International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 9002-9012

https,//doi.org/10.70517/ijhsa463771

Data Analysis of Hot Trends in Journalism and Communication Based on Knowledge Graph

Bingqi Yin[1] and Yanxue Wang[2,3,*]
[1] School of Drama, Film and Media, Dalian Art College, Dalian, Liaoning, 116699, China
[2] School of Communication, Baicheng Normal University, Baicheng 137000, Jilin, China
[3] Center For Research in Media and Communication, Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia, UKM Bangi 43600, Selangor, Malaysia
Corresponding authors: (e-mail: 19904366766@163.com).

Abstract In the process of gradually deepening the study of journalism and communication in China, scholars of journalism and communication have made great theoretical improvement and good development in their continuous study. In order to explore the research hotspots in the field of journalism and communication in recent years, this paper, based on the bibliometric analysis method, collected and counted the data of the articles in the top ten journals in journalism and communication, classified them systematically, and used CiteSpace software to draw the relevant knowledge graph. This paper made a detailed analysis of the research hotspots of journalism and communication by using the knowledge graph through the selection of ten major journals of journalism and communication. According to the data, the frequency of media convergence, social media, short video, international communication and digital news was 55, 48, 39, 36 and 24 respectively, and the centrality was 0.51, 0.40, 0.31, 0.22 and 0.21 respectively. The study found that from December 2021 to November 2022, the five research hotspots of the ten major journals of journalism and communication mainly included media convergence, social media, short video, international communication and digital news. In the future research in the field of journalism and communication, attention should be paid to the intersection and integration of communication, sociology and other related disciplines to increase the proportion of empirical research in the field of journalism and communication.

Index Terms Journalism and Communication, Knowledge Graph, Spectral Clustering Algorithm, Research Hotspot

I. Introduction

This paper takes the research documents in the ten major journals in journalism and communication as the research object, and makes a visual analysis of the research hotspots of China's journalism and communication with the help of CiteSpace. Through the analysis of the annual trend of publication, core authors, high-yield institutions and cited frequency of China's journalism and communication research, the current situation of China's journalism and communication research is sorted and summarized from two aspects of discipline characteristics and hot trends. Based on the in-depth analysis of the research hotspots of China's journalism and communication, this paper reveals the research achievements and development direction in this field, and provides certain reference for further research in this field.

With the continuous development of society, the research of journalism and communication has gradually increased. Yang Jie studied the application of big data analysis and visualization technology in journalism and communication. Big data can make communication science more and more scientific, and even social science more and more scientific. Content visualization is represented by content graphics [1]. Lewis Seth C. believed that news research, especially the research focusing on automated news, had a lot to learn from human-computer communication. This was an emerging conceptual framework and experience-based research field, which was formed in response to more and more technologies [2]. Lin Yistie analyzed that the network was used to spread news, which enriched the diversity of journalism and communication. In network communication, the market environment was highly competitive, and profit drove the change of truth. Moreover, due to the lack of control consciousness in the network communication of news, fake news production may occur in every link of network communication [3]. Greussing Esther studied the visualization of individuals' participation in static and interactive animation data embedded in online science news, and the results showed that interaction and animation were the most attractive to participants [4]. Usher Nikki believed that in order to better understand the challenges faced by today's professional journalism, it must be placed at the center of academic and industry analysis [5]. Although these studies have promoted journalism and communication to some extent, they have not been combined with the actual situation.

At the same time, the knowledge graph has gradually attracted widespread attention from the academic community. Wang Guofeng used knowledge graph and CiteSpace to conduct bibliometric analysis of news discourse analysis, and outlined its scientific landscape according to the English journal articles in the Scopus database from 1988 to 2020 [6]. Carlson Matt believed that algorithmic judgment should be considered as a professional judgment different from that of journalists, and pose a fundamental challenge to news judgment based on dual beliefs, which promoted the depth of knowledge graph in practice [7]. Kruikemeier Sanne's research used eye movement tracking to check the news mode (that is, whether the news was presented in a printed newspaper or on a news website), and to what extent it affects the visual attention to the news, thus affecting knowledge, which had reference significance for the application of knowledge graph [8]. Although these research methods are innovative, a large number of experimental data are needed to prove the reliability of the methods.

This paper first analyzed the determination and analysis methods of research hotspots in the field of journalism and communication, including the development of knowledge graph, knowledge graph based on co-word clustering, hot spot formation and hot spot analysis. Secondly, this paper processed the data of the top ten journals by selecting the popularity ranking of journals by the people engaged in journalism and communication. Finally, this paper made a detailed analysis of the research hotspots of journalism and communication.

II. Methods for Determining Hot Spots in the Field of Journalism and Communication

II. A. Development of Knowledge Graph

In 1955, a researcher published the citation index, which laid the foundation for citation analysis. In 1964, several researchers manually drew the industry knowledge graph based on citation analysis. The concept of "knowledge graph" has not been put forward. However, in fact, the word "knowledge graph" has been widely used in practice with the efforts of researchers.

In 1973, the concept of "co-citation" was introduced, and the intensity was used to measure the degree of co-citation between articles. This idea has important practical significance for the research of knowledge graph. Compared with citation analysis, it can better describe the kinship of the knowledge graph and the topological structure of the graph.

Since the 1990s, with the rapid development of visualization technology, the relationship between knowledge graph and visualization has become more and more close, and has gradually developed into a new research direction in which information science, applied mathematics, computer science and other multi-disciplines integrate and promote each other.

