
AI translation · View original & related papers at
chinaxiv.org/items/chinaxiv-202311.00173

Postprint: Research Topic Mining in China's Green Consumption Field Using Text Mining and Complex Networks

Authors: Liu Jieping, Xu Jinya

Date: 2023-11-07T00:00:00+00:00

Abstract

[Purpose/Significance] The development of green consumption constitutes an important component of China's 14th Five-Year Plan and the 2035 Long-Range Objectives Outline. Mining research topics in the green consumption domain facilitates rapid understanding of current research advances and hotspots, providing reference and guidance for further investigation. [Method/Process] Based on text mining technology and complex network analysis methods, this study proposes an approach that comprehensively considers document titles, abstracts, and keywords, employs text segmentation techniques to extract thematic terms, and determines co-occurrence weights of binary thematic word pairs using the AHP method. To address the inability of the traditional word frequency g-index to effectively exclude high-frequency general terms, this study optimizes the traditional word frequency g-index based on the TF-IDF algorithm and proposes the TI-g index. An empirical study is conducted on academic literature in China's green consumption field from 2010 to 2022. [Results/Conclusion] This study generates a heatmap of research topic evolution in China's green consumption field since 2010, mines research hotspots since 2018, and identifies four major thematic domains in this field.

Full Text

Mining Research Topics in China's Green Consumption Field Based on Text Mining and Complex Networks

Liu Jieping, Xu Jinya

(Chengdu Neusoft University, Chengdu 611844, China)

Abstract

[Purpose/Significance] Promoting “green consumption” represents a fundamental objective in China’s 14th Five-Year Plan and Vision 2035. Mining research topics in the green consumption domain facilitates rapid comprehension of current research progress and hotspots, providing valuable references and guidance for future studies. **[Method/Process]** This study proposes a comprehensive methodology that integrates literature titles, abstracts, and keywords, employing text segmentation techniques to extract subject terms and utilizing the Analytic Hierarchy Process (AHP) to determine co-occurrence weights for binary term pairs. To address the limitation of traditional word frequency g-index in effectively eliminating “high-frequency generic terms,” we optimize it using the TF-IDF algorithm and propose the TI-g index. An empirical study was conducted on academic literature in China’s green consumption field from 2010 to 2022. **[Result/Conclusion]** We generated a heatmap illustrating the evolution of research topics in China’s green consumption field since 2010 and identified research hotspots emerging since 2018, recognizing four major thematic domains within this research area.

Keywords: Green consumption; Research hotspot; Text mining; Complex network; TF-IDF; Word frequency g-index

Introduction

Green consumption research originated in the 1970s. In 1987, British scholars Elkington and Hailes defined green consumption as “consumer behavior that avoids products harmful to health, wasteful of resources, over-packaged, derived from rare animals or natural resources, and detrimental to other countries, especially developing nations” [?]. The World Environmental Organization proposed the 5R principles of green consumption: Reduce, Reevaluate, Reuse, Recycle, and Rescue [?]. Strictly speaking, academia has not yet formed a unified definition of green consumption. At a broad level, green consumption is generally understood as behaviors that reduce waste and avoid pollution throughout the entire process of product purchase, use, and disposal. At a narrow level, green consumption focuses more specifically on the act of purchasing green products themselves [?].

As Chinese citizens’ environmental awareness continues to strengthen, the concept of green consumption has gained widespread dissemination and acceptance. According to the *Research Report on the Current Status of Green Consumption Among the Chinese Public*, 83.34% of respondents expressed support for green consumption behavior, with 46.75% indicating “strong support” [?]. Research by the China Chain Store and Franchise Association and Alibaba’s New Service Research Center suggests that over 60% of respondents are aware of green consumption, with post-2000 and post-1990 generations showing significantly higher awareness than other age groups, reaching 79% and 70% respectively [?]. In 2013, the State Council’s *Circular Economy Development Strategy and*

Recent Action Plan identified “green consumption” as a crucial measure for promoting circular economy development at the societal level. In 2014, Premier Li Keqiang proposed promoting green consumption and expanding the production of energy-saving products at a State Council executive meeting [?]. Since then, China has introduced a series of systems, plans, and standards related to green consumption. In 2021, “developing green consumption” was incorporated into the 14th Five-Year Plan and Vision 2035 [?], establishing it as a national development strategy.

