Proceedings of the 6th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/71/20241632

Overview of knowledge graph construction in the field of fully electronic computer interlocking systems

Yangwenjing Li

Lanzhou Jiaotong University, Anning West Road No. 88, Anning District, Lanzhou, Gansu Province, China

824320325@qq.com

Abstract. As a core technology in railway signal control, the fully electronic computer interlocking system achieves the control of railway signals and switches through computerization, ensuring the safety and efficiency of train operations. With the increasing complexity of railway systems, traditional management and maintenance face significant challenges. Knowledge graphs, as an advanced data management and representation method, can effectively organize and utilize large-scale domain knowledge, enhancing the system's intelligence. This paper introduces the construction methods, core concepts, and applications of knowledge graphs in the field of fully electronic computer interlocking systems, providing references for research and practice in this area.

Keywords: Knowledge Graph, Fully Electronic Computer Interlocking System, Railway Signals, Data Management.

1. Introduction

1.1. Research Background and Significance

The all-electronic computer-based interlocking (CBI) system is a core component of modern railway transportation, controlling signal equipment and switches through computer technology to ensure the safe and efficient operation of trains. With the continuous growth in global railway transportation demand, the complexity and scale of railway systems are also increasing. Traditional manual management and maintenance methods can no longer meet the complex demands of modern railway systems, making it urgent to introduce new technological means to enhance system safety, reliability, and intelligence [1]. Knowledge graphs, an emerging technology for organizing and representing domain knowledge in a structured manner, provide support for intelligent system management.

1.2. Research Objectives and Methods

This paper aims to explore the construction methods, core concepts, and applications of knowledge graphs in the field of all-electronic CBI systems. By reviewing relevant literature, analyzing data, and conducting case studies, this paper systematically summarizes and analyzes the current state of research and development trends in this field, proposing an intelligent management framework for all-electronic CBI systems based on knowledge graphs.

2. Overview of All-Electronic Computer-Based Interlocking Systems

2.1. System Components and Functions

The all-electronic CBI system controls railway signal equipment and switches in real-time through computer programs, ensuring that trains run safely along predetermined routes and speeds[2]. The system comprises central control systems, signal equipment, sensors and detection devices, and monitoring and diagnostic systems.

2.2. Working Principles

The all-electronic CBI system controls signal equipment and switches in real-time via computer programs, using sensors to collect real-time data. It combines preset rules and logic to control the signal equipment and switches, preventing train collisions or derailments. The interlocking system utilizes its core interlocking logic to make comprehensive judgments about the status of signal equipment and switches, ensuring the safe operation of trains.

2.3. Main Features of All-Electronic CBI Systems

Compared to traditional mechanical and electrical interlocking systems, all-electronic CBI systems are characterized by high reliability, high safety, flexibility and scalability, and efficiency[3].

High Reliability: The all-electronic CBI system employs redundancy design and fault-tolerance technology to ensure normal operation during hardware and software failures. Technologies such as redundant control units and dual-system hot backup enhance system reliability.

High Safety: The system ensures train safety through strict interlocking logic and real-time monitoring. All control commands and operations must meet the requirements of the interlocking logic to prevent train collisions and derailments.

Flexibility and Scalability: The all-electronic CBI system can be flexibly configured and expanded according to actual needs. Software updates and system expansions allow easy addition of new signal equipment and switches, meeting the development needs of the railway network.

Efficiency: The system can quickly process large amounts of real-time data and make rapid decisions based on interlocking logic, ensuring efficient train operation. Optimized control algorithms and high-speed communication networks significantly improve train operation efficiency and punctuality.

3. Overview of Knowledge Graphs

3.1. Definition and Classification

A knowledge graph, also known as a scientific knowledge graph, is a series of graphs that display the development processes and attribute relationships of knowledge, using corresponding visualization techniques to represent the relationships between knowledge entities and attributes in a structured manner. Essentially, it is a networked knowledge base composed of knowledge triples[4]. By systematizing and structuring complex domain knowledge, knowledge graphs support various application scenarios such as information retrieval, question answering systems, and decision support.

Currently, knowledge graphs are classified into two types: general knowledge graphs and specialized knowledge graphs. This classification is based on the scope of application. General knowledge graphs cover a wide range of fields and related content, serving as a structured, multidimensional encyclopedic knowledge base. However, general knowledge graphs have lower precision and face challenges in standardizing relationships and attributes between knowledge entities[5]. Specialized knowledge graphs, constructed top-down within a specific domain, provide strong domain expertise by defining the scope of related knowledge entities in advance and using domain-specific data for construction.

3.2. Construction Methods

The construction of knowledge graphs generally includes the following steps: knowledge acquisition, knowledge fusion, and knowledge graph storage[6].

