DSI DATA SCIENCE AND INFORMETRICS
VOL.1, NO.2, MAY. 2021

Research on identification of potential knowledge growth points in information science

Changling Li*, Guoyang Rong, Yanxin Pai, Dehui Du, Lu Xu, Qingqing Fan, Fuzhong Xiang

Institute of Information Management, Shandong University of Technology, Zibo, China

ABSTRACT

The scientific system's complexity makes it impossible to solve social problems by a single discipline independently, and interdisciplinary knowledge cooperation and innovation become an indispensable research mode of modern science. Identifying the potential interdisciplinary knowledge association is the key to promoting interdisciplinary cooperation. In this paper, based on analyzing the growth points of science, "knowledge growth point" is defined as the growth point of science that produces new knowledge, and its fundamental attributes and evaluation indexes have been analyzed. In contrast, the "interdisciplinary knowledge growth point" is defined as the introduction of related interdisciplinary concepts, theories, techniques, and methods, to conduct integrated research of key knowledge points of active disciplines, to generate growth point of innovative knowledge, and analyze its related research status. The identification of "potential interdisciplinary knowledge growth points" is helpful to promote knowledge innovation. Therefore, it is intended to analyze the identification methods of the generation of key knowledge nodes of the element of disciplines and interdisciplinary related knowledge, and explore quantitative and qualitative consultation to identify potential interdisciplinary knowledge growth points.

KEYWORDS

Knowledge Growth Point; Potential Interdisciplinary Knowledge Growth Points; Identification Methods

Introduction

Growth point is a scientific term derived from botany, which initially refers to the apical meristem of roots and stems. Furthermore, the growing point's cell division activity is vigorous (The Biology Editorial Committee of EOC, 1991). Natural science defines scientific growth as closely related to a certain thing and has an apparent inheritance or innovation relationship with it (Baidu Encyclopedia, 2020). The scientific growth point has active growth factors and conditions for generating new disciplines and new knowledge (Zhang, 1986). Therefore, "knowledge growth point" is defined as the scientific growth point that produces new knowledge, and "discipline growth point" is defined as the scientific growth point that produces new disciplines. The generation of knowledge growth points is in Figure 1.

* Corresponding author: lichl69@163.com

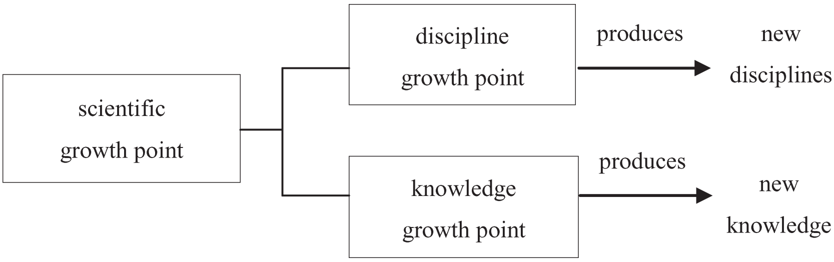

Figure 1 Generation of knowledge growth points

Since knowledge growth point is a scientific growth point that produces new knowledge, it has the same essential nature as the scientific growth point (Zhang, 1986): (1) Advanced nature, at the forefront of disciplines; (2) critical nature, at the key link of discipline development; (3) comprehensiveness, occurring at the intersection of disciplines. Thus key knowledge nodes such as discipline frontiers, discipline hotspots, and discipline intersections are the most likely knowledge growth points.

The conditions for generating key knowledge nodes, including discipline frontiers, discipline hotspots, and discipline intersections, to become knowledge growth points is to complement the intradisciplinary or interdisciplinary knowledge. Cooperation with interdisciplinary knowledge is more conducive to scientific research and innovation. Therefore, the "intradisciplinary knowledge growth point" is the knowledge growth point generated by collaborating with knowledge within the discipline. Furthermore, the "interdisciplinary knowledge growth point" is the knowledge growth point generated by collaborating research with interdisciplinary knowledge (Li et al., 2020). The generation of interdisciplinary knowledge growth points is in Figure 2.

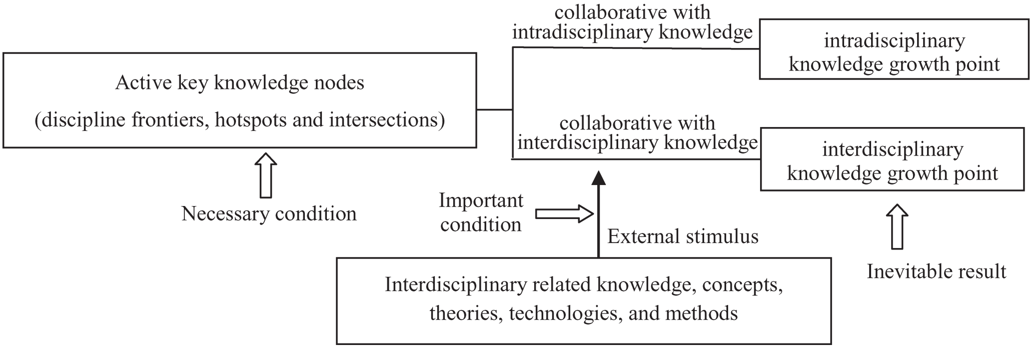

Figure 2 Generation of interdisciplinary knowledge growth points

Thus interdisciplinary knowledge growth point is defined as the knowledge growth point created by active key knowledge nodes and the introduction of interdisciplinary related knowledge, concepts, theories, technologies, and methods. Active key knowledge nodes such as discipline frontiers, hotspots, and intersections are necessary conditions for its creation. Interdisciplinary related knowledge, concepts, theories, technologies, and methods are important conditions. Moreover, the interdisciplinary knowledge growth point is the inevitable result of collaborative research on key knowledge nodes and interdisciplinary related knowledge.

We have published a paper on the definition and features of the above concepts (Li et al., 2020).

