government affairs, and special personnel management system. Each relatively independent application and project management system provide engineering conditions and personnel behavior information. Through the collection, filtering, and mining of above information, and analyzing the relationship between the entities and attributes of each project participants and projects, a domain special knowledge graph will be constructed including the knowledges about projects and participants.

The construction of domain special knowledge graph can facilitate people to query the knowledges of all kinds of projects and participants in the field of transportation engineering management, and thus improving the capability and efficiency of quality and safety supervision.

3.3 Structured Information Extraction from Transportation Engineering Data

The above ontology construction has known the entities and relations (attributes) in the constructed knowledge graph. Now, we introduce how to extract the entities and relations (attributes) from the raw data. These data for the knowledge graph construction of quality and safety supervision of transportation engineering are from various sources, including structured, semi-structured, and unstructured data. So we must extract knowledges from heterogeneous multi-sources data. We utilize a series of automatic or semi-automatic technologies to extract knowledge facts from the original database and third-party databases, which are stored in the data layer and model layer of the knowledge base. Structured data are integrated directly into the knowledge base after the validity of the data is verified. Semi-structured and unstructured data are always complicated.

In order to effectively extract knowledges from these raw complicated data, we design a software system to preprocess semi-structured and unstructured data. The software comprises two primary functions: text extraction and file extraction. The primary function of text extraction is to select and extract the critical information by copying and pasting the text from the raw data for later analysis. The primary function of document extraction is to select and extract the critical information by uploading the document of the transportation engineering projects for later analysis. When the extracted contents are wrong, some minors can be automatically corrected. And the software can be automatically optimized according to the error report to improve the accuracy of information extraction. If the extraction information is correct, the extracted information related to the knowledges of quality and safety supervision of transportation engineering including the full name, abbreviation, location, cost, start time, end time data, and so on, will be stored in the server database for the subsequent processing.

After preprocessing the semi-structured and unstructured data, firstly an entity recognition method is carried out. And the knowledges contained in these two types of data are extracted through relations extraction and attribute extraction methods. Then through entity alignment technology, the entities extracted from multiple data sources are aligned, and calibration of entity attributes to avoid data duplication. Finally, the edges of the knowledge graph are completed by knowledge inference. After the quality evaluation, the triples of (entity, relation/attribute, entity) are added to the knowledge graph to update the knowledges for the following services.

4 Visual Information Query from the Constructed Knowledge Graph

4.1 Query Design

The constructed knowledge graph used for supervising the quality and safety of transportation engineering projects helps to push the related content and events according to different members and show the relationship among various members. From a technical point of view, it is to find the corresponding entities from the immense knowledges stored in the knowledge graph and complete the knowledge retrieval work, that is, it is necessary to efficiently perform sub-graph matching and query. Consequently, we study the application of the constructed knowledge graph in the supervision and management of the quality and safety of transportation engineering construction projects. So we construct a domain special knowledge graph and employ subgraph matching technology, finally, we implement a visual information query software for highly effective supervision and management of the transportation engineering projects.