Since 1997, a number of scholars from the United States and Leiden, the Netherlands, have explored a new approach to the knowledge graph analysis of mathematical science based on the co-cited analysis, using the method of multivariate statistics and computer graphics technology. Since then, the application of knowledge graph has become more and more extensive and mature, as shown in Figure 1.

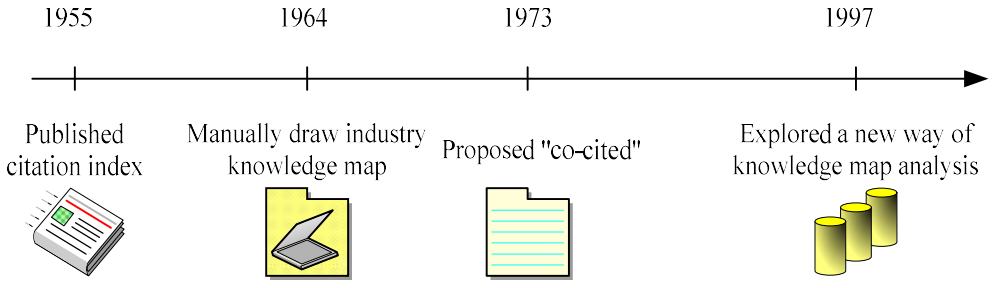


Figure 1: Development of knowledge graph

In 2003, many scholars used the view of "time series" for the first time, presenting the research frontier of a discipline in the form of "time series" in a "time series" way.

CiteSpace was led by researchers from the School of Information Engineering of Drexel University in the United States in 2004. It can analyze and visualize the core elements of a discipline, thus realizing the "panorama" display of a discipline's knowledge. It is one of the best knowledge graph construction tools at present. In the same year, a scholar first published an article in the Journal of the Academy of Sciences in the United States, explaining the working principle and role of CiteSpace. In 2006, a researcher quoted the keyword burst detection algorithm based on word frequency proposed in 2002 in another article, added his own understanding of the research frontier, and embedded it into CiteSpace software to facilitate the analysis of the frontier of field research. At the same time, a

paper was published in 2010, introducing an automatic extraction algorithm for classification using Term Frequency-Inverse Document Frequency (TF-IDF). American scholars put forward a new calculation method-h index, which is used to measure the influence of a node in the knowledge base. The original definition of the h index clearly pointed out that the academic journal groups were ranked according to their cited times from high to low.

The above researchers have conducted in-depth research on the knowledge graph and achieved certain results. The visualization results are presented in various forms to make the knowledge graph more clear, intuitive and readable, and can mine more information from it, providing valuable reference for the theoretical and empirical analysis of this paper.

II. B.Knowledge Graph Based on Co-word Clustering

To understand the central idea of a text, the core words of the text need to be understood. The core words in a paragraph are words with high frequency. This “frequency” can be understood in two aspects, as shown in Figure 2.

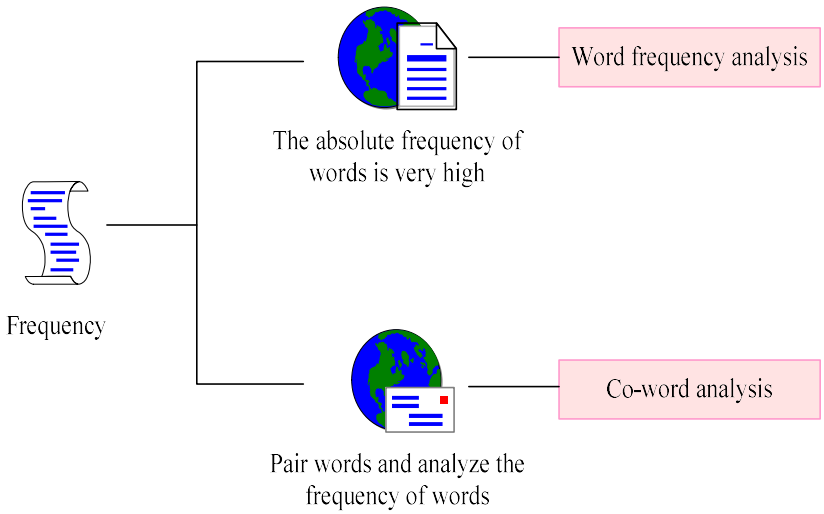


Figure 2: Understanding of “frequency”

It can be seen from Figure 2 that the absolute frequency of words is very high, which is consistent with the calculation method of “word frequency analysis”. The second is based on the analysis of word frequency, matching words in pairs, and analyzing the frequency of words, which corresponds to the “co-word analysis method”.

Co-word analysis means that on the basis of word frequency analysis, all subject words are grouped in pairs to form several word pairs, and then the frequency of a group of word pairs in the same document is calculated [9], [10].

In this paper, the co-word matrix can be obtained by using the co-word analysis method. The numerical value in the matrix represents the frequency of each pair of words in the same text [11], [12]. Therefore, the common word matrix is a completely symmetric matrix.

Social network analysis is a method based on sociological theory. Its theoretical basis is to reflect various things in people’s daily life, focusing on the relevance between the two objects and their relative contribution to the overall. Its goal is to describe the relationship between entities (nodes) in the network and further analyze the relationship between them.

In the co-word network, the nodes are words, and the connections are the connections between words, which are transformed into the relationship matrix, that is, the connection between words obtained from the co-word analysis. Because of its symmetry, the network diagram that can be transformed into each other is the undirected diagram [13], [14].

