Research on Structure Analysis Framework of Industrial Cooperation Network Based on Knowledge Graph

She Junni[1], Yin Jun[1], Ji Zhe[1], Qian Ping[2]

[1]School of Economics and Management, Jiangsu University of Science and Technology, Zhenjiang, China

[2]School of Computer Science, Jiangsu University of Science and Technology, Zhenjiang, China

Keywords: Industrial cooperation network, Knowledge graph, Shipbuilding industry

Abstract: To effectively integrate multidimensional data and analyze industrial cooperation networks more comprehensively. A framework for analyzing the structure of industrial cooperation network based on knowledge mapping is constructed. A representative shipbuilding industry is selected for example validation. An industrial knowledge map capable of portraying complex relationships was established through industrial element identification and multidimensional data fusion. The study shows that the knowledge graph of shipbuilding industry has obvious advantages in grasping the direction of regional layout optimization, identifying community groups with high intimacy and key enterprises with high authority and strong innovation-driven capability.

1. Introduction

With the deepening and development of the global economic, enterprises constantly seek broader and more stable cooperation. This form of interaction unites enterprises and constitutes an industrial cooperation network involving regions, resources and enterprises, ensuring the continuity of their reproduction cycles and enhancing their competitive advantages. However, the cascade of cooperative relationships among enterprises in the industry and the increasing complexity of network connections have brought new challenges to the research and management of industrial cooperative networks. Faced with the current situation that information data in the industry is complicated, highly dispersed and difficult to show comprehensively, this paper proposes an analytical framework for the structure of industrial cooperation networks, verifies the effectiveness of the proposed method using the shipbuilding industry as an example, and provides new insights for the research and management of industrial cooperation networks.

2. Related Work

Previous studies on industrial cooperation networks have mainly used questionnaire forms or input-output correlation tables to construct networks [1, 2], and have mainly portrayed network configurations in terms of a single relationship dimension, whereas the relationships between firms in industrial cooperation networks are multidimensional and complex, and these different relationships can mutually enhance their effectiveness [3]. As Beckman and Haunschild found, the more complex the relationship between a firm and its network partners, the lower the price the firm pays for its acquired assets [4]. The resources and trust that firms share in these different relationships may enhance the social capital associated with the network [5], so measuring only a single network of relationships would overlook the broader possibilities of different types of relationships merging in a collaborative network.

Traditional single association analysis is difficult to portray the complex and multi-level enterprise relationships within the industrial cooperation network, and knowledge mapping provides a solution for the combing of multi-dimensional relationships. This linked data model can break through the limitation

  DOI: 10.25236/ieesasm.2021.093

of homogeneity of connected edges in traditional single networks, integrate complex knowledge from heterogeneous data, and realize multi-perspective and multi-level connections, as shown in Fig. 1. Meanwhile, it can efficiently traverse information and provide structures that cannot be highlighted by other types of data formats, and has more advantages in realizing basic applications such as visualizing multidimensional information, providing analysis of investment solutions, and realizing intelligent assisted decision-making [6]. By establishing an industrial knowledge map and performing aggregate analysis of multiple relationships, we can systematically and objectively present the regional economic structure of industries, determine the direction of industrial structure optimization, discover intimate groups and control nodes, and provide inspiration for collaborative innovation.

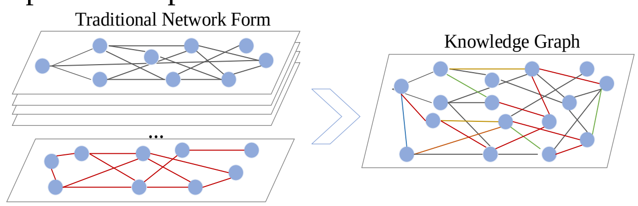

Fig.1 Diagram of Industrial Knowledge Graph

3. Method

The framework for analyzing industrial cooperation networks proposed in this paper is summarized into three parts: industrial element identification, multidimensional data fusion and industrial structure analysis, as shown in Fig. 2.