Literature review

The research and methods on the key knowledge nodes of discipline and their interdisciplinary knowledge are shown in the following aspects:

(1) Identifying key knowledge nodes of discipline by citation analysis method. Citations include the vertically inherited citation/cited relationship and the horizontal co-citation/co-cited relationship. This crisscross and mutual citation relationship structure is called a citation network. Key knowledge nodes, such as core nodes and bridge nodes in the intra-discipline citation network, represent the discipline's research hotspots and frontiers. The common research content of the interdisciplinary cross-citation network represents the interdisciplinary research theme. Thus many scholars identify key knowledge nodes of discipline by citation analysis methods. For example, use the frequency of keyword citations, combined with indicators to identify research hotspots by methods such as citation, co-citation, and coupling analysis. The indicators include the number of recent documents, average publication year, etc. (Small & Sweeney, 1985; Vlachy, 1984a', 1984b). CiteSpace, which is based on the citation frequency of scientific papers and the citation analysis theory, has been designed to detect discipline research hotspots (Chen, 2005). The method of analyzing highly cited papers has been used to identify research hotspots and development trends in information science (Zhu & Leng, 2014; Ye, 2007). A sub-network of highly cited paper relationships, which used network clustering and text mining methods, is constructed to determine the discipline's frontier (Fajardo-Ortiz et al., 2017). The citation network is constructed by a large-scale network clustering algorithm, which can identify key research directions of the discipline from the perspective of network topology (Deng & Wang, 2018). A new index calculation method, which improves the z-index based on the time factor, is applied in hot spot identification research. The method analyzes emerging research hotspots' changes and divides three types of research hotspots with different development trends: upward, stable, and downward (Li et al., 2019). The disciplinary cross-citation matrix method is used to measure the interdisciplinary relationships and evolutionary laws of biological sciences (Meyer, 2010).

(2) Identifying key knowledge nodes of discipline by co-word analysis method. Keywords represent the paper's research theme and can directly reflect the core research content of the paper. The set of keywords includes the main research content of a certain discipline area. The cross-keywords of two or more disciplines can reflect the research themes of the interdisciplinary. Therefore, co-word analysis is an informatics method that counts the frequency of keyword pairs appearing in the same article and builds a common word matrix. The method identifies the hotspots and frontiers of disciplines research and discovers interdisciplinary research topics through word frequency analysis, cluster analysis, association analysis, sudden word frequency analysis etc. For example, six knowledge networks are constructed to identify emerging topics in economics (Ma & Liu, 2014). A new measurement index TI, based on the high-frequency word co-occurrence network, is used to explore the intersection of information science and related disciplines (Xu et al., 2015). The method of screening hot keywords for rapid growth and decline in a co-word network is used to identify the changes in discipline hotspots (Xu & Bi, 2019). CFinder, an overlapping community analysis tool, is used to identify the intersection of information science and computer science (Li et al., 2013). The LDA topic model is introduced to identify cross-research topics between transition economies and emerging markets (Elkin & Anke, 2015). A multi-modal topic network is constructed to identify interdisciplinary research topics between digital libraries and related disciplines (Shang, 2018).

(3) Identifying key knowledge nodes of discipline by altmetrics method. Driven by the mature development of online academic resource platforms and online social media, not only has the academic communication patterns changed dramatically, but the data of online media has grown exponentially (Li et al., 2018). Scholars actively adopt changes in the number of keywords in online social media and Altmetrics data to identify the discipline's frontiers and hotspots. For example, the irregular scores of blogs and the changes in the number of keywords in blog posts are used to identify emerging words (Takahashi et al., 2014). A discipline hot spot discovery, tracking, and analysis mechanism are proposed to find hot topics in the text data of Sina Weibo (Sheng, 2012). The reply posts on the "Essence Page" of the "Informatics" topic on Zhihu's website are used as a sample to explore the discipline of high concern for the user group (Lai et al., 2017). The real-time download information from the scientific literature is also used as a sample to track research trends in a certain field, find research hotspots, and detect research frontiers (Wang et al., 2014). A data screening and filtering mechanism based on seven indicators in Altmetrics, is established to identify research hotspots and cutting-edge topics (Zhao & Yang, 2016). Topic attention heat and intensity monitoring indicators, based on social media data, are constructed to identify emerging themes of disciplines with high attention growth potential (Duan & Pan, 2017). The zt index model, which replaced the citation frequency in the z-index by Altmetrics Attention Score value (AAS), is constructed to identify high-profile research topics (Pai et al., 2019).

Which concepts, technologies, and methods of which disciplines can cooperate with key knowledge nodes in interdisciplinary research to promote discipline development and social progress? This question is a prerequisite for interdisciplinary cooperation, and it is also a difficult problem for researchers in interdisciplinary cooperation. The discovery of interdisciplinary knowledge is the basis of interdisciplinary collaborative research.

"Strength of weak ties" originated in the field of sociological research and is defined as a short-term social contact between two actors. Professor M. Granovetter, who proposed this concept, pointed out that strong ties make the organization's internal ties close and stable. In contrast, weak ties provide an important way for information exchange between different groups and organizations so that isolated subgroups begin to establish connections. Moreover, with the continuous strengthening of this weak connection, the scope of different information exchanges has been further expanded. The dissemination, integration, development, and innovation of information has been accelerated (Granovetter, 1973). This view is also discussed by other scholars in similar researches (Bakshy et al., 2012; David & Jon, 2012; Onnela et al., 2007; Zhao et al., 2010). The research results show that: compared with strong connections, weak ties can transmit more potential, diversified, and non-redundant knowledge resources (Genuis, 2005). Many studies have also found that weak connections are easier to establish more extensive relationships between different individuals in other subnets of the knowledge network, which has a positive effect on promoting scientific research cooperation (Abbasi et al., 2011; Bettoni et al., 2008; Yang et al., 2009). Although weak ties are not as strong as strong ties, they have extremely fast, possibly low-cost, and high-efficiency transmission efficiency.

Therefore, we believe that the key knowledge nodes of disciplines and interdisciplinary related knowledge often exist in weak ties, whether it is when new interdisciplinary knowledge growth points begin to appear or before they appear. Moreover, weak ties can be used as an entry point for interdisciplinary knowledge discovery. Both the co-word network and the citation network have weak connections between knowledge, and both can be used to identify interdisciplinary related knowledge in theory. At present, relevant research mainly focuses on the identification of weak relationships in the keyword co-occurrence network.

