

# Research and Application of Methodologies for Constructing Knowledge Graphs in Secondary School Mathematics

Hao Wang<sup>1</sup>, Ke Xiong<sup>1,\*</sup>, Pengzhang Yu<sup>2</sup>

<sup>1</sup>College of Electronic and Information, Southwest Minzu University, Chengdu 610225, China

<sup>2</sup>Key Laboratory of Electronic Information Engineering, Southwest Minzu University, Chengdu 610225, China

\*Corresponding Author: [wh991009@gmail.com](mailto:wh991009@gmail.com)

## ABSTRACT

The knowledge graph of secondary school mathematics aims to enhance the efficiency and quality of mathematics teaching at the secondary level by constructing relationships between mathematical concepts. Traditional mathematics teaching often follows a linear approach, lacking flexibility and personalization. In contrast, methods based on knowledge graphs organize mathematical knowledge into graphical structures, enabling students to understand the connections and dependencies between various concepts more clearly. Utilizing Neo4j to store and manage the relationships between mathematical concepts provides comprehensive visual data support for the knowledge graph of mathematics. This approach may enhance students' interest and efficiency in learning mathematics and assist teachers in delivering better lessons.

## KEYWORDS

Secondary School Mathematics; Knowledge Graph; Neo4j.

## 1. INTRODUCTION

For the subject of junior high school mathematics, the hierarchical structure of knowledge exhibits clarity and orderliness. However, different students vary in their understanding of mathematical knowledge and learning needs. Each student possesses their own learning style, interests, preferences, and knowledge background, leading to diverse requirements for teaching resources. Traditional uniform teaching methods struggle to meet the personalized learning needs of each student.

With the development of big data and artificial intelligence, various technological teaching platforms and methods related to disciplines are gradually emerging. However, as the number of online education platforms increases, most of them overlook the integration of unstructured information [1]. Therefore, applying knowledge graphs to these platforms enables them to effectively acquire knowledge.

The technology of secondary school mathematics based on knowledge graphs aims to propose a method that utilizes secondary school mathematics textbooks and curricula, under the guidance of subject experts, to construct a knowledge graph of secondary school mathematics. Meanwhile, leveraging the Neo4j graph database and related visualization tools, the technology can effectively display the structure of disciplinary knowledge and the relationships between various knowledge points, which align well with learners' cognitive habits.

## **2. TECHNICAL PRINCIPLES**

### **2.1. Knowledge Graph Technology**

Knowledge Graph is a structured graph model used to effectively represent and organize knowledge. Since Google introduced its first version of Knowledge Graph in May 2012, there has been a surge of research interest in knowledge graph both in industry and academia. Following Google, many companies have also embarked on constructing knowledge graphs. However, apart from general and open-world knowledge graphs, most knowledge graphs currently are domain-specific. Unlike open-world and general knowledge graphs, most domain-specific knowledge graphs constructed and populated are not publicly available on the web, thus inaccessible to other researchers. In vertical domain knowledge graphs, deep learning research in knowledge graph construction is more extensive in fields such as finance and business, but relatively less applied in academic domains. This may be due to the relative difficulty in acquiring knowledge points and datasets in academic domains, as well as the need for further improvement in deep learning algorithms. Continuous efforts are needed both domestically and internationally to conduct in-depth research and unearth more application value.

In the era of big data, knowledge graphs are seen as an innovative and efficient way of organizing knowledge, and significant progress has been made. General knowledge graphs such as Freebase, DBpedia, and Wikidata cover a vast amount of common-sense knowledge in the real world. Additionally, knowledge graph technology has been the subject of research and application in various fields, including search engines, legal regulations [5, 6], medical diagnosis [7, 8], and others. They have actively contributed to the development of various fields and are considered one of the core driving forces behind the advancement of the Internet and artificial intelligence.

This article focuses on the field of secondary school mathematics, integrating knowledge graph methods to construct a knowledge graph for subject development. It comprehensively integrates information for the construction of secondary school mathematics subjects while researching technologies such as knowledge question answering and visualization displays. This enables students and teachers to gain a more intuitive understanding of the content of secondary school mathematics.