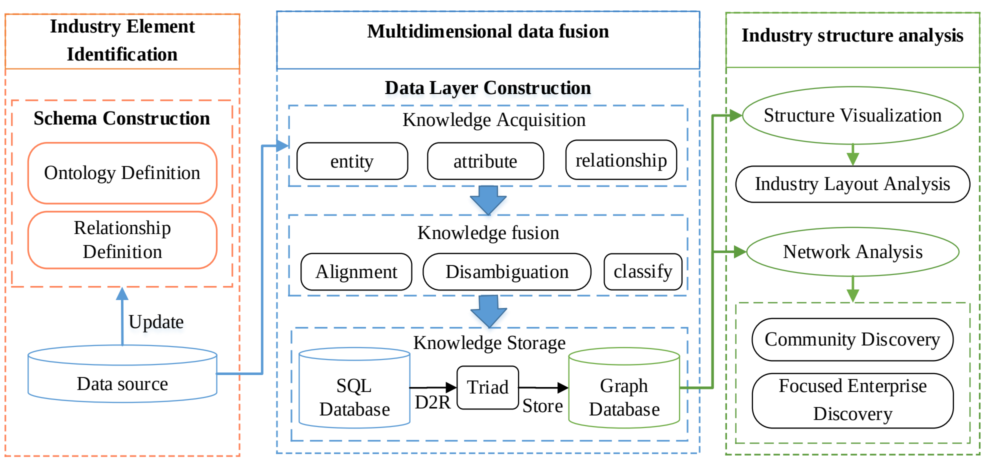

Fig.2 Analysis Framework of Industrial Cooperation Network

3.1 Industry Element Identification (Schema Construction)

The industrial cooperation network has a certain domain nature and pays more attention to the hierarchical structure of knowledge. It is necessary to build the schema in advance based on the knowledge and data of specific industries and refine the knowledge in the domain according to the characteristics of the enterprises in the industry. The schema predefines the ontologies in the industrial knowledge graph and the relationships among the ontologies, provides paradigms for the subsequent acquisition of data from different data sources, and clarifies the overall scope of the proposed graph. Meanwhile, updating the schema based on the characteristics of the acquired data sources can ensure the data quality.

3.2 Multidimensional Data Fusion (Data Layer Construction)

The construction of the data layer can be divided into three parts: knowledge acquisition, knowledge fusion, and knowledge storage. Knowledge acquisition mainly includes entity extraction, attribute extraction, and relationship extraction [7]. The industrial cooperation network has a strong territoriality and can acquire entities, attributes and links among entities directly from relevant data sources according to the definition of the schema to complete knowledge extraction. Knowledge fusion mainly aims at integrating entities, attributes and relationships obtained from different knowledge sources, solving the data conflict problem existing in the data, and then forming a standardized and complete knowledge base by aligning, disambiguating and classifying the entities. In terms of knowledge storage, the structured data in the relational database is transformed into a triad format by the method of D2R mapping, that is, entities are abstracted as nodes, relationships between entities are abstracted as edges, and attributes are abstracted as attributes of nodes and edges. And finally, it is stored in the graph database to form a complete industrial knowledge graph.

3.3 Industrial Structure Analysis (Graph Analysis)

Knowledge graph has obvious advantages for the structural analysis of industrial cooperation networks. First, knowledge graph can visualize the local structure of industry and provide reference for the optimization and adjustment of industrial layout. Second, compared with the traditional single relationship network, knowledge graph is not only more accurate in community discovery, but also can present rich information. Using community algorithms to divide the enterprises in the industry, we can discover some intimate groups, which are useful for curating the development of innovation clusters in the industry or discovering inappropriate competition behaviors. Using community algorithms, some intimate groups can be discovered, which is very useful for planning the development of innovative clusters in the industry or discovering inappropriate competitive behaviors. Finally, knowledge graph provides a more comprehensive assessment of focal firms. Quantitative analysis is performed using centrality metrics, such as eigenvector centrality for describing the long-term influence of nodes; Betweenness centrality portrays the information flow control and innovation performance of nodes [8]. The focal enterprises can be assessed to collaboratively guide other enterprises and promote the overall upgrading and development of the industrial cooperation network.

4. Empirical Research

The shipbuilding industry is highly comprehensive and has a large supporting system, which is representative in the study of industrial cooperation networks. Therefore, this paper further verifies the validity of the above framework with the example of the shipbuilding industry.

4.1 Shipbuilding Industry Element Identification

In this paper, the model is defined according to the characteristics of the shipbuilding industry. Among them, the entities include nodes such as ship enterprises and regions, and according to the characteristics of the ship industry, the enterprises are defined into three categories: one is ship manufacturing enterprises, which mainly carry out manufacturing and repair of ships and offshore engineering equipment as well as the development of engineering technologies; the second is shipping equipment enterprises, which mainly involve the research and development of relevant supporting equipment and the sales of parts and components; the third is service enterprises, which are mainly responsible for various information system services, investment management, import and export services and technical consulting. The regions include North China, Northeast China, East China, Central China, South China, Southwest China and Northwest China. The relationship of the schema is defined according to the upstream and downstream production and sales linkage of the ship industry cooperation network and the interaction in t capital flow, material flow and information flow, mainly the following five relationships: holding, investment, supply, winning bid, and the region to which the enterprise belongs.