With increasing public awareness of green consumption and growing government emphasis on environmental protection, academia has conducted extensive research on green consumption, accumulating substantial research outcomes. To better advance green consumption research and support the implementation of China’s green consumption strategy, it is necessary to comprehensively analyze previous research findings, trace the evolution of research topics in recent years, identify current research hotspots, and examine existing problems in the field. Our preliminary investigation reveals that existing research on green consumption topics has primarily focused on theoretical discussions and qualitative analyses, lacking quantitative analysis and systematic summaries of current status and trends [?]. Furthermore, existing literature mining techniques suffer from issues such as single research objects and subjective or ineffective methods for selecting high-frequency keywords that fail to eliminate generic terms [?].

Therefore, this study employs text mining technology and complex network analysis methods, using academic literature on green consumption in China from 2010 to 2022 as the research object, to mine research topics in this field, trace the evolution of research themes, identify current hotspots, and recognize four major thematic domains. We also propose future research directions and methodological improvements to address deficiencies in each thematic domain. Additionally, we offer optimization measures and improvement suggestions for existing literature mining techniques. The research framework, methodology, and key technologies proposed in this paper can also serve as references for literature mining in other fields.

1. Related Research

Complex networks originated in the 1980s at the Santa Fe Institute in the United States. Real-world complex networks generally exhibit characteristics such as self-organization and small-world properties. Complex network theory can describe and study complex systems and their topological structures. Since its inception, complex network analysis has been widely applied in various complex systems research [?], including talent mobility networks [?], transportation networks [?], power grids [?], financial networks [?], disease transmission [?], public opinion dissemination [?], and literature mining [?].

Research thematic domains in literature studies are similar to community characteristics in complex networks. Consequently, complex networks have been ex-

tensively applied in the field of literature mining, becoming one of the three major approaches to literature analysis [?, ?]. For instance, Holeab et al. proposed a complex literature mining method combining semantic and network analysis to examine research trends in Future-oriented Technology Analysis (FTA) [?]. Wang et al. used complex network theory to mine literature from the Scopus database, quantitatively describing significant features of international talent mobility [?]. Ortega et al. mined literature from Google Scholar Citations and found that the United States dominates the world scientific landscape [?]. Chae et al. conducted keyword network analysis on the semantic network structure of human resource management research in South Korea, revealing that the entire network structure exhibits a complex socially constructed semantic structure [?]. Xin et al. employed community detection techniques based on complex networks to identify eight major thematic research areas in forestry literature [?]. Liu et al. studied research hotspots in the surveying and mapping field using surveying journals as research objects [?]. He et al. examined the evolution of research trends in China's management field over 28 years based on complex network theory [?].

In the domain of green consumption, although some scholars have studied the dissemination of green consumption concepts using complex network analysis methods, literature mining of green consumption research remains scarce. Social network analysis represents an application of complex network theory in social relationship systems. Liu and Gan analyzed the research status and trends in China's green food field from a social network perspective [?]. Du et al. investigated research hotspots and thematic trends in China's green consumption behavior field using social network analysis and co-word analysis [?].

In summary, complex networks serve as a typical literature mining method and have been widely applied across various disciplines. However, their application in green consumption literature mining remains limited. Through reviewing relevant studies, we identify two major deficiencies in current research:

First, most scholars directly use literature keywords as subject terms for literature mining [?], yet literature keywords suffer from subjectivity and semantic ambiguity [?]. Some scholars combine titles, abstracts, and keywords for literature mining [?], but they treat these elements as a single "whole." We argue that titles, abstracts, and keywords should carry different weights when reflecting literature topics. Therefore, this paper proposes comprehensively considering literature titles, abstracts, and keywords, using text segmentation technology to extract subject terms, and employing the AHP method to determine the co-occurrence weights of binary term pairs. These weights directly influence edge weights in the final subject term network construction.