As early as 1979, people introduced the concept of centrality into social networks and applied it to the research of social networks. In this paper, the concept of intermediary centrality is mainly used. In an undirected graph theory, the intermediary centrality is a measure of the media role played by a node, which can directly reflect the importance and influence of the node. The formula is as follows:

$$V_Q(M_o) = \frac{\sum g_{kl}(m_o)}{g_{kl}} \tag{1}$$

g_{kl} is the number of shortest paths from node k to node l. $g_{kl}(m_o)$ represents the number of paths through node o in all the shortest paths from node k to node l.

After obtaining the co-word network, this paper classifies similar words into several categories by clustering method, and generates a knowledge graph, which can make the map clearer and facilitate finding the hot spots and core of field research [15], [16].

This paper adopts a spectral clustering algorithm, that is, by dividing a weighted undirected graph into two or more subgraphs, and making the subgraphs as similar as possible in their interior. The difference between subgraphs should be as large as possible, and the "distance" should be as large as possible. The steps of this algorithm are illustrated in Figure 3.

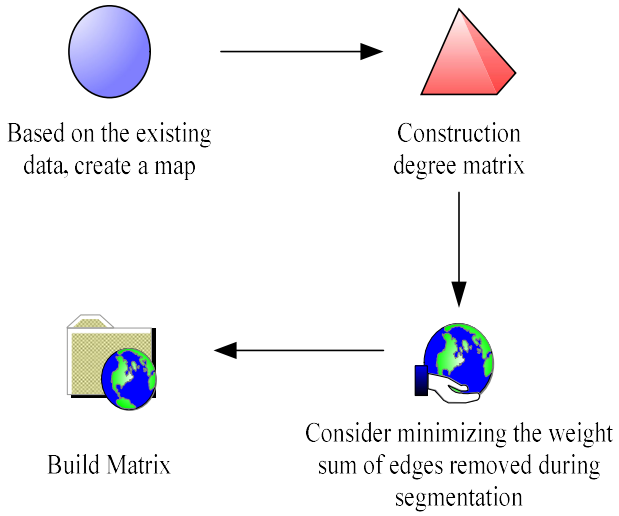


Figure 3: Specific steps of spectral clustering algorithm

The first step is to create a graph based on the existing data. Each node on the graph represents the data. In a graph, all nodes are connected together, and the weight of each edge of the graph is measured by the similarity of these data. This figure and adjacent matrix can be converted to each other:

Here, the similarity is defined by Gaussian kernel function:

$$S(x_o, x_k) = e^{-\frac{\|x_o - x_k\|^2}{2\delta^2}} \tag{2}$$

Among them, x_o and x_k are the original eigenvectors of the two samples. By analyzing the correlation coefficients of the two samples, the correlation coefficients of the two samples are obtained, and the adjacent matrix W is obtained.

The second step is to construct the degree matrix D. According to the definition of Laplace matrix, a similar matrix Z is constructed, namely:

$$Z = D - W \tag{3}$$

The third step is that the clustering purpose of this algorithm is to combine the weighted sum of several edges removed in the blocking process with the block equalization (to avoid one or two points forming a class), and express it as:

$$Cut(A_1, A_2, ..., A_l) = \frac{1}{2}\sum_{o=1}^{l} \frac{W(A_o, \bar{A}_o)}{|A_o|} \tag{4}$$

It can be seen that this formula can well minimize the similarity between different categories. In addition, this formula also reflects the "penalty" principle, that is, its value is inversely proportional to the number of vertices in a certain category. In other words, the smaller the division of a class, the greater its value. Therefore, on this basis,

this paper proposes to divide into larger classes as far as possible, rather than small classes with only one or two points. At this point, only the minimum problem of this formula can be solved to obtain the best segmentation effect.

The forth step is to arrange all the l eigenvectors into a matrix, thus forming a $M*l$ matrix. Then, each row in the matrix is regarded as a vector in an l-dimension space, that is, every sampling is an l-dimension. Then, each row in the matrix is regarded as a point, and then all the points are gathered together using the K-means algorithm. After such processing, the class of each row in the clustering result is the class of M sample data nodes in the original graph.

After clustering, it is also necessary to determine a cluster label for each category in order to summarize the main idea of this category to the maximum extent. In this paper, TF-IDF method is used to select classification marks. TF-IDF is a commonly used text mining method based on weighting, which is used to evaluate the importance of a topic in a specific corpus, and also to classify and extract specific topics [17], [18].

The main idea of the algorithm is that the importance of a keyword is positively related to its frequency in an article, and inversely related to its frequency in an article.

The TF-IDF method is applied to document clustering to find out the words with the largest TF-IDF in each document cluster, and use them as the marks of the document cluster to facilitate classification [19].

II. C.Hot Spot Formation and Hot Spot Retrieval

(1) Hot spot formation

The knowledge graph obtained by the above methods contains a large number of cluster labels. These keywords are hot topics that the article is looking for. The knowledge graph also shows the main high and middle level vocabulary contained in each category, and shows the changes of each category vocabulary over time (in years), and gives the activity years of each category.

Nowadays, the world is in a period of rapid development, and countries are more and more closely connected in various aspects such as culture. In academic circles, the study of a certain discipline is no longer limited to a specific country, but more and more stay in a certain discipline. The research results of various countries are not limited to the citation of individual research results, but more from the perspective of internationalization, absorbing the relevant theories of other countries. From the mutual reference of countries in the literature, the cooperation of countries can be clearly recognized, and then the current development of this discipline among regions is analyzed. Therefore, it is necessary to pay attention to the hot spots in the field and provide guidance for future cooperation between countries in relevant fields.