(4) Identifying key knowledge nodes of discipline by weak ties in the co-word network. Non-relevant literature knowledge discovery methods were proposed in 1986, originated from biomedical literature, and mainly used for related knowledge discovery in biomedical co-word networks (Swanson & Fish, 1986). For example, the symptom-mediated co-occurrence relationship is used to analyze the drug-disease interaction and effect (Song et al., 2018). Scholars also apply the non-relevant literature knowledge discovery methods in other disciplines. For example, the co-word network was used to discover the connection between the World Trade Organization and fundamentalism in agricultural economics (Huang & Ma, 2009). The weak co-occurrence network of high-frequency words is constructed to analyze the interdisciplinarity mode and explore the interdisciplinarity of information science at the micro-level (Wei et al., 2015). The discipline correlation analysis method has been improved to extract the common themes and the respective independent themes of the two disciplines (agronomy reproductive biology and veterinary science) and combined with the correlation measurement method to quantify the correlation between independent disciplines (Wu et al., 2017). The non-relevant literature knowledge discovery method of the co-word network was used to identify relevant knowledge across disciplines, and the knowledge pairing between information science and computer science was discovered for the first time. The results of identification include think tanks and CMB algorithms, knowledge services and mouse behavior, etc.(Li et al., 2018; Liu et al., 2017)

(5) Knowledge key identifying nodes of discipline by weak ties in the citation network. There are not only weak relationships in the co-word network but also weak relationships in the citation network, such as the relevance of keywords in the source document and reference, source document and cited document, co-citation and co-cited, etc. The comparison found that citations can identify more relationship pairs than co-occurrences, and the pairing relationships discovered in this way are more unique and diverse (Song et al., 2018). However, there are currently few researches based on citation identification-related knowledge, especially interdisciplinary related knowledge. We have published related research. For example, source-references and source-cited correlation analysis methods are proposed to identify related interdisciplinary authors by analyzing interdisciplinary authors' five citation relationships (Li et al., 2018). The Index of Discipline-related Novelty (IDN), based on extracting key node keywords-interdisciplinary reference data, is constructed to identify interdisciplinary keywords with high cooperation potential. Recognition results include a temporal and spatial data model, policy process theory and random forest algorithm, etc.(Du et al., 2020)

Methods and Results

We have published researches on the recognition of discipline key knowledge nodes and interdisciplinary related knowledge. The research methods and conclusions are as follows.

Identification of Key Knowledge Nodes in Highly Cited Disciplines Based on Time Factor

Z-index is a new type of evaluation index that multiplies the equal weights of the citation frequency, the average citation frequency, the consistency of citation distribution, and other indicators for comprehensive evaluation. It can comprehensively measure the estimated objects from quantity, quality, as well as the consistency of quantity and quality (Prathap, 2014). The calculation formula is as follows:

$$\mathrm{z} = \left[c \times \frac{c}{n} \times \frac{\frac{c^2}{n}}{\sum_{k=1}^{n} {c_k}^2}\right]^{\frac{1}{3}} \qquad (1)$$

Subsequently, Professor G. Prathap demonstrated the applicability of the z-index in scientific evaluation from the perspectives of countries, institutions, scholars, and journals. We tried to introduce the time factor of the old and new degree of the paper as the citation frequency weight, improve the z-index method, and identify the research hotspots of the discipline. In order to effectively identify the recent and long-term novel and continue the research hotspots of the discipline from the perspective of cited quantity-quality-distribution consistency. The z-index provides an important reference for scholars to grasp the field's research content (Prathap, 2014c; Prathap & Gangan, 2017; Prathap, 2014a, 2014b).

Introduce a time factor, which is expressed as the proportion of time lag in a cumulative time lag. Different weights are assigned to the citation frequency of academic papers within a certain time range of the discipline, according to the degree of newness of the publication time. The time factor is defined as follows (Yu & Guo, 2019).

$$\mathrm{TF}_{t-\mathrm{i}} = \frac{y-i+1}{\sum_{i=1}^{y}(1+2+\cdots+y)} = \frac{y-i+1}{\frac{y(y+1)}{2}} \qquad (2)$$

In formula 2, t represents the annual statistical time; i represents the time lag from the statistical year t; y represents the length of the research time range (t-y, t-1), i∈[1,y]. For example, if the statistical year is 2019, and the research scope is sample data from 2014 to 2018 for five years, t=2019, y=5, and then i∈[1,5].

The total citation frequency of a research topic Cy (tf), which based on the time factor, is expressed as:

$$C_y(\mathrm{tf}) = \sum_{i=1}^{y} \mathrm{TF}_{t-i} \times C_{\mathrm{t-i}} \qquad (3)$$

Based on the time factor to improve z-index to zy (tf) index, the expression is as follows:

$$z_y(\mathrm{tf}) = \left[C_y(\mathrm{tf}) \times \frac{C_y(tf)}{N} \times V_y(\mathrm{tf})\right]^{\frac{1}{3}} = \left[C_y(\mathrm{tf}) \times \frac{C_y(\mathrm{tf})}{N} \times \frac{\frac{C_y(\mathrm{tf})^2}{N}}{\sum_{K=1}^{N} C_K\,(\mathrm{tf})^2}\right]^{\frac{1}{3}} = \left[\frac{\frac{C_y(\mathrm{tf})^4}{N^2}}{\sum_{k=1}^{N} C_K\,(\mathrm{tf})^2}\right]^{\frac{1}{3}} \qquad (4)$$

Under the action of the time factor, the zy(tf) index improves the defect of the traditional z-index for equal weighting of the citation frequency of academic papers with different publication times, making papers of newer years have a larger weight for citations, reducing the number of citations. The limitation of the cumulative effect of citation frequency over time increases the possibility of new literature highlighting research hotspots. Therefore, the research hotspots of zy (tf) index ranking are rising rapidly, which are emerging frontier hotspots that are frequently cited in new literature. Different types are classified according to the difference between the z-index and the zy (tf) index ranking changes. The specific classification criteria are shown in Figure 3.