Second, to eliminate interference from numerous low-frequency subject terms, scholars generally focus only on high-frequency terms. Regarding the determination of high-frequency term quantities, many scholars rely on empirical judgment [?, ?, ?], a method lacking theoretical guidance and exhibiting subjectivity. To avoid this subjectivity, Yang et al. proposed the word frequency

g-index based on the scholar influence g-index. Its core principle is that when the cumulative frequency of g keywords is no less than g^2 times, while the cumulative frequency of $g+1$ keywords is less than $(g+1)^2$ times, g represents the word frequency g-index [?]. However, this index relies solely on term frequency. According to the TF-IDF algorithm (Term Frequency-Inverse Document Frequency), a classic text mining approach, many “high-frequency generic terms” exist in practice—terms with high frequency but weak semantic meaning. In this study, for example, “green consumption” appears frequently but fails to reflect specific research topics in the field, making it a high-frequency generic term. Therefore, this paper improves the word frequency g-index based on the TF-IDF algorithm and proposes the TI-g index to address this limitation.

2. Research Framework and Methodology

This study mines research topics in the green consumption field based on text mining technology, complex networks, and AHP theory. The research framework is illustrated in Figure 1 [Figure 1: see original paper].

2.1 Literature Collection and Preprocessing We retrieved relevant academic literature published between January 1, 2010, and December 31, 2022, from CNKI, Wanfang, and VIP databases using “green consumption” and “green purchase” as keywords in “title, abstract, and keywords” precise search mode. After merging, deduplicating, and standardizing the literature data, we obtained 7,318 documents. Subsequently, we used Pandas to clean the dataset, removing non-academic literature such as news reports, industry events, and call for papers, and eliminated documents with missing features, duplicates, or anomalies. After further manual verification, we finalized a dataset of 4,901 valid documents.

Figure 2 [Figure 2: see original paper] shows the annual publication volume from 2010 to 2022. Overall, the number of publications in China’s green consumption field has remained relatively stable, indicating that green consumption has consistently been a key research focus among Chinese scholars.

2.2 Text Segmentation and Subject Term Extraction Literature keywords are inherently independent terms, while titles and abstracts require text segmentation. This study employs the `jieba.lcut()` method for text segmentation of literature titles and abstracts. For stop words such as “research,” “policy,” and “suggestions,” we filtered them using comprehensive stop word libraries from Baidu, Harbin Institute of Technology, Sichuan University, and a custom stop word library. For unknown words, we compiled a dictionary by combining split keyword sets and mainstream input method dictionaries (Baidu, Sogou, QQ) along with custom terms. After replacing synonyms and near-synonyms (e.g., “college students” and “university students,” “green consumption behavior” and “customer green consumption behavior”), we obtained 18,007 subject terms.

2.3 High-Frequency Subject Term Selection Subject terms appear with varying frequencies in the literature corpus. For instance, “green consumption concept” appears most frequently with 4,921 occurrences. Statistics show that the cumulative frequency of the top 3,566 subject terms accounts for 80% of total frequency, while the remaining 14,441 terms account for only 20%, with frequencies below 6. Therefore, empirical research typically focuses only on high-frequency subject terms rather than the entire corpus.

