Comparison of Research on Social Media in China and Foreign Countries Based on Bibliometric Analysis

Peng Yunan¹², Wang Zongshui¹², Feng Yueyang¹², Yang shengqin¹²

¹ School of Economics and Management, Beijing Information Science and Technology University, Beijing 100192, China;
² Beijing Key Lab of Green Development and Decision Based on Big Data, Beijing 100192, China
870431044@qq.com

Abstract. This paper aims to compare the similarities and differences of Chinese and foreign research on social media. Cluster analysis, co-word analysis and social network analysis have been used to compare the research trends, content relevance and network structure of research on social media. (1) Similarities: Social media is a hot topic for scholars both in China and foreign countries; the development of social network technology leads to many new areas; social media research tends to study privacy, commerciality and service; hotspots of social media research in China and foreign countries both include social media, platform and social network. (2) Difference: Chinese research mainly focuses on big data, Wechat and Weibo, while foreign research mainly focuses on Facebook, Twitter and other kinds of social network media; Chinese research is clustered while foreign research is extensive; Chinese research pays more attention on macroscopic perspectives, such as media convergence, big data and libraries; foreign research tends to microcosmic perspectives, such as emotional analysis, crisis information dissemination, social media marketing.

1. Introduction
There were 829 million Internet users by December 2018 in China, of which mobile Internet users account for 98.6%[1]. It is an urgent task to promote the integrated development of media and build All-media with the media of Internet and smart phones[2]. Social media mainly exchanges information through internet platforms, such as Social networking sites(SNS) [3]. Social media changes the way of communication, improves the efficiency of information diffusion and becomes a bridge between users[4]. Social media is used for emotional communication, information sharing, communication and interaction among users which mainly bases on Web2.0. The advantages of social media such as immediacy and easy access[5] make it becomes an important means of enterprise marketing.

At present, the research on social media involves a wide range of contents but they are relatively scattered and lacked systematic comparative research on contents and trends in China and foreign countries. Therefore, this paper collects data from CNKI database and Web of Science (WOS) database, then uses Bicomb2, Pajek, SPSS, Excel and other research software to analysis the data. Then through artificial interpretation, social network analysis (SNA), cluster analysis, co-word analysis[6], bibliometrics[7] and other methods to analysis the research contents and development trend of social media research in China and foreign countries in 2009-2018. This paper divides the development trend into two stages, then summarizes the similarities and differences of social media research in China and foreign countries by comparing keyword clustering, keyword network and co-occurrence relationship.
2. Research design

2.1. Data collection

CNKI and WOS cover the main literature data, which are widely used by scholars in China and foreign countries. Many types of research have selected CNKI and WOS databases, such as network rumors [8], social media [9], social network analysis [10], information security [11], etc. Therefore, CNKI and WOS databases are selected as data sources in the way of advanced retrieval and select the year between 2009-2018. The retrieval date of China part is March 9, 2019. The source journal is designated as CSSCI. The retrieval conditions are as follows: title, keyword and abstract are all "social media" with "or" as the relation retrieval term in CSSCI. 2317 documents are retrieved through the above methods. The retrieval date of the foreign part is April 11, 2019. Using "title" = (Social Media) for data retrieval, database = SCI-Expanded, SSCI, A&HCI. Use article as the document type and English as the language definition to refine, and retrieve 5542 documents (Table.1).

| Database | CNKI | WOS |
|----------|------|------|
| Subject words | Social media | Social media |
| Time interval | 2009-2018 | 2009-2018 |
| Literature type | Journal | Article |
| Language | Chinese | English |
| Research field | Basic science, Engineering technology, Philosophy and humanities, Social science, Information technology, Economy and management science | Communication, Business economics, Information science library science, Psychology |
| Data sources | CSSCI | SCI-EXPANDED, SSCI, A&HCI |

Delete incomplete information and different topics articles through artificial interpretation and data preprocessing. Finally, 2287 Chinese articles and 3003 foreign articles were obtained and the trend chart of the issued articles is shown in Fig.1.

