Knowledge Map Analysis of CBR—Based on CiteSpace

Xian-Wu ZHANG

Wuhan University of Technology, Wuhan City, Hubei Province, China
571612950@qq.com

Keywords: CBR, Knowledge Mapping, CiteSpace, Research Hotspots.

Abstract. CBR is an important branch in the field of artificial intelligence, which is a case-based reasoning system. Rational use of CBR can effectively solve problems such as emergency decision-making, disaster management, medical diagnosis, etc., so the research on CBR has important significance. To show CBR research focus problem and digging a new research direction at the same time, CBR related literatures were analyzed by CiteSpace V software. Finally, we get the information of co-occurrence of research institutions, co-citation of authors, and co-occurrence network of key words. The results show that: 1. The institutions with the highest number of published articles are all universities; 2. The articles of Zhiwei Ni, Jianyang Li and Huiting Liu are cited more frequently; 3. The top high-frequency keywords are case-based reasoning, fault diagnosis, expert system, case retrieval, rule reasoning, etc., which are research hotspots.

Introduction

CBR is an important branch in the field of artificial intelligence. Its research was originated from the cognitive model of CBR and its framework first proposed by professor Roger Shank of Yale University in 1982 in Dynamic Memory: A Theory of Learning in Computer and People [1]. CBR is based on case reasoning system, the database recorded the successful solution to the problem before, through the case matching, then the current examples of possible solutions [2]. CBR is the problem of the past solutions in the form of case, after the new problem occurred, according to the target case prompted to search the source case, the case under the guidance of source case solution to solve the target case answer [3]. About CBR model, the process of widely used is Admodt and Plaza puts forward 4R model, in this model of CBR reasoning cycle consists of four links, including: Retrieve Reuse Revise Retain, as shown in figure 1 [2]. Due to its own advantages, CBR is widely used in emergency decision-making, disaster management, medical diagnosis and other fields. Since CBR can well simulate human cognitive process, its application scope is becoming more and more extensive [4]. Therefore, it is of great significance to use CiteSpace software to carry out visual analysis on the knowledge mapping of CBR research, which is helpful to understand and grasp the historical context of CBR research, reveal the development status and emerging trends of CBR research, and provide help for researchers engaged in this field to explore new research directions.

Citespace, as an information visualization software, has been widely used in management science and engineering, security science, information science and other fields in recent years. It focuses on the analysis of the potential knowledge contained in scientific literature, and obtains the research trends of research hotspots on specific topics and the issues in frontier hotspots [5]. In the field of safety science, cuiping Ren et al. made use of CiteSpace to analyze the research context and research status of dangerous goods transportation [6]. By virtue of CiteSpace, Li Hongxia et al. summarized the research status of global coal mine safety management and predicted the future development trend [7]. With medical malpractice as the theme, chang qing et al. combed Chinese medical malpractice literature based on CiteSpace [8].
Materials and Methods

The data of this study came from the search of CNKI database, and the keywords were Case-based Reasoning and CBR. The time node was from 2005 to March 31, 2019. After screening, a total of 1119 literatures were obtained. TXT documents were preprocessed by CiteSpace, and the previously saved TXT documents were converted into a format convenient for CiteSpace identification and analysis by using the Data function of CiteSpace software. Next, CiteSpace was officially used to analyze the data. Our goal was to obtain the following three results from the data, including co-occurrence analysis of keywords, co-occurrence analysis of authors and co-occurrence analysis of research institutions. The software version used in this study is CiteSpace V 5.3.r4.8.31.2018.

Result

CBR Literature Time Series Distribution

As shown in figure 2, since 2005, CBR related studies have shown a growing trend, and the number of published articles has increased rapidly. In 2007, the number of related research literatures increased to the first peak of 104, and then the number of literatures decreased slightly, but the overall trend remained stable. During the period from 2007 to 2013, the number of related literature publications remained at a high level, ranging from 90 to 100 every year. After that, a turning point appeared in 2013, and the number of published articles turned to decrease year by year, and maintained until 2019.
CBR Research Institutions’ Co-occurrence

It can be seen from the figure that the institutions with the highest number of publications are all universities, and the top three universities with the highest number of publications are HeFei University of Technology, Nanjing University Of Aeronautics And Astronautics and Northeastern University, with 62 papers, 41 papers and 37 papers respectively.

After that, the co-occurrence distribution of CBR literature research institutions was analyzed. In CiteSpace, the time span of published papers was set as 2005-2019, “Per slice” was set as 1, “Node Types” were selected as “Institution”, and the co-occurrence network of research institutions was obtained, as shown in figure 4. For co-occurrence visualization results, the manifestation is that the mechanism is the display node and the line is the connection between different nodes. Among them, the size of the node label display represents the frequency of its corresponding organization, and the line of different colors represents different publication time. The publication time changes from far to near, and the color of the line changes from dark to light.