As previously discussed, high-frequency term selection can be optimized using the TF-IDF algorithm to address limitations of traditional word frequency g-index. TF-IDF is a text mining weighting technique that effectively prevents the selection of high-frequency generic terms based solely on frequency. Its core principle is that a term’s weight increases with its frequency in a document but decreases with its prevalence across the document corpus [?]. For a document set $D = \{d_i | i = 1, 2, \dots\}$ and subject term set $W = \{w_j | j = 1, 2, \dots\}$, where $D_{w_j} = \{d_i | i = 1, 2, \dots\}$ represents documents containing term w_j , the TF-IDF value for term w_j is calculated as:

$$\text{TF-IDF}_{w_j} = \sum_{d_i \in D_{w_j}} \left(\frac{n_{w_j}}{n_{d_i}} \times \lg \frac{N_D}{N_{D_{w_j}}} \right) \quad (1)$$

where n_{w_j} denotes the frequency of term w_j in document d_i , n_{d_i} represents the total number of terms in document d_i , N_D is the total number of documents in corpus D , and $N_{D_{w_j}}$ is the number of documents containing term w_j .

Combining TF-IDF with the traditional word frequency g-index concept, we propose the TI-g index. The TI-g index is defined as: when subject terms are sorted by TF-IDF values in descending order, if the cumulative TF-IDF value of the top g terms is no less than g^2 , while the cumulative TF-IDF value of $g+1$ terms is less than $(g+1)^2$, then the top g terms constitute the high-frequency subject terms. The TI-g index calculation proceeds as follows:

1. Calculate TF-IDF values for all terms w_j in set W using equation (1).
2. Sort terms w_j in set W by TF-IDF values in descending order, assigning each term w_j a rank r_j .
3. Sequentially accumulate TF-IDF values; the cumulative sum for term w_j is $\sum_{k=1}^{r_j} \text{TF-IDF}_k$.
4. Compute the square of each term’s rank: r_j^2 .
5. When the absolute difference between the cumulative TF-IDF sum and the squared rank is minimized, the corresponding rank r represents the TI-g index, indicating that the top r terms are high-frequency subject terms. The TI-g index formula is:

$$\text{TI-g} = \arg \min_r \left| \sum_{k=1}^r \text{TF-IDF}_k - r^2 \right| \quad (2)$$

2.4 Construction of Ternary Subject Term Groups Complex network-based literature topic mining is essentially co-word analysis, which pairs subject terms (SH1, SH2) to form binary term groups. By traversing the frequency W at which (SH1, SH2) co-occur in document d_i , we use this frequency as weight to form ternary term groups (SH1, SH2, W), ultimately constructing a subject term network.

As noted earlier, mining research hotspots in a field builds upon topic identification in individual documents, typically relying solely on keywords. However, since keywords are author-assigned, they suffer from subjectivity and semantic ambiguity, limiting the effectiveness of word frequency and co-word analysis [?]. To address this, this study comprehensively analyzes titles, keywords, and abstracts, avoiding the limitations of keyword-only extraction and ensuring more complete and reliable topic identification.

There are six possible co-occurrence scenarios for subject terms SH1 and SH2 within a document's title, abstract, and keywords, each carrying different weights. To determine these weights, we employed the AHP method using yaahp software. Twelve experts evaluated pairwise comparison matrices using a 1-9 scale (consistency ratio CR = 0.0176), establishing the relative importance of titles, keywords, and abstracts as 0.47 (w_t), 0.34 (w_k), and 0.19 (w_a), respectively. Based on these weights, we calculated co-occurrence weight coefficients for binary term groups under different scenarios, as shown in Table 1 .

The weight W for ternary term groups (SH1, SH2, W) is calculated as:

$$W = \sum_{i=1}^n w_i \quad (3)$$

where n represents co-occurrence frequency and w_i denotes the weight coefficient for each co-occurrence scenario.

2.5 Complex Network-Based Research Topic Mining Research hotspots are not composed of single subject terms but rather clusters of closely connected terms, analogous to communities in complex networks [?]. Therefore, community detection algorithms in complex networks can be used to mine literature hotspots. Analyzing small-world and scale-free properties reveals network characteristics, while modularity Q values evaluate community partition quality.