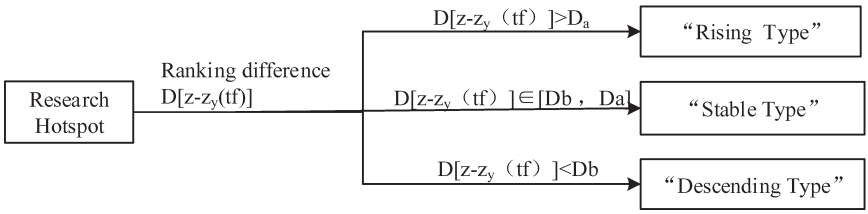

Figure 3 Classification model of disciplinary research hot topics based on ranking order difference

The research topics distributed in the new literature, due to the assignment of higher time factor weights, the citation frequency value increases, and the z_y(tf) value increases, the ranking premise, and the ranking number is smaller than the z-index ranking number. Therefore, the ranking order calculated results of difference D[z-z_y(tf)] between z-index and the z_y(tf) index is used to reflect the degree of research interest in the topic. If D[z-z_y(tf)] is greater than the positive threshold D_a, the ranking of the z_y(tf) index is rising rapidly, which means the popularity of research topics is increasing. Moreover, if D[z-z_y(tf)] is less than the negative threshold D_b, the z_y(tf) index ranking dropped rapidly, which means the popularity of research topics decreased.

We choose domestic information science as an example for demonstration. After data processing and calculation, the research results are shown in Table 1.

Table 1 Classification results of different types of domestic information science research hotspots in 2018

Research Hotspots	Classification Criteria	The Research Topics
" Rising Type" of Research frontier	D [z–z_y (tf)] >3	Think –tank, Altmetrics, Social media, The network public opinion, Micro, User behavior, Knowledge service, Health information, Privacy, Internet of things, Microblog, Journal evaluation, Scientific data, Interdisciplinary, Information behavior, Core author, Information security, Research collaboration, Highly cited papers
" Stable Type" of Research frontier	D [z–z_y (tf)] ∈ [–2,3]	Big data, Citation analysis, Empirical study, Impact factors, Cloud computing, Knowledge management, The citation frequency, Open access, Information resources, Knowledge sharing, Substance, Complex network, Visual analysis
" Descending Type" of Research frontier	D [z–z_y (tf)] <–2	Competitive intelligence, Co–word analysis, Emergency, Sentiment analysis, Information service, Social network analysis, Text mining, Academic influence

We have published the derivation process and verification of the improved z-index based on the time factor (Li et al., 2019).

Identification of Key Knowledge Nodes of High Attention Discipline Based on AAS Value

The Altmetrics Attention Score indicator is a relatively mature and widely recognized Altmetrics measurement algorithm (Kuang & Zhang, 2019). It is an aggregated index obtained by weighting all Altmetrics indicators on the Altmetric.com comprehensive platform through a certain algorithm. These indicators include the number of news reports, the number of blog posts, the number of Mendeley readers, the number of peer reviews and other 17 online policy documents and mainstream media, online social media, online reference management tools, and other online sources of use, download, promotion and other data. (Altmetric. com, 2020). In 2017, Altmetric.com incorporated the citation frequency from Web of Science into its index system. Therefore, the AAS value is a comprehensive reflection of the traditional evaluation of scientific literature and the evaluation of online media, which can more comprehensively evaluate the academic level of scientific literature and the online attention received after publication.

Based on the z-index method, we propose the zt index to identify academic papers' research topics with high AAS attention. The calculation formula is:

$$z_t = \left[S(T_i) \times \frac{S(T_i)}{n_i} \times C(T_i)\right] = \left[S(T_i) \times \frac{S(T_i)}{n_i} \times \frac{\frac{S(T_i)^2}{n_i}}{\sum_{j=1}^{n_i}\left[S_j(T_i)\right]^2}\right]^{\frac{1}{3}} = \left\{\frac{[S(T_i)]^4}{{n_i}^2 \times \sum_{j=1}^{n_i}\left[S_j(T_i)\right]^2}\right\}^{\frac{1}{3}} \quad (5)$$

If a topic Ti is distributed in ni academic papers, then:

$S_j(T_i)$ is the single AAS score of topic Ti in the jth paper, j=1,2...ni; $S(T_i)=\sum_{j=1}^{ni} S_j(T_i)$ is the total AAS score of the research topic T_i, describing the overall scale of the attention of topic T_i.

$\frac{S(T_i)}{ni}$ is the average article score, which reflects the average attention level of topic T_i at a single article level. Furthermore, it is an index describing the quality of topic T_i attention.

$C(T_i)$ is the distribution consistency index, which is used to describe the distribution consistency of the topic T_i in the papers of its distribution, and $C(T_i) \in (0,1]$ is used as an adjustment factor to solve the extreme value problem. For example, if the online attention score of a topic T_i is significantly higher in only one or a few articles, and the attention score of other articles is small or even 0, it will make the total score of $S(T_i)$ and $\frac{S(T_i)}{n_i}$ average scores relatively high. The consistency index $C(T_i)$ is introduced, and only when T_i gets higher and more uniform attention in distributed papers can higher Z_t value be obtained. In this way, the z_t index describes T_i's degree of attention from the three levels of overall scale-average level-distribution consistency. It can comprehensively and effectively identify the hot frontier topics of the disciplines.

First, according to the high attention topic word set, the number of distributed articles ni, the total AAS score $S(T_i)$, the average article score $\frac{S(T_i)}{n_i}$ and the distribution consistency index $C(T_i)$ corresponding to each topic word T_i, the average number of each index value is calculated as the reference value.

Secondly, according to the rules shown in Figure 4, the research topics of AAS high attention are divided into potential class, emergent class, core class, and edge class.

Take information science as an example to demonstrate. After data processing and calculation, the research results are shown in Table 2.