(2) Hot spot retrieval

In terms of the interpretation of the hot category, this paper explains and analyzes the meaning of the name of the hot category and the evolution of the vocabulary of the hot category. In view of the research progress in research hotspots in various countries, this paper plans to establish a complete set of retrieval database on CiteSpace software and connect to the database of the database through hyperlinks.

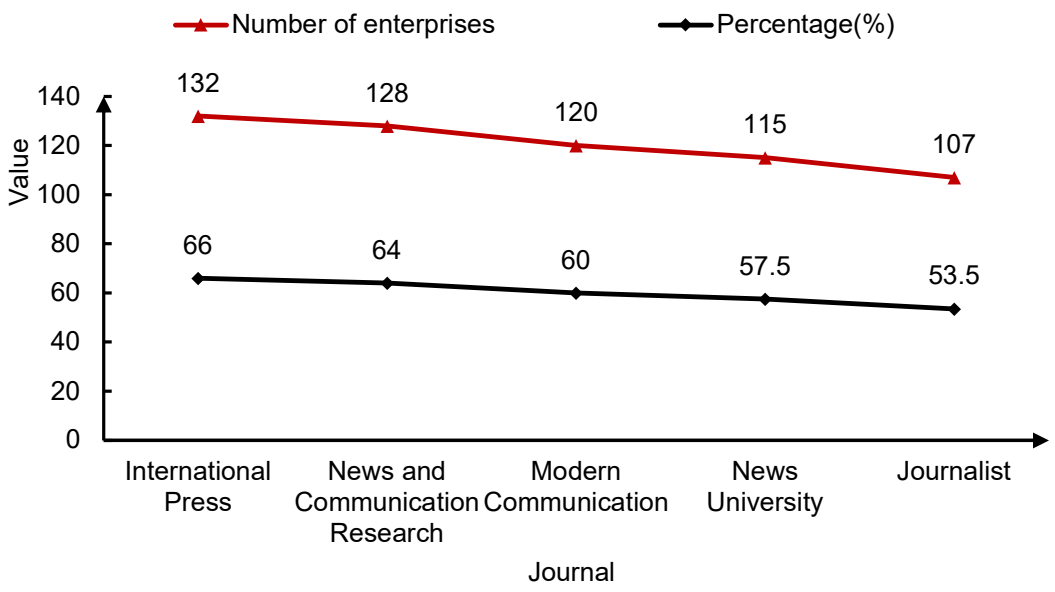
The co-citation analysis is based on the premise of co-citation of the two target documents. In a sense, it can be judged that there is a certain relationship between the two target documents. The “relationship” here refers to the close academic background and research content between the two target documents. The strength of “relationship” is defined by the frequency of common reference (frequency of common reference). The co-citation relationship between a group of target documents can be transformed through the citation network, and the “relationship” in the citation network is determined by the connotation of “relationship”.

An important indicator to measure the influence of a country in a certain discipline is its literature citation rate. On this basis, this paper introduces the h index proposed by an American scholar in 2005, which is related to the citation frequency of the article. The initial definition of the h index clearly shows that the academic journal group is ranked according to the cited frequency from high to low.

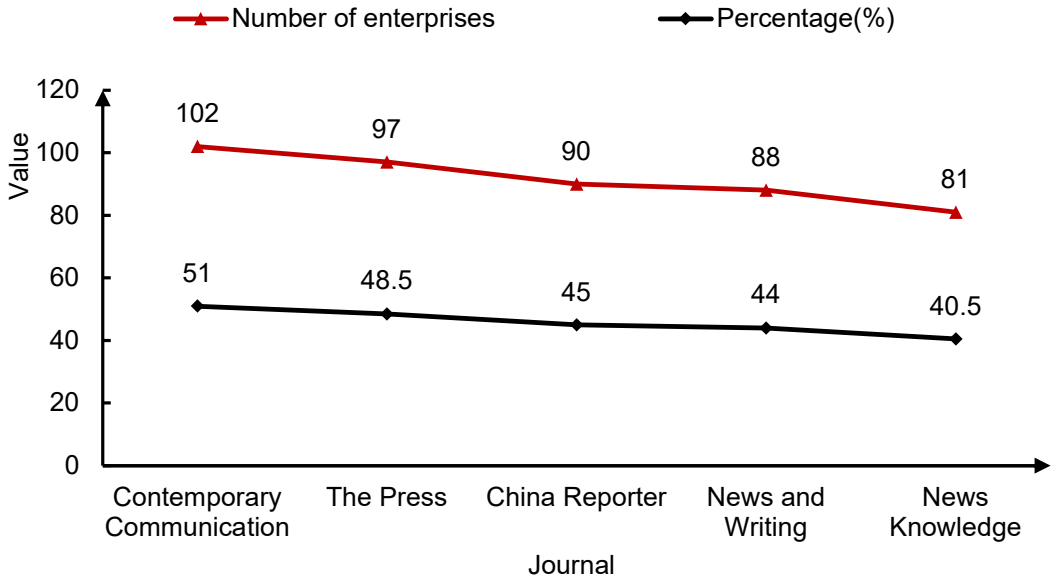
III. Experiments on the Hot Spots and Trends of Journalism and Communication

III. A. Data Source and Processing

In the early stage of the experiment, this paper selected 200 people engaged in journalism and communication, and selected ten most popular journals from many journals. Figure 4 shows the number and percentage of journals selected by the top 10 in popularity among 200 people.



