

Research Trends and Hotspot of Collective Intelligence Using Co-word Analysis

Jinfa Li^{1,a}, Shimeng Wang^{1,b} and Hongbing Jiang^{1,c,*}

¹School of Management Engineering, Zhengzhou University, Zhengzhou 450001

^a18806024@qq.com, ^bwsm805@qq.com, ^cjhbymx@foxmail.com

*Corresponding author

Keywords: Collective intelligence, Co-word analysis, Cluster analysis, Strategic coordinate analysis.

Abstract. Taking the relevant literatures in the field of collective intelligence as the research object, which identifies and explores the core knowledge and research hotspots of domestic collective intelligence research, predicts the future development trend of collective intelligence, and provides reference and use for reference by the research and development of collective intelligence. The paper tries to conduct co-word analysis method and the related computer applications on different matrixes taking advantages of three multivariate statistic methods—cluster analysis, multi-dimensional scaling analysis and strategic coordinate analysis. The analysis of strategic coordinate shows that 15 keywords are hot topics in the field of collective intelligence research, such as “personalized recommendation”, “knowledge management” and “virtual learning community”, and 4 keywords have been regarded as the trend of collective intelligence in the future, such as “social network”, “competitive intelligence” and “un-confirmed information transmission”, are the trend of future collective intelligence research.

1. Introduction

In 1999, Heylighen put forward the idea of developing collective mental map to realize collective problem-solving ability, and pointed out that the best metaphor for such a global intelligent network might be "global brain"[1]. The group is the main motive force that the mankind promotes the social development and progress. It is the important magic weapon for mankind to deal with the nature and the social challenge, in which the wisdom of the group is of the utmost importance[2].

Domestic and foreign academic have carried out a series of researches on collective intelligence. Foreign scholar Salminen accumulated literature information related to collective intelligence by Keyword Retrieval Method, and divided the abstract concept of collective intelligence into three levels: micro level, level and emergent level by rough analysis method[3]. Li Zili and colleagues collected relevant literature from the Science Citation Index from 2000 to 2015, and the research focus and research trend of Collective Intelligence was quantitatively evaluated using bibliometric methods, such as co citation analysis, semantic clustering, and frequent occurrence detection of keywords[4]. Based on the literature of collective intelligence from 2012 to 2015, Grimon and Meza et al. identify three main areas of research for the application of collective intelligence by Content analysis : learning, technology and decision making[5]. To summarize, it can be seen that foreign scholars have begun to pay attention to the study of the subject structure of collective intelligence.

In China, Dai Yang and Zhou Lei inspected the cognitive process of collective intelligence in foreign countries. By using the method of Bibliometrics, five major topics of foreign research on collective Intelligence were summarized[2]. Liu Haixin and Liu Renjie combed the research of collective intelligence both in China and foreign country, and constructed the theoretical framework of collective intelligence[6]. At present, due to the immature research and development of collective intelligence in China, there are less studies on development and more studies on application.

To sum up, collective intelligence is a new innovation mode in the era of knowledge economy and Web 2.0[7]. Most of the existing literature is on the application of collective intelligence, and few studies are on the development of the field of collective intelligence. Moreover, the research on the

subject structure of collective intelligence is mainly concentrated in foreign countries, and the research hotspots and trends of localization in the field of domestic collective intelligence are focused on. There are few literatures on potential analysis. Therefore, this paper puts forward two research questions: ① What are the research hotspots of collective intelligence? ② What is the future development direction of collective intelligence? By answering these questions, we can not only improve the efficiency of user-centered innovation, but also promote users' participation in the whole process of product development, product design, collaborative innovation, decision-making, and constantly improve the quality of collective intelligence decision-making in various fields.

In order to answer the two research questions raised in this paper, this paper will use bibliometric method to analyze the Co-word of keywords in the research of collective intelligence in China, showing the research hotspots and future research trends in the field of collective intelligence in China, and effectively complement the research of collective intelligence literature.

2. Methodology

2.1 Research methods

As a qualitative and quantitative analysis method, co-word analysis is one of the important methods advocated by complex scientific management[8]. In this study, the keyword matrix and difference matrix are constructed, which are used in quantitative analysis and qualitative analysis. The keyword matrix is mainly analyzed by word frequency analysis, clustering analysis and multi-dimensional scale analysis. Finally, the social network analysis tool UCINET 6.0 is used to visualize high-frequency keywords. More scientifically demonstrates the current research hotspots in the field of collective intelligence. The combination of quantitative analysis and qualitative analysis also makes the research process more rigorous and the conclusions more scientific and reasonable. The research tools used in this study mainly include SATI3.2, SPSS 24.0 and UCINET 6.0.