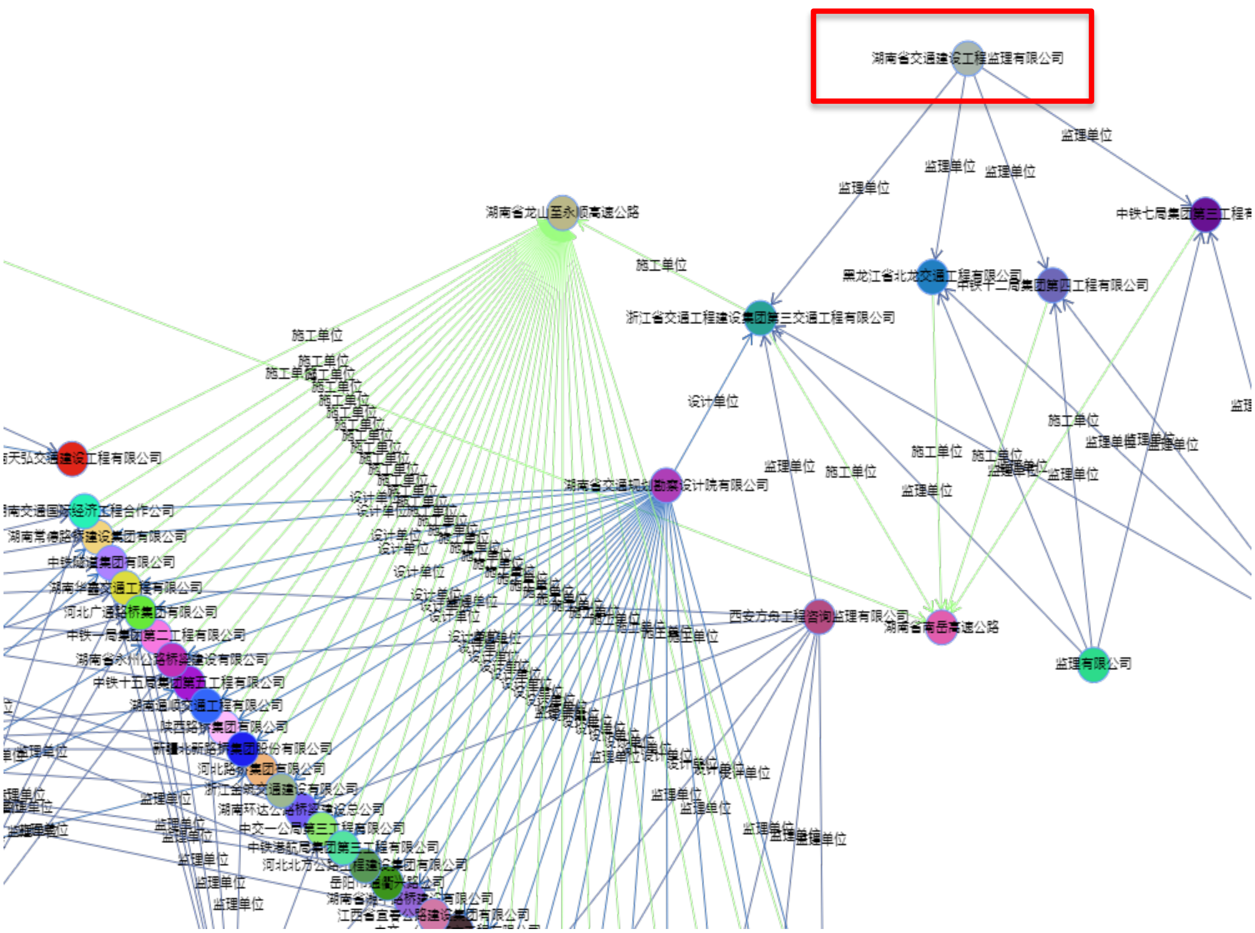

Figure 3: An example of querying the constructed knowledge by the company

The visual information software system mainly provides two primary query functions: company query and road section query (i.e., transportation engineering construction project query). The company query refers to acquire the knowledges about the specific information of a company and its associated construction projects with their associated information. Meanwhile, road section query mainly searches about the specific information of a particular construction project and the construction companies associated with it. For the company query, when the entered company has no associated construction project, only the company name will be displayed, and no associated knowledge graph will be displayed. If the entered company queries the associated construction projects, which have a construction relation with the entered company. In this case, it will display all relevant information associated with the construction projects. When the user query by the road section (the construction project), there will be no relevant graph display if the entered road section does not exist; otherwise, all relevant information of the construction companies associated with it will be displayed through the corresponding knowledge graph.

4.2 Results and Analysis

We collected data from multiple data sources including the information of some transportation engineering projects in China. The knowledges in the constructed knowledge graph are represented in Chinese. So as that the input of the visual information query system should be Chinese. Inputting the name of a company or a road section can query the knowledges associated it in the form of a graph, which takes entities as nodes and attributes or relations as edges. When we input the company name "湖南省交通建设工程监理有限公司"(Hunan Transportation Construction Engineering Supervision Co., Ltd.), the resulting knowledge graph is shown in Fig. 3. While we input the name of a road section "湖南省龙山至永顺高速公路"(Longshan-Yongshun Expressway in Hunan Province), the resulting knowledge graph is shown in Fig. 4. The inputted entities are marked with a red box.