(a) Popular 1-5 journals



(b) Popular 6-10 journals

Figure 4: Top 10 popular journals (multiple choices)

It can be seen from Figure 4 (a) that the number of people who choose International Press was 132, accounting for 66.00%. The number of people who chose News and Communication Research was 128, accounting for 64.00%. The number of people who chose Modern Communication was 120, accounting for 60.00%. The number of people who chose News University was 115, accounting for 57.50%. The number of people who chose Journalist was 107, accounting for 53.50%.

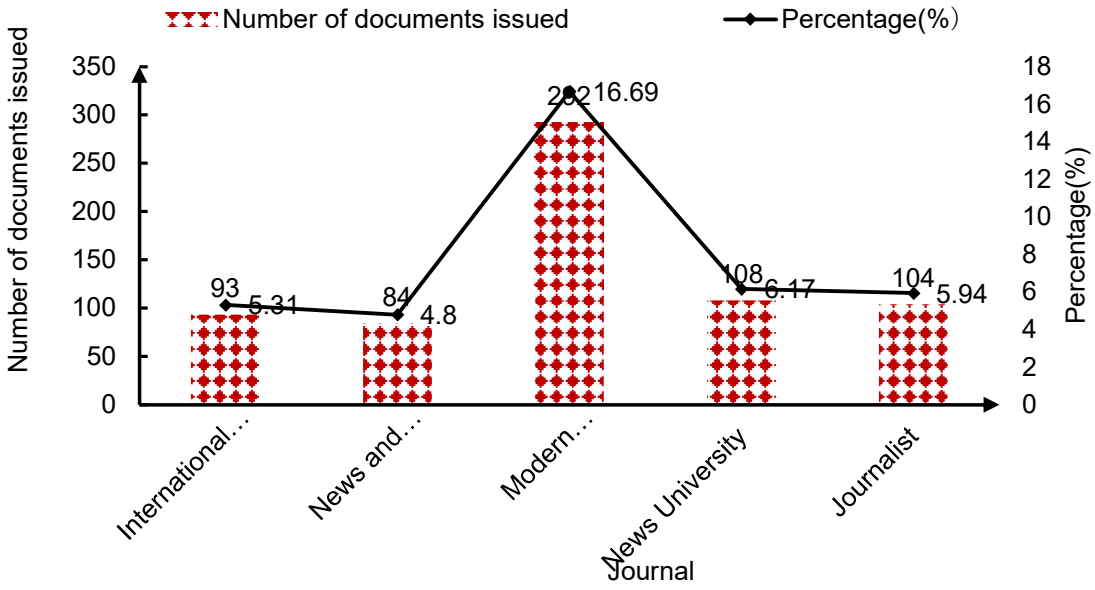
It can be seen from Figure 4 (b) that 102 people chose Contemporary Communication, accounting for 51.00%. The number of people who chose The Press was 97, accounting for 48.50%. The number of people who chose China Reporter was 90, accounting for 45.00%. 88 people chose News and Writing, accounting for 44.00%. 81 people chose News Knowledge, accounting for 40.50%.

Based on the above analysis, this paper selected all the articles from 2021.12-2022.11 of the ten major news and communication journals, including International Press, News and Communication Research, Modern

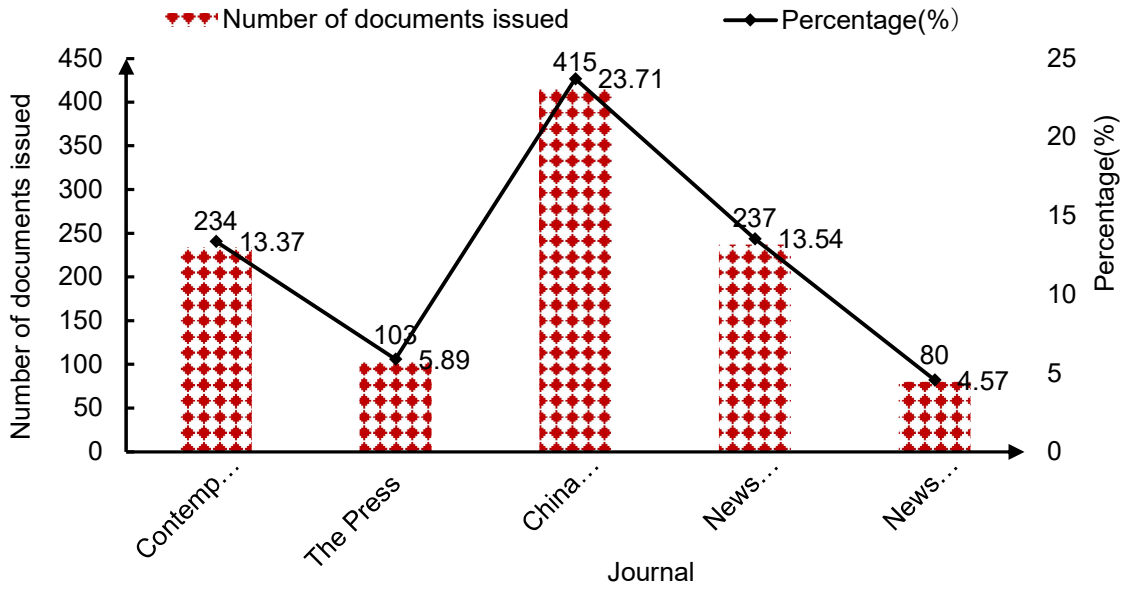
Communication, News University, Journalists, Contemporary Communication, The Press, China Reporter, News and Writing, and News Knowledge, as samples, and collected 1750 article data.