Comparing the research results in Table 1 and Table 2, we find that there are many same results in the same category, such as Altmetrics, social media, collaboration, user behavior, citation analysis, impact factors, etc. Due to the difference in the scientific research environ-

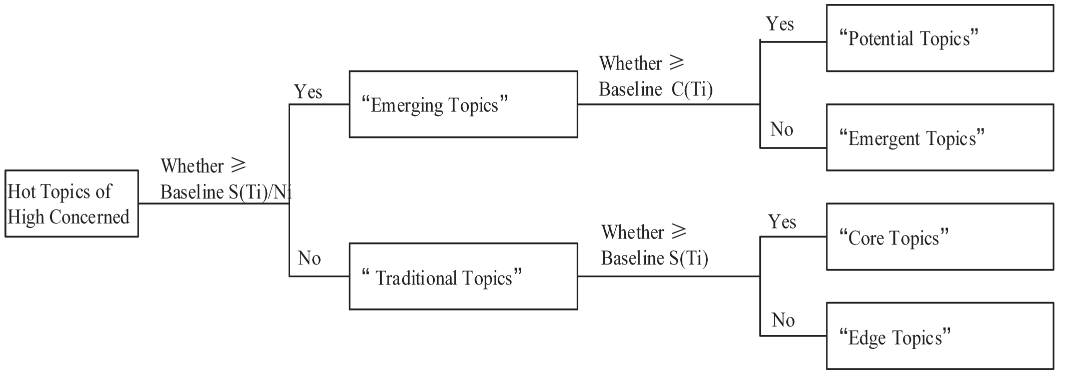

Figure 4 Classification model of the disciplinary research topics of AAS high concerned

Table 2 Classification table of international information science of the research topics of AAS high attention in 2018

Topics Type	Topic Subdivision	Research Topics
"Emerging Category"	"Potential Class"	Altmetrics; social media; collaboration;user–behavior; open access
	"Emergent Class"	Bibliometrics; citation analysis; impact factors
"Traditional Category"	"Core Class"	Bibliometrics; citation analysis; impact factors
	"Edge Class"	Machine learning; research evaluation; research productivity; Google Scholar

ment, the recognition results also have some differences. For example, international information science pays more attention to peer review and gender research such as Peer Review and Gender. In contrast, domestic information science studies on "WeChat" and "Weibo" are more popular.

Therefore, the z-index method identifies highly cited and high network attention research topics in information science at home and abroad. The identification results have similarities and differences. On the one hand, it shows that domestic and foreign information science research hotspots have similarities and differences; on the other hand, it shows that the zy (tf) index to identify highly cited disciplinary hotspots. Furthermore, the z- index to identify the research topics of AAS high attention are both feasible and effective. The two mutually verify and complement each other and comprehensively and synthetically identify domestic and foreign information science research hotspots from the highly cited scientific literature and network media's high attention.

We have published detailed conclusions of the AAS improvement z-index research (Pai et al., 2019).

Identification of Key Knowledge Nodes in Interdisciplinary Based on Overlapping Community

CFinder is a software developed based on the CPM algorithm. It can not only locate and visualize large-scale sparse network communities very effectively but can also be used to quantitatively describe social networks' evolution (Adamcsek et al., 2006; Palla et al., 2005).

It is found that the clique filtering CPM (clique percolation method) algorithm of an overlapping community can process and present complex network data quickly and efficiently. A clique is a collection of vertices connected to any two points, that is, a complete subgraph. The community nodes are closely connected and have high edge density, making it easy to form cliques. Therefore, the edges within the communities are more likely to form a large complete subgraph. In contrast, the edges between the communities are almost impossible to form a large complete subgraph, so that the communities can be discovered by finding the cliques in the network. Simultaneously, a node's cliques may belong to different communities so that the clique filtering method can find overlapping communities. The idea of the CPM algorithm is to find a complete subgraph of a pre-specified scale and then merge them by expand or adjacent cliques to obtain overlapping community structures.

Taking information science and computer science as an example, CFinder is used to intuitively study the cross-research topics and potential research directions of information science and computer science. The visualization results are shown in Figure 5.

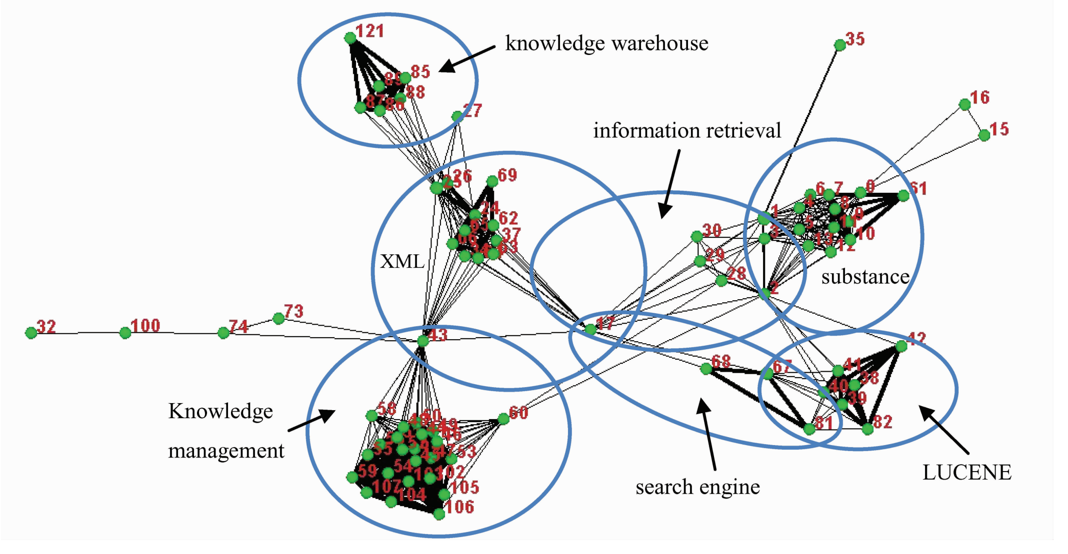

Figure 5 The largest overlapping community network diagram of information science and computer science 5-clique

We have published detailed visualization results of the overlapping communities research (Li et al., 2013).