Knowledge Acquisition: This involves extracting specific information from collected data, including entity extraction, attribute extraction, and relationship extraction[7]. Named entity recognition (NER) is the core technique for identifying special terms or proper nouns from collected texts.

Knowledge Fusion: This step involves integrating the acquired knowledge entities and relationships into knowledge triples.

Knowledge Graph Storage: The storage formats include knowledge tables, Resource Description Framework (RDF), and graph storage. Knowledge tables store triples in a two-dimensional format, RDF represents data in a subject-predicate-object format, and graph storage models each knowledge entity as a node and relationships as edges[8]. Graph databases like Neo4j, gStore, and JanusGraph are commonly used for storing knowledge graphs.

3.3. Advantages of Knowledge Graphs

As an advanced tool for knowledge representation and management, knowledge graphs structure knowledge in a way that is easily understandable and processable by computers. They integrate knowledge from different data sources, solving data silo issues and providing comprehensive and consistent knowledge representations to support cross-domain applications. Knowledge graphs are highly scalable, capable of adding new entities, attributes, and relationships to meet growing knowledge needs. They support rule-based and logical reasoning, enabling the derivation of new knowledge from existing knowledge, which is crucial for complex decision support and intelligent applications.

4. Construction of Knowledge Graphs for All-Electronic CBI Systems

Given the strong specialization of all-electronic CBI systems, dispersed professional knowledge and technical barriers within signal equipment manufacturing companies can lead to inconsistent understanding and erroneous applications among technical personnel. Thus, establishing a systematic knowledge framework for knowledge sharing within enterprises is necessary to summarize, analyze, and refine the knowledge related to all-electronic CBI systems.

4.1. Data Collection and Processing

4.1.1. Data Sources

To construct a knowledge graph in the domain of the fully electronic computer-based interlocking system, information must be collected from various data sources. This includes extracting specialized knowledge from relevant technical documents, deriving domain knowledge from industry standards and normative documents, standardizing the collected knowledge based on the expertise and experience of domain experts, and labeling the actual operational data records of the system.

4.1.2. Data Processing

The collected data needs to undergo data cleaning to remove redundant and noisy data and to fill in missing values. Data normalization ensures that data formats, units, and representations are consistent, allowing seamless integration of data from different sources. This includes standardizing date formats, converting numerical units, and normalizing names.

4.2. Knowledge Extraction

Knowledge extraction in the domain of the fully electronic computer-based interlocking system aims to retrieve relevant domain knowledge hidden in unstructured, semi-structured, or structured texts and third-party databases. This extraction primarily focuses on three aspects: domain knowledge entity extraction, domain knowledge entity attribute extraction, and domain knowledge entity relationship extraction[9].

4.2.1. Domain Knowledge Entity Extraction for Fully Electronic Computer-Based Interlocking System

In the construction of a knowledge graph, entity extraction serves as the foundation, mainly used to identify and define key entities. Common entities in the domain of the fully electronic computer-based interlocking system include signal equipment, control logic, fault types, and maintenance records. Signal equipment encompasses signal lights, switches, track circuits, etc. Control logic describes the rules and logic by which the system controls signal equipment and switches. Fault types describe various potential system faults. Maintenance records encompass the system's maintenance and inspection logs.

Domain-related texts in the fully electronic computer-based interlocking system are annotated with BIOES labeling standards for named entity recognition, assigned corresponding labels, and processed into an initial domain knowledge text dataset. A BERT-BiLSTM-CRF-based named entity recognition model is then constructed, incorporating attention mechanisms in the word vector construction via the BERT model. Feature extraction is conducted using BiLSTM, and the features are input into the CRF model for final training[9].

4.2.2. Domain Knowledge Entity Attribute Extraction for Fully Electronic Computer-Based Interlocking System

Most domain knowledge entity attributes in the fully electronic computer-based interlocking system exist in semi-structured texts. On most encyclopedic websites, semi-structured data is usually found in infoboxes or various table structures at the top of the pages. Processing semi-structured texts involves constructing web wrappers to retrieve semi-structured data from tables or infoboxes based on the corresponding page's HTML code. Entity attribute extraction aims to construct <knowledge entity, attribute, attribute value> triples[10].

4.2.3. Domain Knowledge Entity Relationship Extraction for Fully Electronic Computer-Based Interlocking System

Relationship extraction aims to construct <knowledge entity 1, relationship, knowledge entity 2> triples[11]. Relationship modeling defines the connections between entities, providing semantic associations for the knowledge graph. Identifying the types of relationships between knowledge entities establishes a comprehensive knowledge system in the domain of the fully electronic computer-based interlocking system and lays the foundation for entity relationship extraction. Common relationships in this domain include control relationships, monitoring relationships, fault associations, and maintenance records.