2.2 Data processing

With the full-text database of China HowNet as the data source and the keywords of "collective intelligence" or "wisdom of crowd", this paper retrieves 1330 results, excluding the conference reports and related documents, including more articles about primary and secondary school teachers preparing lessons (because they are not the research focus of this study) and incomplete information. 157 valid papers were retained as research samples, and the title information of 157 papers was derived from CNKI and saved in EndNote format for subsequent analysis using relevant research tools.

Firstly, the title information of selected literature derived from China HowNet is imported into SATI 3.2 for format conversion, keyword extraction and statistical word frequency operations. A total of 439 original keywords are obtained. However, because the same concept is expressed differently in different literature, the use of such non-standardized keywords will interfere with the results of data analysis. Therefore, before choosing high-frequency words and subsequent analysis, we should first standardize the extracted keywords. After a series of standardization treatment, 360 keywords were finally identified for subsequent analysis. According to this, the threshold of high-frequency words is 26. The cumulative number of occurrences of these 26 keywords is 267. The cumulative word frequency accounts for 39.2% of the total word frequency. Among them, the top five are collective intelligence (70), wisdom of crowd (68), social media (19), Web 2.0 (14), crowdsourcing (12).

2.3 Constructing Co-occurrence Matrix and Difference Matrix

SATI 3.2 software is used to count the word frequency of high frequency keywords, then two pairs of pairs are paired. The frequency of their co-occurrence in the same literature is counted, and the high frequency keyword co-occurrence matrix is generated, as shown in Table 1. At the same time, SATI 3.2 is used to construct the required dissimilarity matrix, as shown in Table 2.

Table 1. High Frequency Key Co-occurrence Matrix (Part)

	Wisdom of Crowd	Collective Intelligence	Social Media	web2.0	Crowdsourcing
Wisdom of Crowd	70	0	8	6	5
Collective Intelligence	0	68	9	5	5
Social Media	8	9	19	1	0
web2.0	6	5	1	14	0
Crowdsourcing	5	5	0	0	12

Table 2. Dissimilarity matrix (part)

	Wisdom of Crowd	Collective Intelligence	Social Media	web2.0	Crowdsourcing
Wisdom of Crowd	0	1	0.952	0.963	0.970
Collective Intelligence	1	0	0.937	0.974	0.969
Social Media	0.952	0.937	0	0.996	1
web2.0	0.963	0.974	0.996	0	1
Crowdsourcing	0.970	0.969	1	1	0

3. Analysis of Research Hotspots in Collective Intelligence in China

3.1 Cluster Analysis

The 26 x 26 keyword dissimilarity matrix is imported into SPSS 24.0 for clustering analysis. The clustering analysis method chooses systematic clustering, and uses Ward method as clustering method, thus outputting the clustering analysis tree as shown in Figure 1.