Fig.1 The number of published articles per year from 2009 to 2018

We illustrate such growth in Fig.1, which shows a rapid upward trend of published articles per year. The polynomial trend line equation of China is $y = -0.5602x^3 + 10.933x^2 - 2.9916x - 6.3$ ($R^2 = 0.9839$) but the foreign countries is $y = -1.2294x^3 + 24.183x^2 - 60.269x + 72.633$ ($R^2 = 0.9867$), which means that the trend line has high consistency with real condition. The number of Chinese and foreign publications increased exponentially, but the number of related publications in 2009-2013 was relatively small. Since 2013, the volume of social media's publications has been increasing rapidly, especially the foreign publications, which are distributed in the range of 200 to 700. Therefore, this paper divides the number of documents issued from 2009 to 2018 into two stages: 2009-2013 is the first stage and 2014-2018 is the second stage. The number of Chinese articles is 394 and 1893 and foreign is 530 and 2473 in two stage respectively.
2.2 Research methods
There are methods in this paper.

(1) Cluster analysis is a data mining method, which can analyze the potential association between
data and the data affinity[12]. It is mainly divided into four steps: data collection, feature selection,
clustering process, result evaluation. Construct the association matrix[13] to measure the research and
select available data then use SPSS software to cluster the word matrix.

(2) Social network analysis (SNA) is mainly used in the study of relationships, diffusion, influence,
customers, intelligence, management and other fields[14]. It is an interdisciplinary research paradigm.
SNA analysis network characteristics through centrality, aggregation coefficient, network density and
node average, then analysis social phenomena by quantifying network structure[15]. In this paper,
bicomb2 and pejak software are used to draw social network diagram and analysis corresponding
indicators.

(3) Co-word analysis is based on word frequency analysis[16] to reflect the relevance between
different keywords. The advantage is that it reflects the relationship between individuals through
quantitative analysis.

3. Research analysis

3.1. Cluster result analysis
Keywords can highly summarize the theme, hot spots and trends of the article[15]. In this paper, 2287
(Chinese) and 3003 (foreign) literature were selected by Bicomb2 software in stages and then filter the
keywords. Then we use manual interpretation to delete irrelevant words. About the literature in China
and foreign countries, the top 20 keywords of word frequency are selected to generate the text matrix,
which is then exported to the text. SPSS software is used to cluster the text matrix data. In this paper,
the square Euclidean distance method is used to draw the genealogy as shown in Fig.2 and Fig.3.

![Fig.2 Clustering genealogy of high frequency keywords about Chinese social media](image-url)
According to Fig. 2 and Fig. 3, some key words in China are summarized as SNA, media integration, e-government [16], social network and social website. Foreign part is mainly divided into four aspects, social marketing, social networks, Web 2.0 and social networking sites (SNS). Some research fields have strong similarities in China and foreign countries, such as social networks and SNS. The main characteristics of Chinese research are media convergence, e-government and so on. The development of audio-visual media integration has promoted the deep integration of traditional media and new media. In addition, government information sharing through social media has become a hot spot of current research. The foreign part is characterized by social marketing and Web 2.0 related research.

| Num | Name               | Key words                                                                 | Num | Name               | Key words                                                                 |
|-----|--------------------|---------------------------------------------------------------------------|-----|--------------------|---------------------------------------------------------------------------|
| 1   | SNA                | Newspaper, social network analysis, social capital                         | 1   | social media       | marketing, SNA, engagement, INS, corporate social responsibility, social media marketing, social movements, sentiment analysis, crisis communication |
|     | media integration  | Opinion leaders, News, influencing factors, media integration, government weibo, mobile Internet, library | 2   | social networking  | political participation, public relations, social networking, social networking sites, China, social networks, Internet, Facebook, Twitter |