This paper used the method of knowledge graph analysis to describe the hot spots in the research fields of the ten major journals, and carried out a comprehensive quantitative analysis and visual display of the research fields of the ten major journals, from which five hot spots of the ten major journals were extracted, and also included the co-presence of the author of the article, the co-presence of the institution, and so on.

Figure 5 shows the proportion of the number of papers published by the top ten journals from 2021.12 to 2022.11.



(a) Number of the first five publications



(b) Number of the last five volumes

Figure 5: Proportion of the number of papers published by the top ten journals of journalism and communication during the period from 2021.12 to 2022.11

It can be seen from Figure 5 (a) that the number of articles published by the International Press was 93, accounting for 5.31%. News and Communication Research published 84 articles, accounting for 4.8%. Modern

Communication published 292 articles, accounting for 16.69%. The number of articles published by News University was 108, accounting for 6.17%. The number of articles published by the Journalist was 104, accounting for 5.94%.

It can be seen from Figure 5 (b) that the number of published articles of Contemporary Communication was 234, accounting for 13.37%. The number of articles published by The Press was 103, accounting for 5.89%. The number of articles published by China Reporter was 415, accounting for 23.71%. News and Writing published 237 articles, accounting for 13.54%. The number of articles published by News Knowledge was 80, accounting for 4.57%.

From here, it can be seen that there are certain differences in the proportion of the number of papers issued by the top ten journals. However, it should be noted that the quantity of published papers is related to the nature of the journal, not to the quality of published papers and journals.

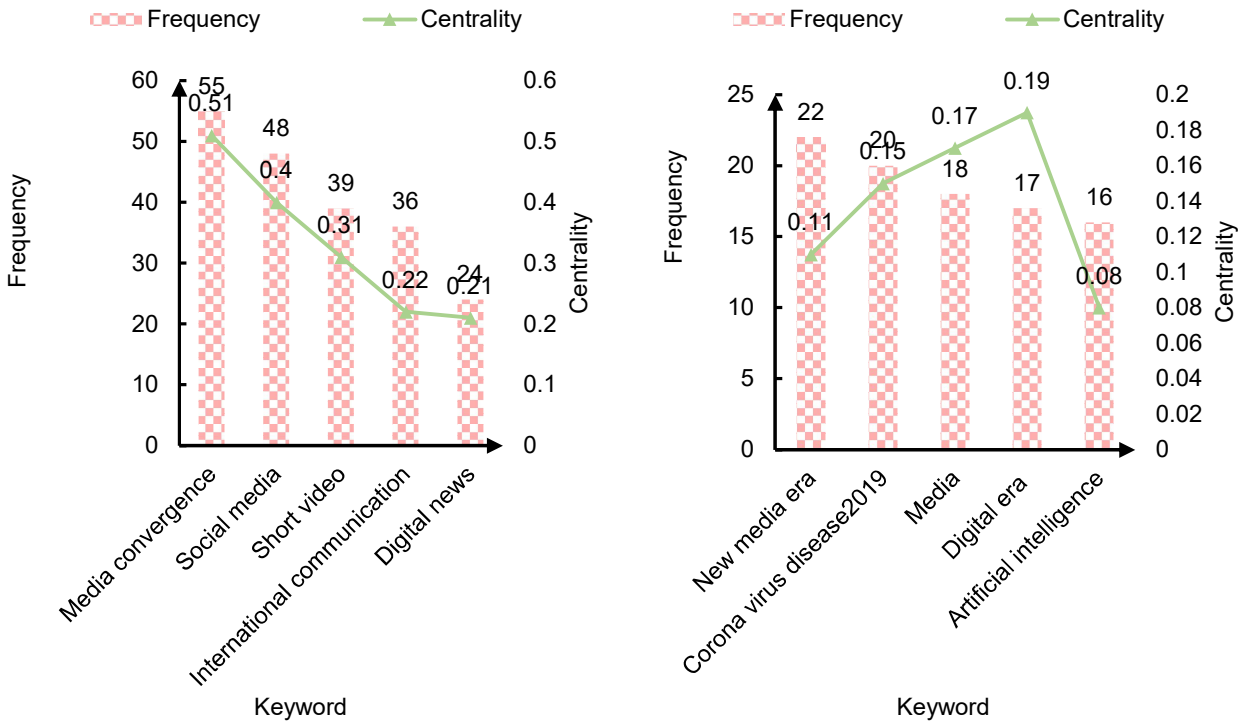
III. B. Hot Spots and Trends of Journalism and Communication

The so-called research hotspot refers to the focus of current academic circles in a specific discipline. Key words are the most frequent words in a document, and also the most core words summarized by the author, which can reflect the main research content and research theme of the document. This paper made statistics on the keywords of ten major news and communication journals from December 2021 to November 2022 in order to have a general understanding of the hot topics that have emerged in the past year.