4.2 Multidimensional Data Fusion for the Shipbuilding Industry

We selected representative state-owned shipbuilding groups and their subordinate research institutes, institutions and listed companies as research subjects. The basic information of enterprises, including enterprise name, address, main business, and shareholder information, was collected from commercial query platforms (e.g., www.tianyancha.com), official websites of enterprises, and ship material procurement information platforms. The relationship data between enterprises are also captured, such as one enterprise investing in another enterprise. After several iterations, a total of 1036 enterprise instance data and 2898 relationship data were collected.

In terms of data processing, entity alignment is performed by numbering the enterprise names with uniquely identified non-empty strings. This method is valid since the State Administration of Industry and Commerce requires that the full name of an enterprise should be unique. For entities with ambiguity, for example, the same enterprise may have different designations. We mainly used manual checking to remove ambiguities and unify different designations. In total, 529 pieces of data were processed. According to the schema definition, the ship enterprises and relationships are classified and stored in SQL database, and the data in the database are transformed into predefined Mapping files using D2RQ, and the database table names are directly mapped to the classes of RDF, the columns are used as attributes, and the relationships are derived from the primary and foreign keys of the tables. Finally, it forms a triplet format, such as “<Enterprise, Region, East China>“. Due to the high efficiency and stable performance of Neo4j graphical database, the data results are stored in Neo4j graphical database in this paper.

4.3 Shipbuilding Industry Structure Analysis

4.3.1 Regional Layout Analysi

Given the characteristics of the shipbuilding industry, China's shipbuilding industry is mainly concentrated in North, East and South China along the coast and rivers. The correlation information of these nodes is selected for visualization and analysis, as shown in Fig. 3. The shipbuilding industry in North and East China regions has better supporting infrastructure, relatively uniform industrial structure distribution, and stronger collaboration conditions compared to South China region. In contrast, large ship repairing enterprises in South China region occupy most of them, and there are fewer related supporting service enterprises, which may not be able to meet the demand of developing and using large ships in a short time, making it difficult to bring into play the cluster benefits. Therefore, on the layout of ship construction in South China region, some resources can be inclined to ship-related supporting service enterprises to optimize the industrial layout and promote the upgrading of shipbuilding capacity. By comparing the industrial layout of different regions, it helps to clarify the future industrial layout planning and form a new industrial pattern with quality and efficiency and regional coordination and support.

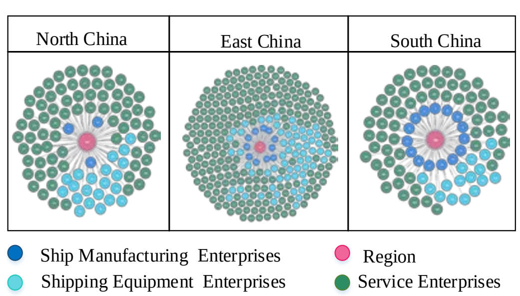

Fig.3 Comparison of Regional Industrial Distribution

4.3.2 Community Discovery

Community division through knowledge graph can reflect the agglomeration phenomenon of enterprises more accurately and present richer information on community discovery. We choose Louvain algorithm to classify the community of enterprises in the ship industry. According to the result, the Cypher statement is used to query the enterprises in the same community, as shown in Fig. 4. The graph clearly shows a small community of intimate enterprises and provides diverse relationship information. In a community, advanced technologies and industry standards are more easily promoted through the "learning effect", and the synergistic effect of industry can be maximized through collaborative management to accelerate the reasonable flow and optimal allocation of technology and talent resources. However, it should also be noted that enterprises in the same community have shareholdings and investments, and have a high degree of intimacy with each other. When the winning enterprises are in this intimate relationship many times, it is more likely that improper competition will occur in the bidding process [9]. This provides a new perspective with some guidance for the relevant regulatory authorities to manage small clusters on the industry.

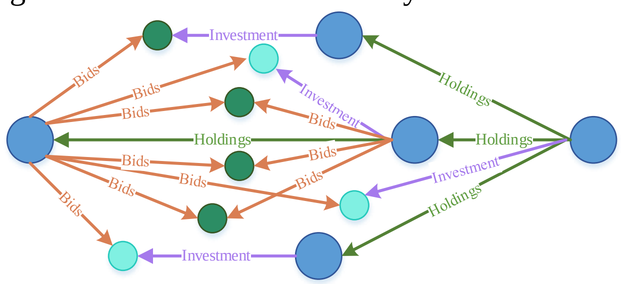

Fig.4 Query Results of Node Relationships in the Same Community

4.3.3 Regional Layout Analysi

The focal firm has richer resources in the network as well as stronger relationship management skills [10], and can promote the overall performance of the cooperative network by guiding and encouraging other firms and synergizing the roles of all parties. In this paper, we use eigenvector centrality and betweenness centrality indicators to accurately locate the firms with strong authority and innovation-driven capabilities in the industry. The results show that the company with the highest eigenvector centrality is a shipbuilding company with a score of 14.30. Within the industry, this company is in a dominant position in areas including maritime defense, development equipment, transportation, and repair and modification, and this comprehensive and synergistic structure makes it more risk resistant and stable, and has a stronger long-term influence in the industry. The enterprise with the highest betweenness centrality is a supporting equipment enterprise with a score of 13.00. This enterprise has strong scientific research support and belongs to the strategic emerging industries supported by the national new energy policy, and its industrial chain covers all major links in the upstream, midstream and downstream of the industry, occupying an important position in the industrial chain and having a certain support role in the innovation drive. Through quantitative analysis, the validity of the method proposed in this study is further verified.