Fig.3 Clustering genealogy of high frequency keywords about foreign social media
From Table 2, it can be seen that the key words about Chinese and foreign research include social network, social capital and e-government, reflecting the consistency of research in these aspects. Chinese research on social media focuses on big data while foreign research focuses on political participation and public relations in recent years. The research in China about SNS mainly focuses on WeChat, Weibo, new media hot spots while foreign research includes social networking, China, Internet, Facebook, Twitter, etc. In addition, Chinese scholars pay more attention to media integration, opinion leaders and libraries, while foreign scholars pay more attention to emotional analysis, crisis information dissemination.

3.2. Co-word analysis of keywords

According to the statistical data of keyword frequency, we use bicomb2 software to extract co-occurrence matrix. Select social media as the main focus keywords then construct co-occurrence coefficient matrix and analysis the Ochia index of two-stage (Table 3).

| Keywords index | Keywords index | Keywords index | Keywords index |
|----------------|----------------|----------------|----------------|
| Weibo 0.62     | WeChat 0.53    | social network 0.26 | Twitter 0.26 |
| New media 0.43 | Weibo 0.54    | Facebook 0.30     | Facebook 0.27 |
| Press 0.39     | Big data 0.57 | Twitter 0.28      | social networks 0.20 |
| Traditional media 0.38 | New media 0.50 | web 2.0 0.32 | internet 0.14 |
| SNA 0.33       | Social network 0.41 | public relations 0.19 | journalism 0.14 |
| Mobile Internet 0.34 | E-government library 0.31 | internet 0.13 | China 0.12 |
| E-government 0.37 | library 0.31 | blogs 0.17 | sentiment analysis 0.09 |
| Media convergence 0.35 | influence factor 0.43 | media 0 | social capital 0.11 |
| News 0.35      | Media convergenc e 0.28 | corporate social responsibility 0.04 | crisis communicatio n 0.12 |
| Web2.0 0.39    | Internet 0.31 | social capital 0.02 | big data 0.10 |
| Twitter 0.31   | social capital 0.28 | knowledge management 0.20 | political participation 0.11 |
| Total          | 4.62           | 4.03            | 1.91           | 1.66 |

In the first stage (2009-2013), the cumulative Ochia index of social media in China is 4.62, among which the Ochia index of Weibo, New media, newspaper, Web2.0 and social media is higher, respectively 0.62, 0.43, 0.39 and 0.39. It shows that Weibo, New media and social media are highly related in this stage. In addition, with the rapid development of social media, the number of new media users such as Weibo increased, and the number of paper media users decreased significantly. The
research on new media and traditional media as well as Weibo become research hotspots. The cumulative foreign Ochiai index of social media is 1.91, among which the Ochiai index of social network, Web 2.0, Facebook and Twitter are higher, respectively 0.26, 0.32, 0.30 and 0.28. Social media in foreign mainly studies the social network relationship between Facebook and Twitter in the era of Web 2.0 from a micro perspective. In the first stage, social media in China and foreign countries is the focus of scholars' research. Chinese research mainly focuses on the alternation of traditional media and new media. Foreign study mainly focuses on social networks and social platforms.

In the second stage (2014-2018), the cumulative Ochiai index of Chinese high-frequency keywords and social media are respectively 4.03, among which the Ochiai index of WeChat, Weibo and big data are higher, respectively 0.53, 0.54 and 0.57. WeChat attracts a large number of social media users. At this stage, with the development of Internet technology, big data and social platforms have become hotspot in social media research.