### **2.2. Technology of Secondary School Mathematics Knowledge Graph**

Constructing a knowledge graph of secondary school mathematics utilizes the fundamental methods of knowledge graphs to represent data in a graphical structure. It organizes various knowledge points in secondary school mathematics as entities, forming a network of entity relationships and attributes, thus establishing a knowledge graph for the convenience of both students and teachers in learning. By constructing the knowledge graph, it becomes possible to integrate different data sources and reveal their correlations, thereby achieving comprehensive and accurate knowledge representation.

This knowledge graph integrates various knowledge points, including those from major chapter sections, sub-knowledge points within these chapters, and miscellaneous knowledge points. It benefits both students and teachers by providing a comprehensive understanding of the content of knowledge points and the relationships between them, thus aiding students in learning and assisting teachers in lesson preparation.

### **2.3. Construction of Secondary School Mathematics Knowledge Graph**

The construction of a junior high school mathematics knowledge graph revolves around the core issue of data. This study relies on authoritative textbook materials, as textbooks are typically developed based on mathematics syllabi issued by educational authorities, ensuring the rigor of the data. Through research on knowledge graph theories and technologies, and addressing the current lack of publicly available knowledge graphs for junior high school mathematics in the educational field, this study successfully constructed a junior high school mathematics knowledge graph using junior high

school mathematics textbooks and syllabi. Leveraging OCR text recognition technology and entity labeling techniques, the construction process was conducted semi-automatically under the guidance of relevant experts.

The main sources of knowledge for this study are the syllabus for junior high school mathematics teaching and six PDF textbooks from the East China Normal University Press. Text in PDF format possesses stable characteristics that make it difficult to copy and modify, which hinders subsequent entity extraction and relationship labeling work.

The conversion of raw data requires the use of OCR (Optical Character Recognition) technology. OCR technology is a technique used to convert characters, such as text and numbers, from images into editable text. It analyzes and processes the image to recognize the characters and then converts them into a text format that can be processed by a computer, facilitating subsequent data processing and analysis. With OCR technology, it is possible to automate the processing and conversion of non-structured data from paper documents, image files, and other sources.

As shown in Figure 1, the images obtained in this study were processed through PaddleOCR for OCR text recognition. In order to better present the OCR text recognition results, the textbooks and the recognition results are juxtaposed for comparison, with the textbook on the left and the recognition results on the right.



**Figure 1.**Textbook PDF Recognition Results

This study aims to construct a knowledge graph of junior high school mathematics, with entity extraction being a key step in this process. To ensure the high accuracy of the knowledge graph, a semi-automatic method employing a top-down approach was adopted, with manual annotation as the main approach and machine algorithms assisting in organizing data to build nodes. This semi-automatic construction method comprises a manual extraction stage and an algorithmic processing stage. Through this approach, although the manual construction stage requires a significant amount of work and time, it ensures the accuracy and reliability of the knowledge graph of junior high school mathematics.

### 3. KNOWLEDGE VISUALIZATION QUERY

#### 3.1. Knowledge Graph Visualization

To help teachers and students better understand the various mathematical concepts and their relationships in the junior high school mathematics knowledge graph, we have designed a visualization interface. In this interface, different types of entities are represented by different colors. When teachers and students hover the mouse over a mathematical concept, the system will immediately display a brief introduction to that concept.

#### 3.2. Knowledge Point Query in the Knowledge Graph

When students or teachers need to query a knowledge point, they enter the query term into the search box. The system automatically retrieves the relevant knowledge points and displays them on the page to assist students and teachers in understanding the knowledge. For example, if a teacher enters "trigonometric functions" into the search box, the interface will display relevant knowledge, formulas, concepts, etc., from the textbook about trigonometric functions. This allows the teacher to have a clearer understanding of trigonometric functions and better impart knowledge to students.

#### 3.3. Relationship Query in the Knowledge Graph

In the mathematics knowledge graph, students and teachers can easily query the mathematical concepts they are interested in and explore the relationships between these concepts and other entities. For example, if a student or teacher wants to understand the relationship between "trigonometric functions" and other mathematical concepts, they can enter "trigonometric functions" into the input box. The system will then quickly retrieve the knowledge entities related to trigonometric functions and display the triple nodes associated with trigonometric functions.