5. Conclusion

Aiming at the lack of clear and all-round industrial structure description in the current research on industrial cooperation network, this paper designs an analysis framework of industrial cooperation network structure based on knowledge graph, and describes the process and method of constructing a knowledge graph of the shipbuilding industry as an example. Through industrial element identification, multidimensional data fusion, and industrial structure analysis, an industrial knowledge graph capable of portraying complex relationships is established, and the application scenarios of industrial cooperation network structure analysis based on knowledge graph are discussed. It is found that the adoption of knowledge graph can break through the traditional single network relationship model and deeply integrate multiple information and data of the industrial cooperation network. Through example verification, the optimized direction of regional layout in industry, community groups with higher intimacy and focal enterprises with high authority and strong innovation drive are discovered. It helps to adjust and optimize the industrial layout, broaden the development space, promote fair competition within the industry, strengthen the driving role of focal enterprises, improve the industrial development potential, and then promote the industrial technology revolution and optimization and upgrading.

6. Acknowledgment

This study was partially funded by the National Natural Science Foundation of China [grant numbers G0115], Postgraduate research and innovation plan project in Jiangsu Province [grant number SJCX21_1738].

References

[1] M. Li, Y. Song, and Z. Li, “Knowledge Transfer in Industrial Cluster Based on SNA: A Case Study of Foshan Lighting Cluster”. Systems Engineering, vol.35, no.1, Jan. 2017, pp. 46-57. doi.org/CNKI:SUN:GCXT.0.2017-01-007.

[2] J. Guang, S. Wu, and L. Xing, “Research on the Competitive Situation of Modern Manufacturing Industry in Beijing-Tianjin-Hebei Region from the Perspective of Complex Network”. Science & Technology Progress and Policy, vol.35, no.7, Apr. 2018, pp. 51-57, doi.org/10.6049/kjjbydc.2017090138.

[3] A. Grandori, “An organizational assessment of interfirm coordination modes”. Organization Studies, vol.18, Nov. 1997, pp. 897–925, doi.org/10.1177/017084069701800601.

[4] C. M. Beckman, and P. R. Haunschild, “Network learning: The effects of partners’ heterogeneity of experience on corporate acquisitions”. Administrative Science Quarterly, vol.47, no.1, Mar. 2002, pp.92–124, doi.org/10.2307/3094892.

[5] M. Smith, S. Gorgoni, and B. Cronin, “International production and trade in a high-tech industry: A multilevel network analysis”. Social Networks, vol.59, no.7, Oct. 2019, pp.50-60, doi.org/10.1016/j.socnet.2019.05.003.

[6] R. Tong, L. Xue, H. Wang, F. Hu, Z. Liang, J. Ding, “Building and Exploring an Enterprise Knowledge Graph for Investment Analysis”. In: Groth P. et al. (eds) The Semantic Web – ISWC 2016. ISWC 2016. Lecture Notes in Computer Science, Springer, Cham. vol. 9982, Sep. 2016, pp. 418-436, doi.org/10.1007/978-3-319-46547-0_35.

[7] Z. Xu, Y. Sheng, L. He, Y. Wang, “Summary of Knowledge Mapping Technology”. Journal of University of Electronic Science and Technology of China, vol.45, no.4, Jul. 2016, pp. 589-606, doi.org/10.3969/j.issn.1001-0548.2016.04.012.

[8] W. Tsai, and S. Ghoshal, “Social Capital and Value Creation: The Role of Intrafirm Networks”. Academy of Management Journal, vol.41, no.4, Aug. 1998, pp. 464-476, doi.org/10.5465/257085.

[9] L. Meng, Analysis of Unfair Competition in Electronic Bidding of Water Engineering Based on Complex Network. Tianjin, Tiangong University, 2018.

[10] W. Jia, X. Lou, X. Dang, “Research on the Growth of Core Enterprises in Technological Innovation Network from the Perspective of Knowledge Matching”. Chinese Journal of Management, 2018, vol.15, no.3, Mar. 2018, pp. 375-381, doi.org/10.3969/j.issn.1672-884x.2018.03.008.