Compared with the two stages, the fields involved in social media are gradually increasing. Based on the traditional research fields in China, they are consistent with the social development trend and then derive new research issues such as WeChat, big data, library, etc. big data, WeChat, Weibo and other related issues have received high attention. The Ochiai index of high-frequency keywords with social media in foreign is 1.66, among which the high Ochiai index keywords are Facebook and Twitter, and the Ochiai index are 0.27 and 0.26 respectively. At this stage, the foreign research on Facebook and Twitter is still hot, but the cumulative Ochiai index has a downward trend. In addition, emotional analysis, crisis information dissemination, big data, etc. have become new foreign research hot spots. In the second stage, the Ochiai index of social media and other high-frequency words in China and foreign countries is lower and the relevance is reduced. Chinese research on social platforms and big data has become a new hotspot while foreign research is still mainly focused on Facebook and Twitter.

Through the calculation of Ochiai index, it can be seen that the research on social media is more concentrated in China, with less fields but more research relevance. In the foreign, the research field is more extensive and scattered. Compared with the two stages, big data is more popular in China. Foreign in-depth study of social software platforms earlier, such as Facebook and Twitter. And the foreign research is more advanced in some aspects, such as emotional analysis, crisis information dissemination, public relations and so on, which provides an important reference for related research in China.

### 3.3 Network centric comparison

Using Pejak software to build a two-stage keyword network Fig.4a-5b. For the convenience of comparison, the network diagrams in China and foreign countries are drawn by the proximity to the center. According to the centrality of nodes, core nodes are represented by●, middle nodes are represented by●, and edge nodes are represented by●. The node is divided into three layers: The core node is the node with \( C_{ij}(v_j) \geq 0.7 \), the middle node is the node with \( 0.7 \geq C_{ij}(v_j) \geq 0.6 \), the edge node is the node with \( < 0.6 \); the cumulative percentage of word frequency is set to be \( \leq 30\% \).

---

*Fig.4a Chinese high frequency keyword network diagram of the first stage*
In the first stage, 55 keywords were obtained in Chinese part. The rank of core vertexes based on their centrality values is (the number within brackets is the centrality value): social media (1.00), Weibo(0.77). In this stage, 8 intermediate nodes account for 14.5% and 45 edge nodes account for 81.8%; In foreign part, 55 key words are obtained, including 1 core node: social media (0.92), 0 intermediate node and 54 edge nodes, accounting for 98.2%.

Fig.4b Foreign high frequency keyword network diagram of the first stage

Fig.5a Chinese high frequency keyword network diagram of the Second stage
Fig.5b Foreign high frequency keyword network diagram of the Second stage

In the second stage, 121 keywords were obtained in Chinese part, including 6 core nodes: social media (1.00), social network (0.81), new media (0.78), big data (0.81), WeChat (0.79), Weibo (0.75). There are 14 central nodes, accounting for 11.6%, and 101 marginal nodes, accounting for 83.5%. There are 81 key words obtained in the foreign part, and 4 core nodes: social media (0.99), Twitter (0.70), Facebook (0.74), and social network (0.72). There are 0 central nodes and 117 edge nodes, accounting for 96.7%.

From Fig.4a - 5b, it can be seen that the number of core nodes and high-frequency keywords in China and foreign countries increases. The research fields in China and foreign countries are more extensive, and relevant research hotspots include social networking sites, such as WeChat in China and Facebook and Twitter at foreign. In addition, the research on social network is also on the rise. However, there are many differences in the research on social media in China and foreign countries. The central node of the two stages in foreign countries is 0, which shows that the research on social network is more divergent and more extensive than that in China. In addition to the high-frequency hot key words of the core node, the relevance between the research hot spots is small. What's unique in China is that the research on big data becoming an upsurge in the new era.