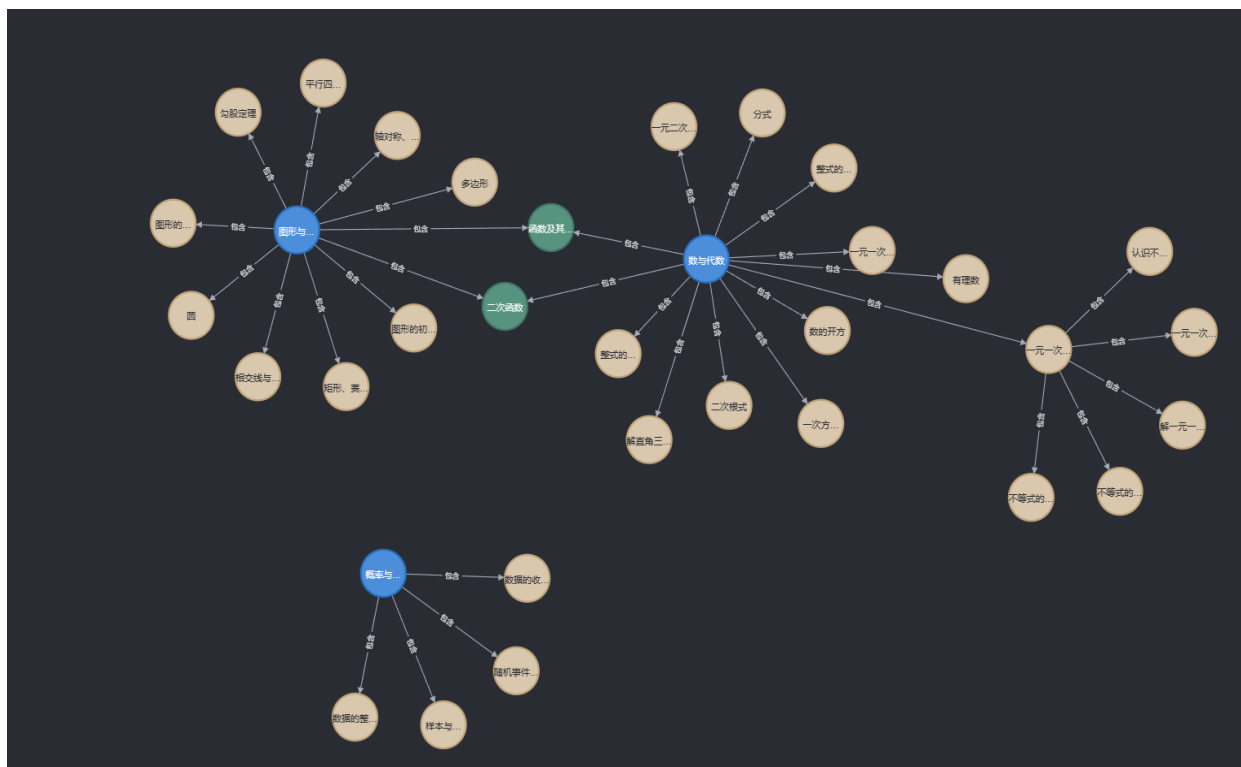


Figure 2. Construction Process of Secondary School Mathematics Knowledge Graph

## **4. INTELLIGENT ANALYSIS OF MATHEMATICS KNOWLEDGE GRAPH**

The mathematics knowledge graph can play various roles in teaching. The secondary school mathematics knowledge graph built using Neo4j is illustrated in Figure 2. Different nodes represent different entities, and the connections between nodes represent the relationships between entities. Combined with knowledge graph visualization, knowledge point queries and relationship queries within the knowledge graph can provide reasonable assistance for student learning, teacher lesson preparation, curriculum planning and design, personalized learning paths, and instructional support.

### **4.1. Curriculum Planning and Design**

During the process of curriculum planning and design, teachers can harness the power of the mathematics knowledge graph as a potent tool. By conducting a thorough analysis of the relationships between mathematical concepts within the knowledge graph, teachers can gain a clearer understanding of the structure and sequence of the curriculum. They can accurately identify key core concepts and determine the logical connections and dependencies between these concepts. This enables teachers to organize and arrange course content in a targeted manner, ensuring the coherence and integrity of knowledge transmission and the learning process.

Moreover, leveraging the mathematics knowledge graph allows teachers to adjust the course structure more flexibly, making corresponding adjustments and optimizations based on students' learning progress and comprehension levels. This approach to curriculum planning and design, based on the mathematics knowledge graph, effectively enhances teaching efficiency, enriches students' learning experiences, and fosters their ability to deeply understand and apply mathematical concepts.