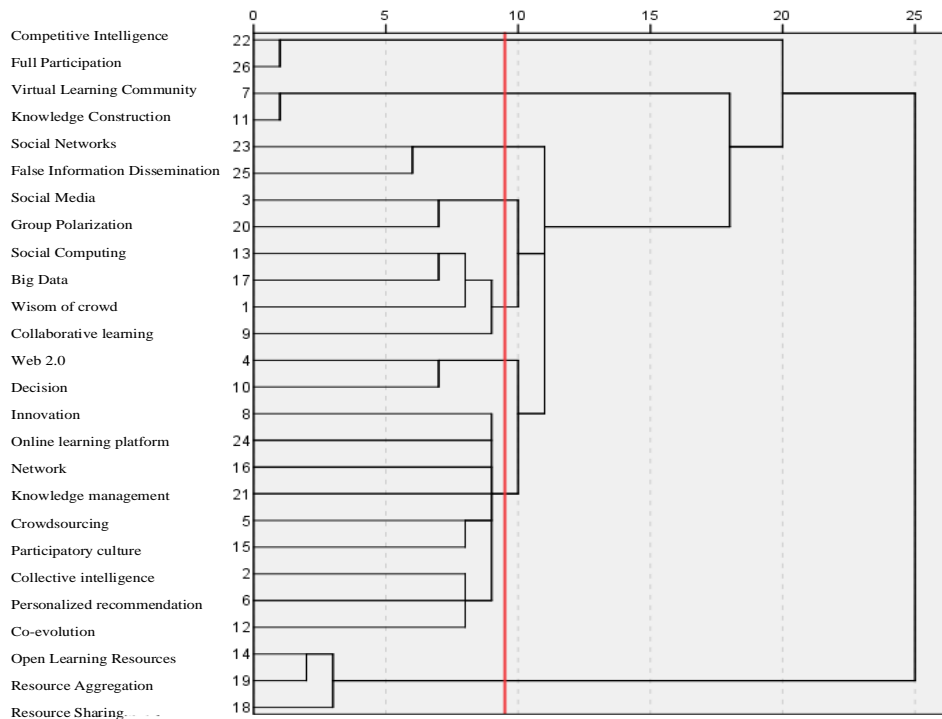


Fig. 1. High Frequency Keyword Clustering Analysis

From the tree graph of cluster analysis, it can be seen that the field of collective intelligence in China is divided into eight categories, as shown in Table 3.

Table 3. Cluster analysis.

Cluster	Number of keywords in a cluster	Keyword
1	4	Wisdom of crowd, Collaborative Learning, Social Computing, Big Data
2	9	Collective intelligence, crowdsourcing, personalized recommendation, innovation, co-evolution, online learning platform, participatory culture, network, knowledge management
3	2	Social Media, Group Polarization
4	2	Web 2.0, Decision Making
5	2	Virtual Learning Community, Knowledge Construction
6	2	Social Networks, False Information Dissemination
7	3	Open Learning Resources, Resource Sharing, Resource Aggregation
8	2	Competitive Intelligence, Full Participation

3.2 Multidimensional Scale Analysis

The obtained high-frequency word dissimilarity matrix is imported into SPSS 24.0 for multi-dimensional scale analysis. The result of multi-dimensional scale analysis is to show the similarity between keyword data on two-dimension. From the two-dimensional coordinate map, it can be seen that the 26 keywords do not form a relatively centralized distribution of several regions, but merge and distribute together without obvious boundaries, forming several large clusters. This also shows that the research in the field of collective intelligence in China is not perfect enough, and can not form a clear knowledge structure, but also shows the future of research in the related fields of collective intelligence. There is more room for development.

On the other hand, we find that the three groups of keywords "knowledge construction" and "virtual learning community", "resource aggregation", "open learning resources" and "resource sharing", "competitive intelligence" and "full participation" are slightly more distant from other keywords, which is the same as the results of cluster analysis. After a series of screening and adjustment, the results of multidimensional scale analysis are shown in Table 4.

Table 4. Clustering analysis results of high frequency keywords (adjusted)

Cluster	Number of keywords in a cluster	Keyword
1	2	Wisdom of crowd, Web 2.0, Crowdsourcing, Network
2	3	Collective Intelligence, Network Learning Platform, Co-Evolution, Knowledge Management
3	2	Social media, personalized recommendation, innovation, collaborative learning, decision-making, social computing, participatory culture, big data, group polarization, social networks, false information dissemination
4	4	Virtual Learning Community and Knowledge Construction
5	4	Open Learning Resources, Resource Sharing, Resource Aggregation
6	11	Competitive Intelligence, Full Participation

3.3 Strategic Coordinate Analysis

The co-occurrence matrix of high-frequency keywords is imported into UCINET 6.0, which is a social network analysis tool. Netdraw, an internal integration tool of the software, is used to draw a high-frequency keyword network map related to collective intelligence.

From the graph, we can see that the key words of collective intelligence, collective intelligence, social media, knowledge management, competitive intelligence, social network are in the middle of the knowledge map.

In order to better grasp the research hotspots and trends of collective intelligence, based on the above clustering analysis, multi-dimensional scale analysis and social network analysis results, combined with the relevant literature in the field of collective intelligence, the high-frequency keywords in the field of collective intelligence are roughly divided into eight groups for strategic coordinate analysis.

Density is based on the sum of the frequencies of keywords in the cluster, and then the average value is taken. Centro-centricity is calculated on the basis of the sum of the occurrence frequency of keywords in one group and those in other groups.

Strategic coordinate maps are drawn according to centripetal degree and density, where X axis represents density, Y axis represents centripetal degree, the mean value of origin density and centripetal degree of coordinates. Strategic coordinate map divides the subject area of two-dimensional space into four quadrants to describe the current status of research in the field of collective intelligence, as shown in Figure 2.