In order to further analysis the keyword network structure, we calculate the network density, clustering coefficient and network average degree of the two stages. The results are shown in Table.4.

| Characteristic indexes | China First stage | China Second stage | Foreign First stage | Foreign Second stage |
|------------------------|------------------|--------------------|---------------------|---------------------|
| Network density        | 0.249            | 0.274              | 0.127               | 0.154               |
| Clustering coefficient | 0.429            | 0.397              | 0.238               | 0.283               |
| Network average degree | 27.382           | 66.331             | 14                  | 24.914              |

It can be seen from table.4 that the clustering coefficient of Chinese high-frequency keywords shows a downward trend, and the network density changes slightly. It shows that the correlation between the hot spots gradually decreases. There are more and more new content present on the basis of the traditional research content, such as WeChat, big data, etc. It reflects that the continuous innovation of social media and its related research. However, the network density and clustering coefficient of foreign
high-frequency keywords are on the rise. This shows that the foreign research focus on social media is gradually increasing, which is in sharp contrast with the foreign research. The average degree of Chinese and foreign networks has increased significantly, which indicates that the research hot pots are still expanding in scale and structure, which is consistent with the upsurge of contemporary social media.

4. Analysis the trend of social media research

In order to further summarize the results of literature analysis, six key words of social media research were selected for analysis in China and foreign. As shown in Fig. 6 and Fig. 7.

![Fig.6 The trend chart of high-frequency keywords in China](image)

![Fig.7 The trend chart of high-frequency keywords at foreign](image)

From Fig.6, we can see that the frequency of high-frequency keywords in these fields in the two stages is increasing, it explains that the relevant research are growing. However, there is a big gap between the two stages of research on WeChat and big data. It shows that big data related content has become a hot topic of social media research. After 2011, WeChat became a hot spot. The development of social media has changed the way of traditional social, enterprise marketing and government office.

From Figure 7, it can be seen that the foreign research hotspots on social networks, Facebook and Twitter rise sharply, especially Facebook and Twitter. However, the research on Web 2.0 is declining. Public relations and social movements are the hot spots with relatively flat growth trend. To sum up, social media has developed rapidly in the past 10 years which provides a convenient, reliable and shared communication platform. Compared with the research of foreign social media, six high-frequency keywords of the research in China have a significant upward trend, while the main hot spots in China and foreign countries are still concentrated on social networks and social platforms. Social platforms are mainly WeChat and Weibo in China, while in the foreign, they are mainly Facebook and Twitter. However, domestic research on big data, new media and e-government is more popular than foreign research. The research in China pays more attention to the specific application of social media in production and life. The research in foreign focus more on public relations and social movements and study the network relationship between objects in society.

5. Research conclusion

This paper compares the research on social media in two stages through the literature data in the past 10
years. It can be seen that many new areas emerging in social media. The relevant research has the characteristics of privacy, commerce and service. The similarity between the research in China and foreign countries is that the scale of related research is increasing and the content is diverse. There are more and more new branches derived in the traditional research field, such as government affairs, business, information dissemination, public relations, etc. Social media promotes social development, it mainly studies the behavior and influence of users based on social websites and social networks, such as Weibo and WeChat in China as well as Facebook and Twitter in foreign. The mobile phone links the users, and social networking becomes an important way of public life. Social media brings a surge of traffic and virtual communication to social communication, but also brings information fragmentation and entertainment.

There are many differences in the research on social media in China and foreign countries. In China, there are different types of research on social media, mainly focusing on WeChat and Weibo. In recent years, WeChat has gradually become a new research hotspot. There are more attention has been paid to social network media such as Facebook and twitter in foreign countries. Since the advent of Facebook and twitter, it has been a hot topic of research and the relevant research content is more and more extensive. The domestic research hotspots on social media are highly concentrated, and each hotspot has a large degree of relevance. The foreign research on social software from the micro perspective which found its social network relationship, and the research content is relatively scattered. In terms of research content, the research on social platforms and social network in social media in China and foreign countries is a hot issue. The agglomeration of domestic hot spots has gradually decreased and more emerging areas of social media have emerged, such as E-government, opinion leaders, etc. The research on media convergence and library will continue to rise. The research in China is constantly exploring and innovating based on the traditional research. Big data, media convergence, WeChat, library, etc. become new research hotspots; The foreign research on social media will increase the relevance between various fields, and more emphasis will be placed on the network relationship between research hotspots and information diffusion, such as crisis information diffusion, emotional analysis, public information diffusion and public relations, etc.