### **4.2. Personalized Learning Paths**

Personalized learning pathways are designed and customized learning schemes based on individual factors such as students' prior knowledge levels, learning objectives, and learning styles. By employing advanced technology such as knowledge graphs, students can accurately identify their knowledge needs and tailor personalized learning pathways accordingly. Teachers can utilize knowledge graphs to analyze students' mastery of knowledge points, existing knowledge gaps, and weaknesses. Subsequently, by integrating students' learning goals and interests, teachers can design personalized learning roadmaps based on relationships and dependencies within the knowledge graph, thereby providing clear guidance on learning direction and progression. Additionally, teachers can leverage knowledge graphs to recommend suitable learning materials and activities, such as textbooks, exercises, and video tutorials, to help students fill knowledge gaps and reinforce learning effectiveness. Through the design and implementation of personalized learning pathways, diverse learning needs can be better addressed, enhancing students' motivation and academic performance while promoting their holistic development and individual growth.

### **4.3. Optimizing Course Structure and Sequence**

Teachers can intricately design the structure and sequence of courses by integrating them with the knowledge graph, focusing on the upcoming content of each lesson. This integration enriches and enhances the course content. Through in-depth analysis of the relationships between entities of various knowledge points within the knowledge graph, the logical framework and knowledge system of the course can be determined. Such analysis aids in organizing related topics into a hierarchical structure, ensuring the intrinsic connections and logical coherence among course contents. With the design of this hierarchical structure, students can better understand and absorb knowledge, progressively delve into learning, and enhance learning efficiency and effectiveness. This course design and arrangement help meet students' learning needs, improve teaching effectiveness, and promote in-depth classroom teaching and comprehensive student development.

## 5. CONCLUSION

By leveraging knowledge graphs, students can more easily grasp knowledge because knowledge graphs can clearly illustrate the relationships and hierarchical structures among knowledge, helping them establish a comprehensive knowledge system. This visual learning approach not only enhances students' understanding and absorption of knowledge but also stimulates their interest in learning. At the same time, knowledge graphs provide valuable references for teachers, helping them determine the structure and sequence of courses, thereby presenting the entire course with a clear logical framework and complete knowledge system. This hierarchical structure design helps students better understand and absorb knowledge, improving learning efficiency and effectiveness. Therefore, employing methods to analyze course structure and sequence using knowledge graphs is of great significance for constructing high-quality teaching systems.

## REFERENCES

- [1] Wang J. Math-KG: Construction and Applications of Mathematical Knowledge Graph[J]. arXiv preprint arXiv:2205.03772, 2022.
- [2] Tian L, Zhou X, Wu Y P, et al. Knowledge graph and knowledge reasoning: A systematic review[J]. Journal of Electronic Science and Technology, 2022, 20(2): 100159.
- [3] Singhal, A., Introducing the Knowledge Graph: things, not strings. 2012, Google.
- [4] Abu-Salih, Bilal. "Domain-specific knowledge graphs: A survey." Journal of Network and Computer Applications 185 (2021): 103076.
- [5] Tang MW, Su C, Chen HH, et al. SALKG: A semantic annotation system for building a high-quality legal knowledge graph. Proceedings of 2020 IEEE International Conference on Big Data (Big Data). Atlanta: IEEE, 2020. 2153–2159. [doi: 10.1109/BigData50022.2020.9378107].
- [6] Schneider JM, Rehm G. Curation technologies for the construction and utilisation of legal knowledge graphs. Proceedings of the LREC 2018 Workshop on Language Resources and Technologies for the Legal Knowledge Graph. Miyazaki, 2018. 23–29.
- [7] Sun HX, Xiao J, Zhu W, et al. Medical knowledge graph to enhance fraud, waste, and abuse detection on claim data: Model development and performance evaluation. JMIR Medical Informatics, 2020, 8(7): e17653. [doi: 10.2196/17653].
- [8] Goodwin T, Harabagiu SM. Automatic generation of a qualified medical knowledge graph and its usage for retrieving patient cohorts from electronic medical records. Proceedings of the 2013 IEEE 7th International Conference on Semantic Computing. Irvine: IEEE, 2013. 363–370. [doi: 10.1109/ICSC.2013.68].