According to the results, there was a cooperation between the State Key Laboratory of Resource and Environmental Information System and Shandong University of Science and Technology between 2007 and 2010. In addition, school of management of Hefei University of Technology has cooperated with school of computer science of Anhui Agricultural University and Key Laboratory...
of Process Optimization and Intelligent Decision-making of ministry of education of Hefei University of Technology. After 2016, Beijing University of Technology has conducted a lot of cooperative research on CBR issues, including the engineering research center of Beijing Key Laboratory of computing intelligence and intelligent system of Beijing Urban Rail Transit and Beijing Key Laboratory of Digital Community of Ministry of Education.

Author Co-occurrence Analysis

CiteSpace was used to analyze the literature authors, and the time span was set as 2005-2019, “Per slice” was set as 1, “Node Types” were selected as “Author”, and the Author co-cited network was obtained, as shown in figure 5. From 2005 to 2010, the articles of Zhiwei Ni, Jianyang Li and Huiting Liu were cited more frequently, and their articles were of high reference value. After 2014, the cooperative research results of Aijun Yan, Pu Wang and Chunxiao Zhang were cited more frequently, and were successively cited by Xiangdong Dai, Pengfei Ni and others.

Keywords Co-occurrence Analysis

Keywords co-occurrence analysis was carried out in CiteSpace, the time span was set as 2005-2019, “Per slice” was set as 1, “Node Types” were selected for “Keywords”, and the knowledge map of keyword co-occurrence network of 150 nodes and 431 lines was obtained. Then keyword clustering is carried out, and the results are shown in figure 6. After clustering, 7 topics are obtained, which are respectively case-based reasoning, case retrieval, fault diagnosis, ontology, genetic algorithm, knowledge representation, case representation and fusion reasoning. In order to effectively display the key nodes in the knowledge map, the top 10 high-frequency keywords are listed in table 1. The higher the word frequency of keywords is, the more important it is, and the keyword with high centrality means that it is the research focus and center.

Table 1. Table of High Frequency Keywords in CBR Literature(Top 10).

<table>
<thead>
<tr>
<th>No.</th>
<th>Keyword</th>
<th>Frequency</th>
<th>Centrality</th>
<th>No.</th>
<th>Keyword</th>
<th>Frequency</th>
<th>Centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>case-based reasoning</td>
<td>936</td>
<td>1.46</td>
<td>6</td>
<td>ontology</td>
<td>47</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>fault diagnosis,</td>
<td>113</td>
<td>0.04</td>
<td>7</td>
<td>similarity</td>
<td>33</td>
<td>0.01</td>
</tr>
<tr>
<td>3</td>
<td>expert system</td>
<td>82</td>
<td>0.07</td>
<td>8</td>
<td>rough set</td>
<td>31</td>
<td>0.01</td>
</tr>
<tr>
<td>4</td>
<td>case retrieval</td>
<td>74</td>
<td>0.07</td>
<td>9</td>
<td>artificial intelligence case</td>
<td>31</td>
<td>0.01</td>
</tr>
<tr>
<td>5</td>
<td>rule-based</td>
<td>53</td>
<td>0.06</td>
<td>10</td>
<td>case</td>
<td>31</td>
<td>0.02</td>
</tr>
</tbody>
</table>
In order to obtain the change trend of high-frequency keywords over time and dig new research directions from research hotspots, only the keywords displaying high-frequency and high centrality are selected in the knowledge map, and the results are shown in figure 7. Similarly, the darker the color of the line in the figure, the closer the age of the keywords on both ends of the line, which is a relatively recent research hotspot. As can be seen from the knowledge map, case retrieval, expert system, fault diagnosis, rule reasoning, ontology, etc. have a high centrality, and these key words are at the center. Although the "centrality" of the surrounding keywords is not high, but the frequency is high, and these keywords are relatively new, which is a recent research hotspot. For example, key words such as artificial neural network and intelligent decision support system,
genetic algorithm and neural network, expert system and case base, knowledge management and similarity calculation are hot topics in recent years. Figure 8 shows the emergence time of high-frequency keywords of different categories in detail, so that we can sort out the development of hot spots in CBR research and provide help for seeking new research directions.

Figure 8. High Frequency Keywords of CBR Literature from Different Time Periods.

Conclusion

This paper expounds the research status and emerging trends of CBR from the perspective of literature metrology, and shows the specific status quo of CBR in China. This paper elaborates on the co-occurrence network of co-authors and co-cited keywords in CBR related literature. It provides help for combing the research vein of CBR and mining the new research direction of CBR.

References


[5] Li J. Guide to CiteSpace use - li jie's blog [J].