A network exhibits stronger small-world properties when its average clustering coefficient is larger and average path length is smaller. In practice, we compare these metrics with those of random networks of the same scale. Let C_a and L_a denote the actual network's average clustering coefficient and average path length, while C_r and L_r represent those of a random network. If equa-

tion (4) yields a value greater than 1, the actual network possesses small-world characteristics, with larger values indicating more pronounced properties [?]:

$$\frac{C_a/C_r}{L_a/L_r} > 1 \quad (4)$$

Let $P(k)$ represent the frequency of nodes with degree k . If $P(k)$ follows a power-law distribution, the network exhibits scale-free properties, emphasizing unequal resource distribution among nodes. The power-law exponent γ in equation (5) typically ranges between 2 and 3:

$$P(k) \propto k^{-\gamma} \quad (5)$$

Modularity Q evaluates community partition quality, with values closer to 1 indicating better community structure. Networks with clear community structures generally have Q values between 0.3 and 0.7 [?]:

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(i,j) \quad (6)$$

where m is the total number of edges (sum of edge weights), A_{ij} is the adjacency matrix representing edge weights between nodes i and j , k_i and k_j are the degrees of nodes i and j , and $\delta(i,j)$ equals 1 if nodes i and j belong to the same community, otherwise 0.

3. Results and Analysis

3.1 Evolution of Research Topics in China's Green Consumption Field

Using the proposed TI-g index, we mined high-frequency subject terms from green consumption literature between 2010 and 2022, visualizing topic evolution through a heatmap shown in Figure 3 [Figure 3: see original paper].

Overall, green marketing, low-carbon economy, green consumption concepts, and ecological civilization have remained persistent research hotspots in China's green consumption field. Before 2016, research primarily focused on foundational concepts (upper-left of Figure 3), with environmental themes such as sustainable development, circular economy, low-carbon economy, environmental protection, environmental issues, and legal systems dominating the discourse. After 2016, research shifted noticeably toward practical applications (lower-right of Figure 3), including green development, psychological studies on specific groups (e.g., college students) regarding green purchase intentions and influencing factors, as well as technical aspects like green design, green production, and green technology innovation.

3.2 Hot Research Topics in China's Green Consumption Field To deeply mine recent research hotspots, we constructed a subject term network for China's green consumption field over the past five years (2018-2022) based on complex network analysis. The subject term network is an undirected graph; detailed network metrics are presented in Table 2 .

3.2.1 Subject Term Network Characteristics Analysis (1) Small-world characteristics analysis

According to Table 2, the subject term network's average clustering coefficient C_a and average path length L_a are 0.614 and 1.906, respectively. For a random network of the same scale, the equivalent average clustering coefficient C_r and average path length L_r are approximately 0.104 and 1.898. Using equation (4), we calculate $5.879 \gg 1$, confirming that the subject term network exhibits small-world characteristics. This indicates that distances between any two nodes are relatively short while tightly connected communities or thematic domains exist within the network.

(2) Scale-free characteristics analysis

The degree distribution of subject term network nodes is shown in Figure 4 [Figure 4: see original paper]. In the log-log coordinate system, the linear relationship is weak, with a coefficient of determination R^2 of only 0.292, indicating that the degree distribution does not follow a power-law distribution. According to equation (5), the subject term network lacks scale-free characteristics, meaning most nodes have similar connection counts without prominent "super nodes."

3.2.2 Community Detection and Analysis in the Subject Term Network We used Gephi for community partition and the OpenOrd algorithm for visualization layout. At standard resolution, four communities were identified (represented by four colors), as shown in Figure 5 [Figure 5: see original paper]. Using Gephi's default algorithm by Blondel et al. [?], the modularity Q value is 0.289. Based on community visualization and equation (6), the subject term network does not exhibit a clear community structure, with substantial cross-fertilization among research thematic domains and no relatively independent research systems or directions. Nevertheless, we can broadly categorize green consumption research hotspots into four thematic domains, with core terms for each domain listed in Table 3 .

Based on the core terms in each domain, we define them as: (1) Green consumption driving factors research, (2) Green consumption values cultivation research, (3) Green consumption institutional mechanisms research, and (4) Other research.