Acknowledgments
This research was supported by the National Key R&D Program of China (No. 2017YFB1400500), Beijing Municipal Education Commission Social Science Foundation (No. SM201911232005) and Qinxin Talents Cultivation Program of Beijing Information Science & Technology University (No. QXTCP B201906).

About the authors
Peng Yunan is a MA student of Industrial Engineering, School of Economics and Management, BISTU. Peng Yunan is the corresponding author and can be contacted at: 870431044@qq.com
Wang Zongshui is a vice Professor of School of Economics and Management, BISTU.
Feng Yueyang is a MA student of Management Science, School of Economics and Management, BISTU.
Yang Shengqin is a MA student of Management Science, School of Economics and Management, BISTU.

References
[1] Statistical report on the development of Internet in China [EB/OL].[2019-03-01]. http://www.cnnic.net.cn/gywm/xwzx/rdxw/20172017_7056/201902/W020190228474508417254.pdf.
[2] Zhu Hongjun. Out of the Structural Dilemma: A Way to Explore the Deep Path of Media Integration. Journalism Review, 2019,03(11): 39-44. (in Chinese)
[3] Wang Zongshui, Zhao Hong, Liu Yu ,Qin Xuzhong . Evolution, Development and Application of Social Network Paradigm: Evidence from CSSCI Database of China. Journal of the China Society for Scientific and Technical Information, 2015,34(12):1235-1245. (in Chinese)
[4] Kietzmann J H, Hermkens K, McCarthy I P, et al. Social media? Get serious! Understanding the functional building blocks of social media. Business Horizons, 2011, 54(3): 241-251.

[5] Khan M A. Social media’s influence on hospitality & tourism management. Journal of Business & Hotel Management, 2012, 1(1): 1-2.

[6] He Q. Knowledge discovery through co-word analysis. Library Trends 48 (1):133–159.

[7] De Bellis N. Bibliometrics and citation analysis: from the science citation index to Cybermetrics. Lanham : Scarecrow Press, 2009.

[8] Tong Wensheng, Wang Jiancheng, Zeng, Runxi. Analysis on the Research Topics and Contents of Internet Rumors in China—Taking Database of CNKI from 2002 – 2013 as Samples. Journal of Intelligence, 2014,33(07):135-140+150.(in Chinese)

[9] Yin Xiangxu, Zhang Gengping, Li Xiaofei. Analysis of the current situation of Information Science Research Based on keyword statistics”. Journal of information,2009,28(11):1-4. (in Chinese)

[10] Lu Jianfeng. Policy leading development technology driven innovation: Reflections on the current practice path of media integration and innovation. Media, 2019 (08) : 56-59. (in Chinese)

[11] Wang Bing, Wu Chao. International research progress of security management information system: typical literature analysis based on web of science database. Intelligence journal, 2018.37(11):131-136. (in Chinese)

[12] Yu Liping. Research on the selection of multi-attribute evaluation methods of journals based on cluster analysis -- the method of selecting the consistency of clustering results. Library and information work, 2018, 62(21):80-86. (in Chinese)

[13] Huang Fu. Clustering algorithm analysis of common tools in scientometrics. Science and technology management research, 2018,38(18): 232-238. (in Chinese)

[14] Wang Zongshui, Zhao Hong, Wang Yan. Social networks in marketing research 2001–2014: a co-word analysis. Scientometrics,2015,105(01): 65-82.

[15] Zhao Lijuan. Basic theoretical methods of social network analysis and its application in Information Science”. Library science research, 2011(20):9-12. (in Chinese)

[16] Wei Ruibin, Jiang Qianwen, Zhang Ruili. A comparative study of research methods based on literature co citation and co word analysis: a case study of CO citation and content analysis. Journal of information,2019,32(08):36-42. (in Chinese)