Interdisciplinary related knowledge discovery based on the weak relationship of co-word network

The disjoint literature-based knowledge discovery method is also called Swanson theory, which was proposed by Professor Don R. Swanson of the University of Chicago in 1986 (Swanson, 1986). The method includes closed and open:

(1)Closed-Type (2)OPen-Type

Figure 6 Disjoint literature-based knowledge discovery method

Use A, B, and C to represent three different concepts, and assume a co-occurrence relationship between A and B, and B and C in the published literature. However, there is no co-occurrence relationship between A and C. Then, the closed-type method can be expressed as Figure 6(1). If B can be retrieved by A, and B can be retrieved by C, A and C are connected. The open-type method can be expressed as Figure 1(2). If B can be retrieved by A and C can be retrieved by B, A and C are connected.

We selected the top 3 journals separately with the highest composite impact factors in information science and computer science as data sources and extracted keywords from six journals in 2015 as sample data for empirical research. After data preprocessing, we obtained 4605 unique keywords in two disciplines and then designed a VBA program to add a suffix to each keyword. Keywords that only appear in information science were suffixed with (Q) and only appear in computer science were suffixed (J), and keywords that appear in both disciplines were not suffixed.

(1) Identifying potential interdisciplinary collaborative research topics based on closed–type disjoint literature –based knowledge discovery. Used closed-type disjoint literature-based knowledge discovery method and combined keyword co-occurrence theory and social network analysis methods to establish a two-disciplinary overall keyword link network. For a certain discipline heading (a_i) of a certain discipline, found the indirect related discipline term (c_i) with a distance of 2 in another discipline through the distance matrix. Then looked for the middle word (b_i) to verify the topic discovery process of $a_i \rightarrow c_i$.

Assuming that the keywords ai and ci have n intermediate keywords, there are n connection paths between them, as shown in Figure 7.

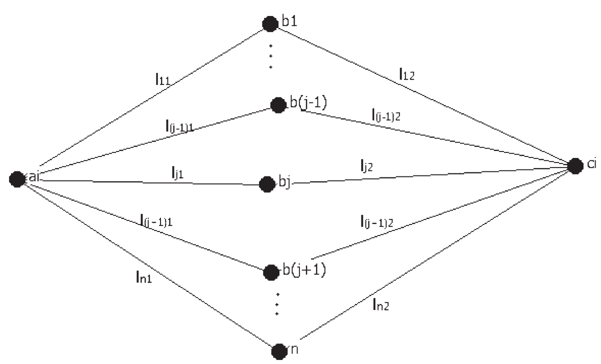

Figure 7 Diagram of each scale between target keywords

Then, we define the topic of interdisciplinary cooperation potential index (TICPI) of ai and c_i as:

$$\mathrm{TICPI} = 0.5 \times n + 0.5 \times \sum_{j=1}^{n} \frac{(l_{j1} \times l_{j2})^2}{|l_{j1} - l_{j2}| + \beta} \qquad (6)$$

We calculated the keywords' TICPI value in the sample data and ranked them from high to low. The relevant data of the top ten pairs of keywords are shown in Table 3.

Table 3 Topic interdisciplinary cooperation potential index between information science and computer science

Initial words (a_i)	Target words(c_i)	TICPI index
think tank (Q)	CMB algorithm (J)	10.53
knowledge service (Q)	mouse behavior (J)	5.25
internet public opinion (Q)	sub–topic (J)	5.17
knowledge graph (Q)	human–computer interaction (J)	4.75
rumor propagation (Q)	social computing (J)	4.5
authorship attribution (Q)	modularity (J)	3.75
library cloud (Q)	incremental learning (J)	3.5
multi–source information fusion (Q)	Map Reduce cluster (J)	3.5
text classification (Q)	spatial filtering (J)	3.5
zero cited (Q)	popularity (J)	3.5
…	…	…

* The data in the table are rounded estimates.

(2) Identifying potential interdisciplinary collaborative research topics based on open–type disjoint literature–based knowledge discovery. The research was carried out on the basis of data and the interdisciplinary paper keyword co-occurrence network. Firstly, a two-dimensional vector space model of nodes in keyword co-occurrence network was constructed, and each keyword (a_i) was expressed in the form of a vector. Secondly, we filtered each keyword's middle word set according to certain criteria to build the core vector model. Then, we defined the substitution method of the keyword core vector model and found the target word (c_i) by substitution method after determining the middle word (b_i). Finally, according to the number of paths and the frequency of mutual co-occurrence between a_i and c_i to calculate the interdisciplinary cooperation potential of each pair of keywords for a_i and c_i.

Substitute sample data for calculation, such as the knowledge graph.

$$\text{knowledge graph(Q)}^{*}=\begin{bmatrix} 3 & \text{information visualization} & 3 & \text{knowledge graph(Q)} \\ 3 & \text{information visualization} & 2 & \text{data mining} \\ 3 & \text{information visualization} & 2 & \text{big data} \\ 2 & \text{data mining} & 2 & \text{big data} \\ 2 & \text{data mining} & 2 & \text{road network model(J)} \\ 2 & \text{data mining} & 2 & \text{competitive intelligence(Q)} \\ 2 & \text{research hotspot(Q)} & 3 & \text{information visualization} \\ 2 & \text{research hotspot(Q)} & 2 & \text{knowledge graph(Q)} \\ 2 & \text{research hotspot(Q)} & 2 & \text{co-citation(Q)} \end{bmatrix}$$

The results are shown in Table 4.

In Table 4, internet public opinion (Q), knowledge service (Q), think tank (Q), rumor propagation (Q), and text categorization (Q) are topics with great potential for interdisciplinary cooperation. Moreover, those recognition results, identified by the open-type disjoint literature-based knowledge discovery method, are the same as the closed-type disjoint literature-based knowledge discovery method. Analyzing the results of the two methods: ①Their results have common parts, which shows that although the research methods of open-type

Table 4 Interdisciplinary cooperation potential index of information science and computer science

Initial words (a_i)	Middle words (b_{ij})	Target words (c_i)	Interdisciplinary cooperation potential
knowledge graph (Q)	data mining	road network model (J)	5.67
internet public opinion (Q)	semantics	sub–topic (J)	5.33
knowledge service (Q)	information retrieval	mouse behavior (J)	4.8
information recommendation (Q)	collaborative filtering	project popularity (J)	4.67
think tank (Q)	big data	CMB algorithm (J)	4.33
rumor propagation (Q)	social network	social computing (J)	3.5
text categorization (Q)	feature extraction	spatial filtering (J)	2.7
information recommendation (Q)	collaborative filtering	shilling attacks (J)	2.67
intelligence 5.0 (Q)	parallel system	traffic waves (J)	2
user–behavior (Q)	mobile Internet	Pervasive computing (J)	2

and closed-type disjoint literature-based knowledge discovery methods are completely different, they are fundamentally similar; ②the recognition results of the two methods are different, which shows that the two methods have their own focus in the research process and there is room for mutual reference.