Green consumption driving factors research examines how consumer green consumption behaviors and lifestyles are formed and identifies their driving factors. These include internal factors such as perceived value, income level,

and environmental awareness, as well as external factors like subjective norms and promotional intensity.

Green consumption values cultivation research investigates how consumers, particularly specific groups like college students and high school students as future consumption mainstays, perceive concepts such as green consumption and human-nature harmony, along with their promotion, education, and cultivation. This domain integrates Marxist ecological views, ideological and political education, and rural revitalization. Further literature review reveals that this research area focuses more on the significance, forms, and methods of values cultivation, with limited studies on evaluation of cultivation effectiveness.

Green consumption institutional mechanisms research explores the relationship between green consumption and the environment, including legal systems, tax policies, evaluation indicators, incentive mechanisms, and related technological innovation mechanisms.

Other research represents applied studies of green consumption concepts in other fields, such as green consumption credit and green consumption finance.

4. Discussion and Conclusions

This study employs text mining technology and complex network analysis with targeted optimizations to mine research topic evolution since 2010 and hotspots since 2018 in China's green consumption field, providing references for future research.

Methodological contributions: Addressing the subjectivity and semantic ambiguity of using literature keywords directly as subject terms, we propose a method that comprehensively considers titles, abstracts, and keywords, using text segmentation to extract subject terms and AHP to determine co-occurrence weights for binary term pairs. To overcome the inability of traditional word frequency g-index to eliminate high-frequency generic terms, we optimize it using the TF-IDF algorithm and propose the TI-g index.

Empirical findings: (1) Before 2016, China's green consumption research focused primarily on foundational concepts related to the environment, such as sustainable development, circular economy, environmental protection, and legal systems. After 2016, research shifted toward practical applications, including green economic development, green purchase intentions of specific groups (e.g., college students), influencing factors of green purchase behavior, and technical aspects like green design, production, and innovation. (2) Since 2018, China's green consumption research can be categorized into four thematic domains: driving factors research, values cultivation research, institutional mechanisms research, and other research. However, complex network analysis reveals that these four domains lack clear community structure and scale-free properties. Therefore, future research should strengthen depth, explore methodological and

interdisciplinary approaches, and develop more comprehensive research systems. (3) At the micro level, each thematic domain exhibits certain research gaps. For example, in driving factors research, methodologies and models are overly uniform, with most studies employing SEM or traditional statistical methods while lacking cutting-edge big data techniques such as text mining, sentiment analysis, and public opinion analysis. Values cultivation research shows insufficient investigation into differentiated cultivation and evaluation methods based on population characteristics. Furthermore, issues related to legal frameworks, tax policies, and practical concerns like green technology, production, and design require deeper investigation.

References

- [?] Elkington J, Hailes J. The green consumer guide: from shampoo to champagne-highstreet shopping for a better environment[M]. London: V. Gollancz, 1988.
- [?] Cui Qiaohuan. Analysis of influencing factors and countermeasures for implementing green consumption in China[J]. Theory Guide, 2007(10): 116-118.
- [?] Hou Haiqing, Yin Li. A review of research on influencing factors of green consumption behavior[J]. China Economic & Trade Herald (Middle), 2021(6): 149-152.
- [?] China Ecological Civilization Research and Promotion Association. Research report on the current status of green consumption among the Chinese public[R]. 2019.
- [?] China Chain Store & Franchise Association, Alibaba New Service Research Center. Toward a new service era—report on the digital development of life service industry (2021)[R]. 2021.
- [?] Ma Weichen, Deng Xu. Research on policy measures for green consumption in China[J]. Environmental Protection, 2017, 45(6): 56-59.
- [?] The 14th Five-Year Plan for National Economic and Social Development and Long-Range Objectives Through the Year 2035 of the People's Republic of China[M]. Beijing: People's Publishing House, 2021.
- [?] Liu Yongsheng, Gan Yingying. Analysis of research status and trends in China's green food field from a social network perspective[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2020, 41(2): 26-34.
- [?] Gao Yang, Xiong Juhua, Wu Hao, et al. Keywords in 2021 NSFC proposals reveal frontier hotspots and development directions in geographic sciences[J]. Scientia Geographica Sinica, 2022, 42(1): 15-30.
- [?] Liu Junnan, Liu Haiyan, Chen Xiaohui, et al. Visual analysis of research hotspots in the surveying and mapping field[J]. Science Technology and Engineering, 2019, 19(32): 43-51.
- [?] Zheng Wei, Liu Yulin. Research on hotspots of university teachers' job burnout based on complex networks[J]. Heilongjiang Researches on Higher Education, 2020, 38(6): 112-118.
- [?] Tan Zhangu, Peng Shengnan, Wang Zhaogang. Research on hotspots and trends of domestic text mining based on cluster analysis[J]. Journal of the