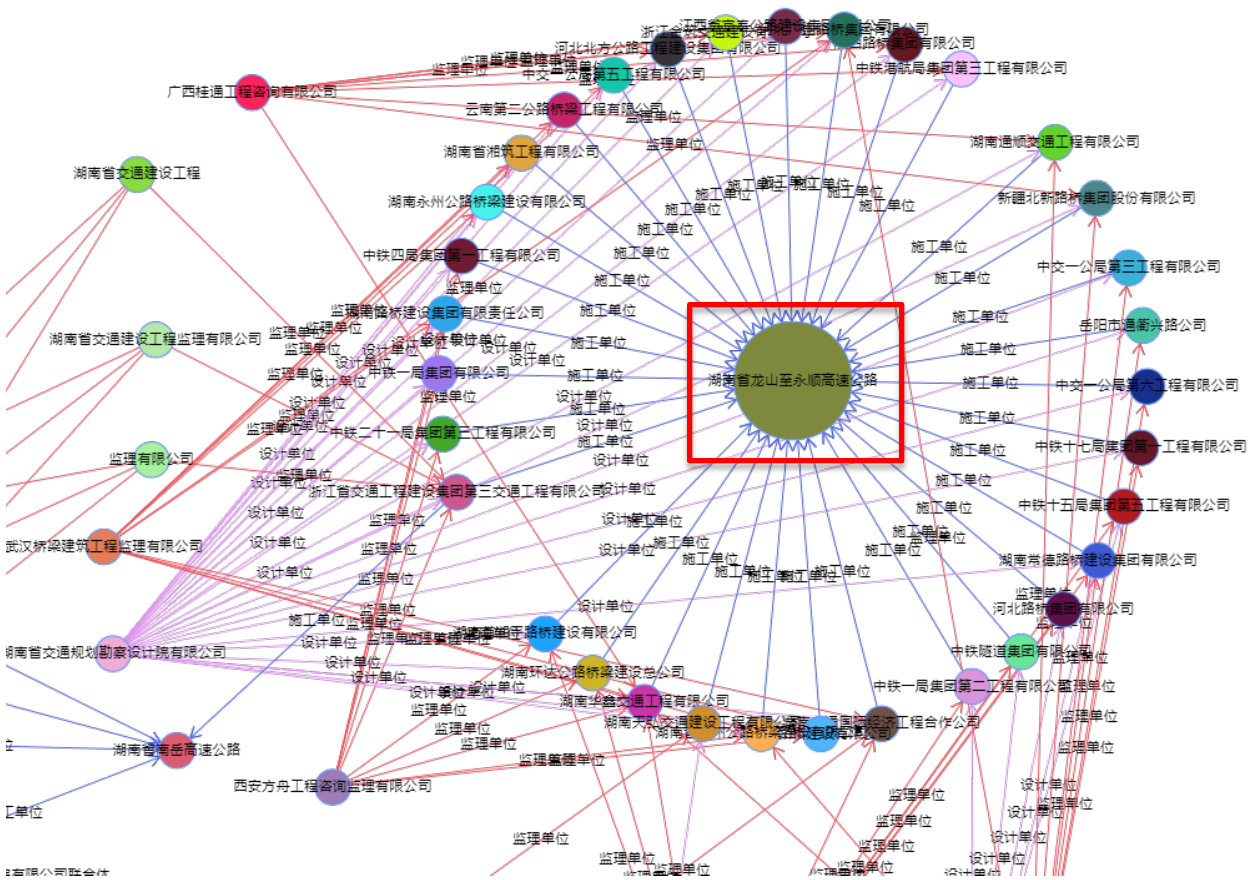

Figure 4: An example of querying the constructed knowledge by the road section

Seen from the Fig. 3 and Fig. 4, we can find that this visual information query system can help users quickly and accurately to find the information they need, which are complicated and associated to several entities. As a result, the quality and safety supervision and management of transportation engineering projects can be promoted to a more intellectual level.

5 Conclusion

Nowadays, a great number of knowledge graphs were constructed and published for general and popular domains in English. The development of Chinese knowledge graphs and their applications are slowing. Moreover, there lacks the knowledge graph research on the domain of transportation engineering construction. Thus, both academia and industry need to pay attention to construct knowledge graphs for improving the capability of quality and safety supervision and management for transportation engineering construction. This paper studies the technologies of constructing general knowledge graphs, which are very good references. And then focuses on construction of the domain special knowledge graph for transportation engineering construction. The acquired knowledges are represented and stored in the form of triples. In order to efficiently use the constructed knowledge graph, the query results from the knowledge graph are visualized in the form of a graph. By means of the domain special knowledge graph, the capability and efficiency of quality and safety supervision and management of transportation engineering can be improved.

Funding Statement: This work was supported by Scientific Research Project of Department of Transportation of Hunan Province under Grant No. 201814.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] J. Lin, Y. Zhao, W. Huang, C. Liu and H. Pu, "Domain knowledge graph-based research progress of knowledge representation," *Neural Computing and Applications*, vol. 33, pp. 681–690, 2021.