We have published the detail of the open-type and closed-type disjoint researches (Liu et al., 2017; Li & Liu et al., 2018).

Interdisciplinary related knowledge discovery based on the weak citation relationship

Based on the weak citation relationship of knowledge nodes indirectly connected between interdisciplinary subnets, identify interdisciplinary related knowledge combinations. The steps are divided into the following three stages: ①Construct a weak relationship network of interdisciplinary knowledge citation/cited; ②identify the weak relationship structure of the target discipline knowledge node (a)-knowledge medium (b)-interdisciplinary knowledge (c) in the interdisciplinary citation network; ③construct an interdisciplinary related knowledge combination (a-c) evaluation and identification model.

Judge whether the target discipline knowledge (a) can form a knowledge associated with a certain interdisciplinary related knowledge (c) through a weak connection. First, look for the

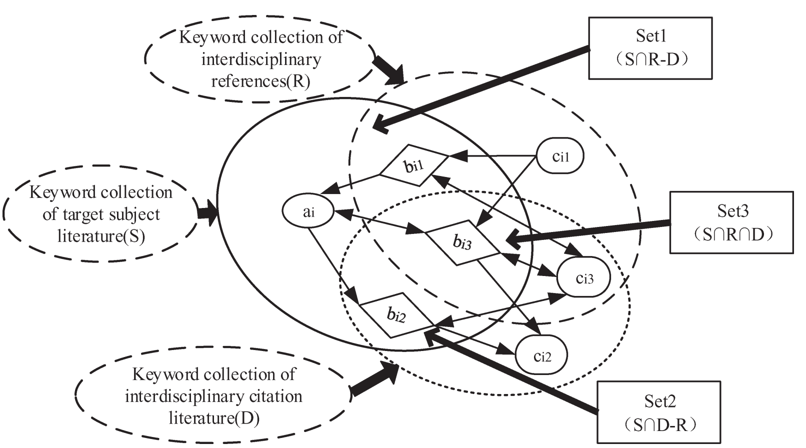

Figure 8 Knowledge medium type b and the a-b-c weak relationship connection path

knowledge medium (b) between them as a bridge. The citing and being cited of interdisciplinary literature is essentially the free combination of knowledge genes from different disciplines to form cross knowledge into the scientific knowledge exchange system, and then generate knowledge chains and knowledge networks linking different disciplines (Wang & Li, 2018). Therefore, interdisciplinary keywords are the knowledge medium that establishes a weak connection between target discipline knowledge and interdisciplinary related knowledge. This is shown in Figure 8.

We set the target discipline literature keyword set S, its interdisciplinary reference literature keyword set R, and the cited literature keyword set D, then the interdisciplinary knowledge weak citation relationship network G is the keywords in the set of S, R, and D, based on the knowledge association established by citation and being cited. Then, the discipline knowledge node a is represented by the keyword a_i, and $a_i \in S$; interdisciplinary related knowledge c is expressed by the keyword c_i, $c_i \in (R \cup D\text{-}S)$; knowledge medium b is represented by the keyword b_i, $b_i \in$ Set 1 $\cup$ Set 2 $\cup$ Set 3, among them, set 1 ($S \cap R$-D), set 2 ($S \cap D$-R), and set 3 ($S \cap R \cap D$) represent keywords of interdisciplinary research in network G.

According to the flow of knowledge, the knowledge medium bi is subdivided into three types: ①Inflow-type knowledge medium $b_{i1} \in$Set1, interdisciplinary knowledge flows into the target discipline through references.②Outflow-type knowledge medium $b_{i2} \in$Set2, the target discipline outputs knowledge through interdisciplinary citing documents. ③Flow-type knowledge medium $b_{i3} \in$Set3, transfer knowledge between disciplines by citing and being cited.

The trend analysis method is used to define the influence index AI of the target discipline knowledge node. The degree of activity in the discipline research is described by judging the trend of the popularity of the discipline knowledge node. The formula is

$$A_I = \frac{Y \times \sum_{y=1}^{Y}(y \times F_y) - \sum_{y=1}^{Y} y \times \sum_{y=1}^{Y} F_y}{Y \times \sum_{y=1}^{Y} y^2 - (\sum_{y=1}^{Y} y)^2} \qquad (7)$$

In the formula, the Y value represents the time span (number of years) of the data sample, and F_y is the research frequency of the keyword a_i in year y in the target discipline. A_I is the slope (rate of change) of the fitted line. If $A_I>0$, it means that this keyword's research interest is on the rise, and the greater the AI value, the greater the rate of change and the higher the activity.

Drawing on Porter's diversity of discipline distribution measurement index (Porter et al., 2007), define the knowledge media influence index B_I:

$$\mathrm{B}_I = \frac{(F_1+F_2+\cdots+F_N)^2}{{F_1}^2+{F_2}^2+\cdots+{F_N}^2} = \frac{\left(\sum_{n=1}^{N} F_n\right)^2}{\sum_{n=1}^{N} {F_n}^2} \qquad (8)$$

In the formula, if the knowledge medium bi has appeared in N disciplines, then Fn represents the number of academic papers in the nth discipline research bi, the sum of F_1, F_2...F_N is the total number of papers in N disciplines of b_i. Among them, $B_I \geq 1$, the larger the value, the higher of interdisciplinary diversity of bi, and the greater the influence. When the knowledge medium b_i only appears in the literature of one discipline, B_I's value is 1.