- China Society for Scientific and Technical Information, 2019, 38(6): 578-585.
- [?] Zhang Ruofan, Shen Yiran, Liu Zehua. Research hotspots and evolution in China's management field based on complex networks[J]. *Statistics & Management*, 2019(8): 104-109.
- [?] Dai Ruwei, Li Yaodong. Hall for workshop of metasyntetic engineering and system complexity[J]. *Complex Systems and Complexity Science*, 2004(4): 1-24.
- [?] Jiang Minqin, Shi Xiaojing, Yao Anyang, et al. Analysis of research hotspots, knowledge evolution and trends in complex networks based on CSSCI data[J]. *Journal of Wuhan Business University*, 2022, 36(2): 74-79.
- [?] Jin Junbao, Qu Jiansheng, Wu Xinnian, et al. Research on complex network characteristics of inter-provincial mobility of high-level scientific and technological talents in China[J]. *Science and Technology Management Research*, 2021, 41(21): 112-118.
- [?] Han P, Yinzheng L, Changxi M. Topology analysis of Lanzhou public transport network based on double-layer complex network theory[J]. *Physica A: Statistical Mechanics and its Applications*, 2022(592): 126694.
- [?] Hu Jun, Wang Yutong, He Xinwei, et al. Analysis and application of global aviation network structure based on complex networks[J]. *Computer Science*, 2021, 48(S1): 321-325.
- [?] Overbye T J, Shetye K S, Hutchins T R. Power grid sensitivity analysis of geomagnetically induced currents[J]. *IEEE Transactions on Power Systems*, 2013(28): 2821-2828.
- [?] Fu Jie, Zou Yanli, Xie Rong. Identification of critical lines in power networks based on complex network theory[J]. *Complex Systems and Complexity Science*, 2017, 14(3): 91-96.
- [?] Wu Desheng, Cao Yuan, Tang Can, et al. Research on debt risk and risk contagion networks under classified management and control[J]. *Management World*, 2021, 37(4): 35-54.
- [?] Xie Chi, He Huimin, Wang Gangjin, et al. Research on extreme risk spillover effects and evolution in pan-financial markets based on complex networks[J]. *Systems Engineering - Theory & Practice*, 2021, 41(8): 1926-1941.
- [?] Angela L, Nicola A, Alfonso M, et al. Complex network modelling of origin-destination commuting flows for the COVID-19 epidemic spread analysis in Italian lombardy region[J]. *Applied Sciences*, 2021, 11(10): 4381.
- [?] Mukul G, Rajhans M. Spreading the information in complex networks: Identifying a set of top-N influential nodes using network structure[J]. *Decision Support Systems*, 2021(149): 113608.
- [?] Jia Hongyu, Zhao Xueyan, Qiu Chenzi. Research on analysis methods of microblog public opinion maps based on complex networks[J]. *Modern Information*, 2015, 35(3): 64-67.
- [?] Peng Cheng, Qi Kai, Li Bingxue. Research on public opinion dissemination and early warning mechanisms in complex networks based on SIR-EGM model[J]. *Information Science*, 2020, 38(3): 145-153.
- [?] Sun Haisheng. Empirical study on literature coupling networks based on Web of Science database[J]. *Journal of Intelligence*, 2018, 37(10): 201-207.