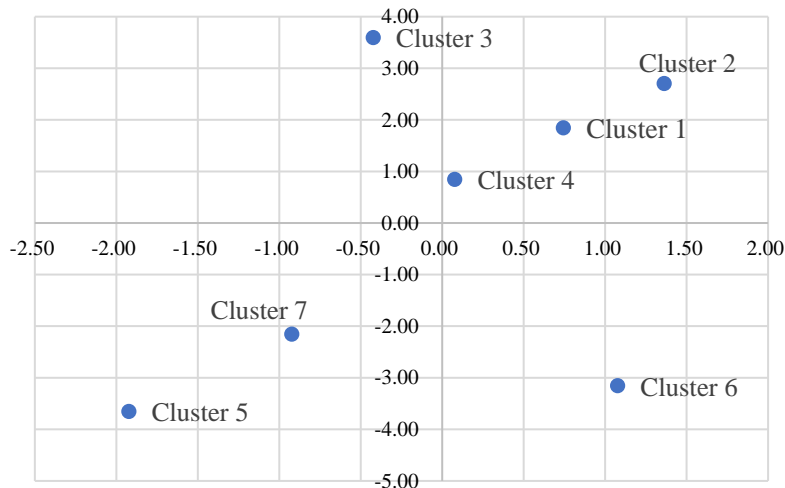


Fig. 2. Strategic coordinate map of 7 fields in Collective Intelligence.

From Figure 2, we can see that the first quadrant contains cluster 1, 2, 4, which is based on collective intelligence. It mainly includes key words such as collective intelligence, personalized recommendation, collaborative learning, social computing, co-evolution, crowdsourcing, network, knowledge management, etc. This kind of topic has a close internal connection, and it also represents research with other group keywords. The research topic has a high degree of relevance. It is an active topic in the study of collective intelligence in China. It is a hot topic in the whole subject area, which is closely related to the research topics in other groups.

The second quadrant consists of group 3, which is group decision-making based on social media. It mainly includes key words such as social media, Web 2.0, group polarization, decision-making and so on. The research topics represented by such topics are closely related to other research topics and have been extensively studied. They are active topics in this field. However, the internal structure of such research topics is relatively loose and can not be well self-contained. Due to the unstable structure, it is easy to be decomposed in the subsequent development process, thus evolving into other related group content [9]. Therefore, although this group has great potential for development, its internal structure is unstable.

The third quadrant consists of 5,7 clusters, which are based on the participation of all members in social networks and information dissemination, including key words such as social networks, false information dissemination, competitive intelligence and full participation. Such topics are not closely related internally and loosely structured, and are not closely related to other clusters. The research is still immature and belongs to marginal researchers with relatively little attention in the field of research. However, it also shows that the attention of this study has great room for development.

The fourth quadrant contains cluster 6, which is the key words of resource aggregation and resource sharing among groups, including open learning resources, resource sharing and resource aggregation. This kind of topic has a high density but a low centrality, which indicates that the internal connection of this kind of research topic is relatively close, and has formed a certain research scale, but the connection with other research topics is not close and the structure is loose. In the follow-up research and development, once such topics develop to a certain stage, they may disappear if they do not get enough motivation.

4. Analysis on the Development Trend of Collective intelligence in China

Combining cluster analysis, multi-dimensional scale analysis, social network analysis and strategic coordinate map analysis, we can see that the current academic research on collective intelligence is still in its infancy and has not formed a systematic discipline system, but it also shows that there is still a great space for the development of collective intelligence in the data age. Therefore, combined with previous research on collective intelligence and our analysis results, the following seven aspects are mainly manifested.

(1) Personalized recommendation based on big data.

For e-commerce enterprises, personalized recommendation can not only help retailers improve product sales, but also create more cross-selling opportunities. At the same time, for consumers, in the face of the same goods, consumers can choose the goods according to their actual situation, and have greater autonomy.

(2) Collective intelligence decision-making based on Web 2.0.

Because of the complexity of decision-making problems, the traditional decision-making model has not been able to solve the complex decision-making problems of enterprises[10]. With the development of network, collective intelligence has gradually become the focus of public attention. When complex decision-making problems can not be solved by traditional decision-making methods, more solutions will be created by relying on collective intelligence because of the differences in learning background, life experience and education level of groups.

(3) Network public opinion dissemination based on social network.

As the carrier of network public opinion dissemination, the public is the extension of actual public opinion. The government, as the administrator of network public opinion dissemination, should control and guide the generation and spread of network public opinion, and attach importance to groups. The opinion leaders in the body should ensure that public opinion is on the right track of development and build a green and healthy cyberspace.

(4) The construction of learning platform based on wisdom education.