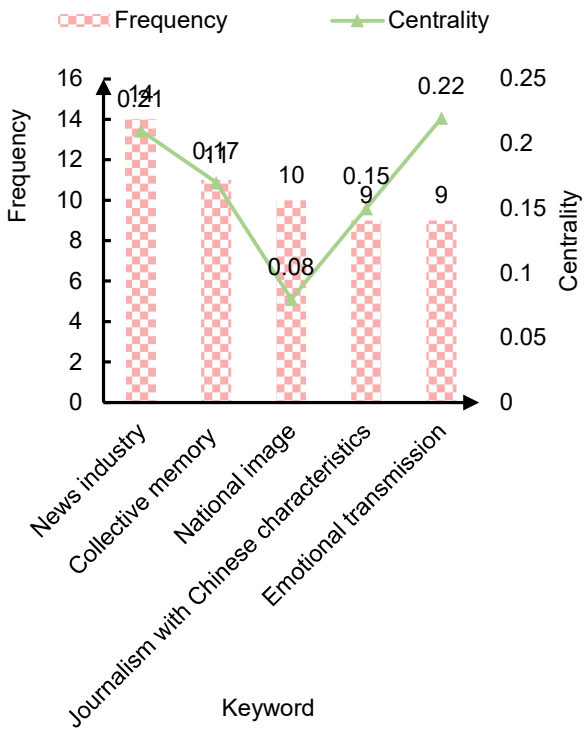
Relevant data was input into CiteSpace software and keyword option was selected. Through the coexistence analysis of the keywords in the top ten journals, the knowledge graph of the keywords in the top ten journals from December 2021 to November 2022 was obtained.

In the knowledge graph, the size of node and text font size indicates the frequency of keyword occurrence. The larger the node and font size, the higher the frequency of keyword occurrence. The thickness of the line reflects the frequency of symbiosis between two nodes, and the thickness of the line indicates the frequency of symbiosis.

In order to more specifically present the research hotspots of the recent year, this paper obtained the keyword data of high frequency and high school heart in this period through CiteSpace, and combined the keywords of high frequency and high school heart. The specific data is shown in Figure 6.



(a) 1-5 high-frequency keywords (b) 6-10 high-frequency keywords



(c) 11-15 High frequency keywords

Figure 6: Key words of high frequency and high school

It can be seen from Figure 6 (a) that the frequency of media convergence was the highest, 55, and the centrality was also the highest, 0.51. The second was social media, with frequency and centrality of 48 and 0.40 respectively. The third was short video, with frequency and centrality of 39 and 0.31 respectively. The frequency and centrality of international communication were 36 and 0.22 respectively, and the frequency and centrality of digital news were 24 and 0.21 respectively. It can be seen that the higher the frequency of the top five keywords, the greater the centrality.

According to Figure 6 (b), the frequency and centrality of the new media era were 22 and 0.11 respectively. The frequency and centrality of COVID-19 were 20 and 0.15 respectively. The frequency and centrality of media were 18 and 0.17 respectively. The frequency and centrality of the digital era were 17 and 0.19 respectively. The frequency and centrality of artificial intelligence were 16 and 0.08 respectively.

It can be seen from Figure 6 (c) that the frequency and centrality of the news industry were 14 and 0.21 respectively. The frequency and centrality of collective memory were 11 and 0.17, respectively. The frequency and centrality of national image were 10 and 0.08 respectively. The frequency and centrality of journalism with Chinese characteristics were 9 and 0.15 respectively. The frequency and centrality of emotional transmission were 9 and 0.22 respectively. From the data in Figure 6 (b) and Figure 6 (c), the higher the frequency of keywords, the less the size of the heart.

The higher the frequency of keywords, the higher the degree of attention of researchers to the research topics involved, and the high-frequency keyword list can directly reflect the hot areas of research. The core of keywords represents the relationship between keywords and other keywords. The higher the centrality of a node, the closer the relationship between this node and other nodes, and the importance of this node in the entire network map.

In order to analyze the relationship between keywords more deeply, this paper took the keyword co-occurrence map as the basis, and then carried out cluster analysis on keywords. The Log-Likelihood Ratio (LLR) algorithm was used to calculate, and finally the knowledge graph of keyword clustering was obtained.

From 2021.12-2022.11, the knowledge graph of ten journals in journalism and communication can be seen that the order is from 0 to 4, namely # 0 algorithm, # 1 digital era, # 2 COVID-19, # 3 international broadcasting, and # 4 digital news.

Table 1 shows the statistical list of ten journal lattice clustering keywords in journalism and communication from 2021.12 to 2022.11.

Table 1: Statistics of clustering keywords

Cluster number	Scale	Cluster average contour value	Label (LLR)				
0	16	0.894	Algorithm	Media convergence	Intelligent media	Artificial intelligence	...
1	14	0.862	Digital era	Social media	Traditional journalism	Communication research	...
2	13	0.836	Corona Virus Disease 2019	Short video	Emotional transmission	Collective memory	...
3	11	0.780	International communication	Cross-media communication	National image	Empathy communication	...
4	9	0.721	Digital news	News practitioners	Integrated communication	Journalism	...

It can be seen from Table 1 that the algorithm covers media convergence, intelligent media, artificial intelligence and other related tags. The digital era covers social media, traditional journalism, communication research and other related labels. The COVID-19 covers short videos, emotional transmission, collective memory and other related labels. International communication covers relevant labels such as cross-media communication, national image, empathy communication, etc. Digital news covers relevant labels such as news practitioners, integrated communication, and journalism. The average clustering contour value of the algorithm was 0.894, and the average clustering contour value of the digital era was 0.862. The cluster average contour value of COVID-19 was 0.836. The cluster average profile value of international communication was 0.780. The clustering average contour value of digital news was 0.721.