- [?] Wang Yanpeng, Han Tao, Chen Fang. Research on key technology identification methods integrating literature knowledge clustering and complex networks[J]. Library and Information Service, 2020, 64(16): 105-113.
- [?] Xin Juanjuan, Cao Jia. Literature hotspot mining and visualization based on complex networks[J]. Computer Engineering and Applications, 2016, 52(12): 106-110.
- [?] Yang Qian. Comparative study on analysis functions of common bibliometric tools[J]. Information Research, 2021(10): 87-93.
- [?] Zhong Huixin. A review of emerging trend detection research[J]. Modern Information, 2017, 37(12): 162-167.
- [?] Holeab C, Paunica M, Curaj A. A complex method of semantic bibliometrics for revealing conceptual profiles and trends in scientific literature. The case of future-oriented technology analysis (FTA) science[J]. Economic Computation and Economic Cybernetics Studies and Research, 2017, 51(2): 23-37.
- [?] Wang Y, Luo H, Shi Y. Complex network analysis for international talent mobility based on bibliometrics[J]. International Journal of Innovation Science, 2019, 13(11): 419-435.
- [?] Ortega J, Aguillo I. Institutional and country collaboration in an online service of scientific profiles: Google Scholar Citations[J]. Journal of Informetrics, 2013, 2(7): 394-403.
- [?] Chae C, Yim J H, Lee J, et al. The bibliometric keywords network analysis of human resource management research trends: the case of human resource management journals in South Korea[J]. Sustainability, 2020, 12(14): 5700.
- [?] He Bo, Zhao Haiyuan, Yuan Yue. A longitudinal study of the evolution of research trends in China's manager field over 28 years[J]. Journal of Chongqing University (Social Science Edition), 2016, 22(3): 100-108.
- [?] Liu Yongsheng, Gan Yingying. Analysis of research status and trends in China's green food field from a social network perspective[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2020, 41(2): 26-34.
- [?] Du Xianyun, Gao Linxia, Xu Zhiying. Research on domestic green consumption behavior based on social network analysis and co-word analysis[J]. Computer Programming Skills & Maintenance, 2018(6): 20-23.
- [?] Mo Fuchuan, Lou Cequn. Theoretical basis and path exploration of applying highly cited papers to research hotspot identification[J]. Information Studies: Theory & Application, 2019, 42(4): 59-63.
- [?] Chen Guiwu. International research trends in library and information science: An empirical analysis based on SSCI research papers from 2000-2009[J]. Journal of Library Science in China, 2011, 37(1): 73-82.
- [?] Yang Aiqing, Ma Xiufeng, Zhang Fengyan, et al. Research on the application of g-index in subject term selection for co-word analysis[J]. Journal of Intelligence, 2012, 31(2): 52-55.
- [?] Salton G., McGill M. J. The SMART retrieval system—Experiments in automatic document processing[J]. Information Storage and Retrieval, 1971, 6(4): 209-242.
- [?] Song Jinghua, Gong Xiandian. Community spatial network evaluation and optimization based on complex network analysis[C]//Architectural Digital

Technology Teaching Committee of the National Architectural Education Steering Committee. Digital Intelligence Construction: Proceedings of the 2020 National Conference on Digital Technology Teaching and Research in Architecture Departments. Changsha: China Architecture & Building Press, 2020: 39-45.

[?] Newman M, Girvan M. Finding and evaluating community structure in networks[J]. Phys. Rev. E, 2004, 69(2): 026113.

[?] Blondel V D, Guillaume J L, Lambiotte R, et al. Fast unfolding of communities in large networks[J]. Journal of Statistical Mechanics: Theory and Experiment, 2008, 2008(10): P10008.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.