For example, MOOC Online Education Platform pays attention to the construction of community, including not only online discussion but also offline communication. It relies on the network community for communication and interaction to continuously improve students' interest and motivation [11]. But for the answer of open questions, the machine is difficult to score according to specific criteria, so collective intelligence is particularly important for such open and flexible questions.

(5) Data mining and knowledge discovery based on collective intelligence.

How to find the data information we need from a large amount of data and transform it into valuable information and knowledge is the most difficult problem we face. Mass Participation Collective intelligence provides a new idea for the study of new data mining methods [12].

(6) Construction of open source innovation projects based on co-evolution.

Open source innovation emphasizes that all kinds of people at different levels can freely and freely join the open source innovation community, play their professional skills, and modify the code of open source software projects. Under this kind of innovation mode, the innovation results are more in line with the public's own research interests and users' needs.

(7) Competitive Intelligence Acquisition Based on full participation.

Web 2.0 not only enhances the ability, consciousness and opportunity of individual information intelligence, but also promotes the top-down cultural transformation of enterprises and universities, lays a foundation for all personnel to participate in competitive intelligence, and further improves the level of enterprise competitive intelligence based on collective intelligence[13]. Thus, when enterprises use collective intelligence to obtain information and intelligence, they can not only reduce the cost of seeking professional personnel, but also improve the quality of information.

5. Conclusion

In this paper, we use Co-word analysis, cluster analysis, strategic coordinate analysis and visual network analysis to explore the status quo of research in the field of collective intelligence in China, and reveal the research hotspots and key points of collective intelligence in decision-making, knowledge mining, network public opinion dissemination and other fields.

In the process of Co-word analysis, 26 keywords are taken as the object of this study. The keyword matrix is clustered and multi-dimensional scaling is analyzed by SPSS software. At the same time, the results of strategic coordinate analysis are presented in a more intuitive way. The results show that the application fields of collective intelligence based on Internet can be divided into seven aspects: personalized recommendation, collective intelligence decision-making, network public opinion dissemination, data mining, open source innovation, Competitive Intelligence Acquisition and intelligent learning platform.

References

- [1] F. Heylighen. Collective Intelligence and its Implementation on the Web: algorithms to develop a collective mental map. *Computational & Mathematical Organization Theory*, vol.5(3), pp.253-280, 1999.
- [2] Y. Dai, L. Zhou. Literature Review on Collective Intelligence in Foreign Countries. *Documentation, Information & Knowledge*, vol.2, pp.120-127, 2014.
- [3] J. Salminen. Collective Intelligence in Humans: A Literature Review. *Computer Science*, 2012.
- [4] Z. Li, L. Zeng, H. Zhong, et al. Research Hotspots and Trends in Collective Intelligence: From 2000 to 2015. Literature Review. vol.15(6), pp.24-35.2016.
- [5] F. Grimon, J. Meza, V. C. Mónica, et al. Research and Trends in the Studies of Collective Intelligence from 2012 to 2015[M]//E-Learning, E-Education, and Online Training. *Springer International Publishing*, 2017.
- [6] H. X. Liu, R. J. Liu. The Connotation of Collective Wisdom and Its Research Summary. *Chinese Journal of Management*, vol.10(2), pp.305-312,2013.
- [7] T. Ge, Z. J. Xia, Y. Zhai. An Analysis of the Factors Influencing the Dissemination of False Information in Social Media by Group Wisdom. *Journal of Intelligence*, vol.7, pp.148-152,2017.
- [8] Q. Zhang, X. S. Xu. Qualitative and Quantitative Analysis Method-Co-word Analysis. *Technology Economics*, vol.29(6), pp.20-24,2010.
- [9] W. J. Zhong, J. Li. Study on Co-word Analysis (2) - Cluster Analysis. *Journal of Information*, vol.6, pp.141-143,2008.
- [10]H. Su, X. X. Hu. Decision-making Model of Complex Problems Based on Group Intelligence. *Chinese Journal of Management Science*, vol.20(2), pp.783-789,2012.
- [11]H. S. Guan, Q. S. Gao. University Curriculum Alliance under MOOC Wave. *Higher Education of Sciences*, vol.1, pp.44-52,2014.
- [12]J. X. Cai, D. W. Xu. Application of Massive Data Mining Technology Based on Cloud Computing Platform in Plastic E-commerce Platform. *Guangdong Science & Technology*, vol.20 (16) , pp.58-59,2011.
- [13]X. P. Song, Y. Yang, B. Z. Li. Research progress of full participation in competitive intelligence under Web 2.0 environment. *Journal of Intelligence*, vol.34(11), pp.69-73,2015.