Control relationships describe the connections between the control system and signal equipment and switches. Monitoring relationships describe the connections between sensors and signal equipment and switches. Fault associations describe the relationships between various faults. Maintenance records describe the connections between maintenance logs and fault types, signal equipment, and switches.

4.3. Knowledge Graph Storage Tools

Neo4j graph database is used to store knowledge data in the domain of the fully electronic computer-based interlocking system. Neo4j provides a user-friendly query and display interface, supports large data sets, and has database characteristics. Neo4j stores data in the form of directed graphs composed of nodes and edges, where nodes represent entities and edges represent relationships between the start and end nodes.

5. Challenges and Future Development

This paper primarily studies the construction of a knowledge graph in the domain of the fully electronic computer-based interlocking system. Despite the convenience provided by the knowledge graph, several challenges remain, including:

1. Data Quality and Completeness: The construction of a knowledge graph relies on high-quality and complete data, which is often difficult to guarantee in practical applications. Future research needs to focus on methods for data cleaning and integration to improve data quality and consistency.

DOI: 10.54254/2755-2721/71/20241632

2. Knowledge Update Maintenance and Cross-System Integration: The railway system is continuously evolving, requiring the knowledge graph to be updated and maintained accordingly. Future research should explore automated methods for updating and maintaining the knowledge graph to ensure its timeliness and accuracy. Railway signaling systems are usually closely related to other systems (e.g., traffic management systems, communication systems). Future research should investigate cross-system integration methods for knowledge graphs to achieve more comprehensive intelligent management and control.

3. Security and Standardization: When applying knowledge graph technology, particular attention must be paid to system security and privacy protection. With the increasing digitalization of railway signaling systems, the risk of cyber-attacks and data breaches also increases. Future research should focus on developing more secure methods for constructing and applying knowledge graphs to ensure system security and user privacy. The construction and application of knowledge graphs require unified standards and norms to ensure interoperability and compatibility between different systems[12]. Future efforts should aim to establish relevant industry standards and norms to promote the widespread application of knowledge graph technology in railway signaling systems.

6. Conclusion

The knowledge graph in the domain of the fully electronic computer-based interlocking system provides new ideas and methods for the intelligent management of railway signaling systems. By constructing and applying knowledge graphs, system operational efficiency and safety can be improved, fostering innovation and development in railway signaling technology. Future research can further optimize the methods for constructing and applying knowledge graphs and explore their potential applications in broader fields.

References

[1] Dou Baoqin. Discussion on the Status and Role of Railway Signals in Railway Transportation [J]. Science and Technology Wind, 2014(1): 230.
[2] Lu Nanmei. The Development of Railway Signal Technology in China [J]. Communication World, 2015(4): 18.
[3] Tian Jiaming, Liu Hongyan. Discussion on the Fully Electronic Modular Computer-Based Interlocking System [J]. Engineering Science II, 2009, 56(23): 1-11.
[4] Xu Zenglin, Sheng Yongpan, He Lirong, et al. Overview of Knowledge Graph Technology [J]. Journal of the University of Electronic Science and Technology of China, 2016, 45(04): 589-606.
[5] Liu Qiao, Li Yang, Duan Hong, et al. Overview of Knowledge Graph Construction Technology [J]. Journal of Computer Research and Development, 2016, 53(3): 582-600.
[6] Wang Xin, Zou Lei, Wang Chaokun, et al. Overview of Knowledge Graph Data Management Research [J]. Journal of Software, 2019, 30(07): 2139-2174.
[7] Wang Qingwen. Research on Railway Signal Basic Equipment Maintenance Decision Method Based on Knowledge Graph [D]. Gansu: Lanzhou Jiaotong University, 2023.
[8] Wang Yang. Research and Implementation of Knowledge Graph in the Railway Domain [D]. Inner Mongolia: Inner Mongolia University, 2019.
[9] Meng Hanjiao. Design and Implementation of an Intelligent Q&A System for Railway Signal Accidents Based on Knowledge Graph [D]. Hebei: Hebei University of Science and Technology, 2020: 23-45.
[10] Feng Haojun, Duan Li, Zhang Biying. Overview of Knowledge Reasoning for Knowledge Graphs [J]. Computer Systems Applications, 2021, 30(10): 21-30.
[11] Yang Lianbao. Research and Application of Key Technologies for Big Data Analysis of Railway Accident Fault Texts [D]. Beijing: China Academy of Railway Sciences, 2018: 26-39.
[12] Wang Zhiyue, Yu Qing, Wang Nan, et al. Overview of Intelligent Q&A Research Based on Knowledge Graph [J]. Computer Engineering and Applications, 2020, 56(23): 1-11.