Based on the above analysis, this paper summarized the five research hotspots of the ten major journals of journalism and communication from December 2021 to November 2022, including media convergence, social media, short video, international communication and digital news.

IV. Conclusions

CiteSpace can help researchers understand the current research situation in this field more intuitively and accurately by analyzing the hot spots of journalistic and communication research in China. According to the co-occurrence relationship of high-frequency keywords, it can be speculated that the core authors and high-yield institutions in this field can better understand the main research directions and dynamic changes in this field. At the same time, it can also have a deeper understanding of the important role of journalism and communication in China's social, political, economic and cultural fields, and help further study the development direction of journalism and communication. Through the retrieval and data collection of the ten journals in the field of journalism and communication, this paper used CiteSpace software to draw a knowledge graph of research hotspots. Through comparative analysis, the research hotspots in the field of journalism and communication were found, which provided an important reference for future research on journalism and communication. In general, there are many scholars studying journalism and communication, but the number of documents is small. Moreover, most of them are theoretical articles written by media personnel, and empirical research is relatively small. Therefore, in future research, it is necessary to pay attention to the cross-integration of communication, sociology and other related disciplines to increase the proportion of empirical research in the field of journalism and communication. At the same time, in terms of topic selection, it is necessary to pay more attention to and research on issues related to online public opinion, and be paid attention to the integration and development of traditional media and new media to promote the sustainable development of journalism and communication from multiple levels.

References

[1] Yang Jie, and Hua Jin. "Application of big data analysis and visualization technology in news communication." Computer-Aided Design and Applications 17.2 (2020): 134-144.

[2] Lewis Seth C., Andrea L. Guzman, and Thomas R. Schmidt. "Automation, journalism, and human–machine communication: Rethinking roles and relationships of humans and machines in news." Digital journalism 7.4 (2019): 409-427.

[3] Lin Yisitie, and Shitao Zhang. "The proliferation of fake news in network communication and the reconstruction of media credibility." Academic journal of humanities & social sciences 4.11 (2021): 6-10.

[4] Greussing Esther, Sabrina Heike Kessler, and Hajo G. Boomgaarden. "Learning from science news via interactive and animated data visualizations: An investigation combining eye tracking, online survey, and cued retrospective reporting." Science communication 42.6 (2020): 803-828.

[5] Usher Nikki. "Putting “place” in the center of journalism research: A way forward to understand challenges to trust and knowledge in news." Journalism & Communication Monographs 21.2 (2019): 84-146.
[6] Wang Guofeng, Xiuzhen Wu, and Qiao Li. "A bibliometric study of news discourse analysis (1988–2020)." Discourse & Communication 16.1 (2022): 110-128.
[7] Carlson Matt. "Automating judgment? Algorithmic judgment, news knowledge, and journalistic professionalism." New media & society 20.5 (2018): 1755-1772.
[8] Kruikemeier Sanne, Sophie Lecheler, and Ming M. Boyer. "Learning from news on different media platforms: An eye-tracking experiment." Political Communication 35.1 (2018): 75-96.
[9] Corrales-Garay Diego, Marta Ortiz-de-Urbina-Criado, and Eva-Maria Mora-Valentin. "Knowledge areas, themes and future research on open data: A co-word analysis." Government information quarterly 36.1 (2019): 77-87.
[10] Shao Yuhong, Hu Zhongyi, Luo Mingzhi, Huo Tingting, Zhao Qingxue. "What is the policy focus for tourism recovery after the outbreak of COVID-19? A co-word analysis." Current Issues in Tourism 24.7 (2021): 899-904.
[11] Tan Luc Phan, Lan Pham Xuan, Le Angelina Nhat Hanh, Trang Bui Thanh. "A co-citation and co-word analysis of social entrepreneurship research." Journal of Social Entrepreneurship 13.3 (2022): 324-339.
[12] Sedighi Mehri. "Using of co-word analysis method in mapping of the structure of scientific fields (case study: The field of Informetrics)." Iranian Journal of Information Processing and Management 30.2 (2022): 373-396.
[13] Zhu Xiang, and Yunqiu Zhang. "Co-word analysis method based on meta-path of subject knowledge network." Scientometrics 123.2 (2020): 753-766.
[14] Soheili F., A. Khasseh, and P. Koranian. "Thematic trends of concepts in Knowledge and Information Science based on co-word analysis in Iran." Librarianship and Information Organization Studies 29.2 (2018): 171-190.
[15] Katsurai Marie, and Shunsuke Ono. "TrendNets: mapping emerging research trends from dynamic co-word networks via sparse representation." Scientometrics 121.3 (2019): 1583-1598.
[16] Tang Muh-Chyun, Weijen Teng, and Miaohua Lin. "Determining the critical thresholds for co-word network based on the theory of percolation transition: A case study in Buddhist studies." Journal of documentation 76.2 (2019): 462-483.
[17] Qaiser Shahzad, and Ramsha Ali. "Text mining: use of TF-IDF to examine the relevance of words to documents." International Journal of Computer Applications 181.1 (2018): 25-29.
[18] Pimpalkar AMIT PURUSHOTTAM, and R. Jeberson Retna Raj. "Influence of pre-processing strategies on the performance of ML classifiers exploiting TF-IDF and BOW features." ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal 9.2 (2020): 49-68.
[19] Zhuohao W. A. N. G., W. A. N. G. Dong, and L. I. Qing. "Keyword Extraction from Scientific Research Projects Based on SRP-TF-IDF." Chinese Journal of Electronics 30.4 (2021): 652-657.