[2] Y. Tang, T. Liu, G. Liu, J. Li, R. Dai *et al.,* "Enhancement of power equipment management using knowledge graph," in *Proc. of the 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*, Chengdu, China, pp. 905–910, 2019.

[3] P. Chen, Y. Lu, V. W. Zheng, X. Chen and B. Yang, "Knowedu: A system to construct knowledge graph for education," *IEEE Access*, vol. 6, pp. 31553–31563, 2018.

[4] T. Wu, G. Qi, L. Cheng and W. Meng, "A survey of techniques for constructing Chinese knowledge graphs and their applications," *Sustainability*, vol. 10, no. 9, 3245, 2018.

[5] Q. Liu, Y. Li, H. Duan, Y. Liu and Z. Qin, "Knowledge graph construction techniques," *Journal of Computer Research and Development*. vol. 53, no. 3, pp. 582–600, 2016.

[6] L. F. Rau, "Extracting company names from text," in *the Seventh IEEE Conf. on Artificial Intelligence Application*, Miami Beach, FL, USA, pp. 29–32, 1991.

[7] X. H. Liu, S. D. Zhang, F. R. Wei and M. Zhou, "Recognizing named entities in tweets," in *Proc. of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, vol. 1. Association for Computational Linguistics, Portland, Oregon, USA, pp. 359–367, 2011.

[8] S. Misawa, M. Taniguchi, Y. Miura and T. Ohkuma, "Character-based bidirectional LSTM-CRF with words and characters for Japanese named entity recognition," in *Proc. of the First Workshop on Subword and Character Level Models in NLP*, Copenhagen, Denmark, pp. 97–102, 2017.

[9] N. Kambhatla, "Combining lexical, syntactic, and semantic features with maximum entropy models for extracting relations," in *Proc. of the ACL 2004 on Interactive Poster and Demonstration Sessions*, Association for Computational Linguistics, Barcelona, Spain, 2004.

[10] G. D. Zhou, J. Su, J. Zhang and M. Zhang, "Exploring various knowledge in relation extraction," in *Proc. of the 43rd Annual Meeting on Association for Computational Linguistics*, Stroudsburg, PA, USA, pp. 427–434, 2005.

[11] S. B. Zhao and R. Grishman, “Extracting relations with integrated information using kernel methods,” in *Proc. of the 43rd Annual Meeting on Association for Computational Linguistics*, Stroudsburg, PA, USA, pp. 419–426, 2005.

[12] D. J. Zeng, K. Liu, S. W. Lai, G. Y. Zhou and J. Zhao, “Relation classification via convolutional deep neural network,” in *Proc. of COLING*, Dublin, Ireland, pp. 2335–2344, 2014.

[13] D. J. Zeng, K. Liu, Y. B. Chen and J. Zhao, “Distant supervision for relation extraction via piecewise convolutional neural networks,” *EMNLP*, Lisbon, Portugal, 2015.

[14] E. Alfonseca, M. Pasca and E. Robledo-Arnuncio, “Acquisition of instance attributes via labeled and related instances,” in *Proc. 33rd Int. ACM SIGIR Conf. on Research and Development in Information Retrieval*, New York, NY, USA, pp. 58–65, 2010.

[15] M. Yang, K. Chen, S. Sun, Z. Han, L. Kong *et al.,* “A pattern driven graph ranking approach to attribute extraction for knowledge graph,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 2, pp. 1250–1259, 2022

[16] T. Wang and H. Li, “Coreference resolution improves educational knowledge graph construction,” in *2020 IEEE Int. Conf. on Knowledge Graph (ICKG)*, Beijing, China, pp. 629–634, 2020.

[17] K. Zhang, Y. Zhu, W. Gao, Y. Xing and J. Zhou, “An approach for named entity disambiguation with knowledge graph,” in *2018 Int. Conf. on Audio, Language and Image Processing (ICALIP)*, Shanghai, China, pp. 138–143, 2018.

[18] N. Jia, X. Cheng, S. Su and L. Ding, “CoGCN: Combining co-attention with graph convolutional network for entity linking with knowledge graphs,” *Expert Systems*, vol. 38, no. 1, pp. 1–13, 2021.

[19] X. P. Han, L. Sun and J. Zhao, “Collective entity linking in web text: a graph-based method,” in *Proc. of the 34th Int. ACM SIGIR Conf. on Research and Development in Information Retrieval*, ACM, New York, NY, USA, pp. 765–774, 2011.