Define the interdisciplinary related knowledge relevance index C_I. The calculation formula is as follows.

$$C_I = \frac{(I_{ab} \times I_{bc})^2}{|I_{ab} - I_{bc}| + \beta} \qquad (9)$$

I_{ab} is the citation/cited frequency of a_i and b_i, and Ibc is the citation/cited frequency of b_i

and c_i, then the degree of correlation between node c_i and a_i is positively correlated with I_{ab} and I_{bc}. To make the above formula meaningful, introduce β ,when $I_{ab}=I_{bc}$, set β =1, when $I_{ab} \neq I_{bc}$, set β =0.

In a combination of interdisciplinary related knowledge , the degree of possibility of collaboration between discipline knowledge nodes a_i and interdisciplinary related knowledge c_i through knowledge media bi can be regarded as the size of the gravity of ai and c_i nodes in the entire citation association network, which is affected by the key knowledge node influence index A_I, knowledge media influence index B_I, and interdisciplinary knowledge relevance index C_I. Therefore, we define interdisciplinary related knowledge combination potential cooperation index, and its formula is:

$$P = A_I \times B_I \times C_I = \frac{Y \times \sum_{y=1}^{Y}(y \times F_y) - \sum_{y=1}^{Y} y \times \sum_{y=1}^{Y} F_y}{Y \times \sum_{y=1}^{Y} y^2 - (\sum_{y=1}^{Y} y)^2} \times \frac{(\sum_{n=1}^{N} F_n)^2}{\sum_{n=1}^{N} F_n{}^2} \times \frac{(I_{ab} \times I_{bc})^2}{|I_{ab} - I_{bc}| + \beta} \tag{10}$$

We selected information science journals as empirical samples for research. Through the processing and analysis of sample data, the following results are obtained:

Table 5 Relevant data on the recognition result of the interdisciplinary knowledge combination of the information discipline (part)

ranking	Knowledge node (a)	Knowledge media (b)	Interdisciplinary related knowledge (c)	Flow path	Influence index of a (A_I)	Influence index of b (B_I)	Relevance index of c (C_I)	Potential cooperation index (P)
1	research cooperation	knowledge flow	population dynamics model	a⟷b⟵c	2.60	2.73	5.33	37.35
2	citation network	multi–data fusion	finite element analysis	a⟵b⟵c	2.50	2.34	6.25	35.85
3	interdisciplinary	link prediction	random forest algorithm	a⟵b⟷c	2.60	1.99	6.25	32.48
4	highly cited literature	complex network	herding effect	a⟵b⟵c	2.70	2.83	4.00	30.87
5	think tank	big data	MongoDB database	a⟷b⟶c	2.30	3.24	4.00	29.82
6	Altmetrics	open access	information network communication right	a⟷b⟶c	3.10	2.08	4.50	29.25
7	knowledge graph	association rules	multi–label learning	a⟷b⟵c	2.00	2.17	6.25	26.77
8	internet public opinion	entity similarity	information cocoon room	a⟷b⟶c	2.50	2.09	4.00	20.60
9	sleeping beauty literature	superstring theory	superstring gravity model	a⟵b⟶c	2.40	1.68	4.50	18.37
10	social media	interpersonal intelligence network	concept lattice	a⟶b⟵c	2.30	1.73	4.00	16.02
…	…	…	…	…	…	…	…	…

The weak citation relationship research has been published online on CNKI (Pai et al., 2020).

Identifying interdisciplinary knowledge based on references

As is shown in Figure 9, related knowledge's citation relationship between disciplines can be presented by reference paths of node document keywords-interdisciplinary reference keywords.

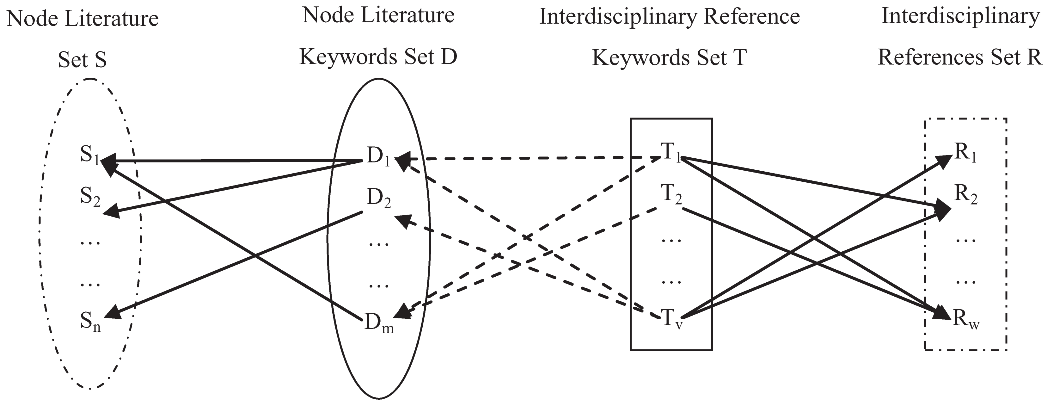

Figure 9 Paths of node literature keywords-interdisciplinary reference keywords

C_i represents the number of times the interdisciplinary citation keyword T_i be cited in the node literature set S, N_i represents the number of target discipline documents in the node literature collection S that cite the interdisciplinary citation keyword T_i. The strength of the correlation between T_i and the target discipline is positively correlated with Ci and N_i. Therefore, we can multiply the T_i and C_i as a numerator. K_i represents the number of documents published in the target discipline of the node literature collection S with T_i as a keyword (published before the publication year of the node literature collection), the novelty of T_i in the target discipline is inversely proportional to K_i, and K_i can be used as the denominator. Therefore, the Index of Discipline-related Novelty (IDN) calculation formula is as follows:

$$\mathrm{IDN}(T_i) = \frac{C_i \times N_i}{K_i + 1} \qquad (11)$$

Add 1 to Ki in the denominator to make the above formula meaningful.

We choose information science as the object of empirical research. Through the processing and calculation of sample data, the results are as follows:

Table 6 Part of interdisciplinary related knowledge identification results

No.	Interdisciplinary Reference Keywords	C_i	N_i	K_i	$C_i \times N_i$	K_i+1	IDN
1	Spatiotemporal data model	9	2	0	18	1	18.000
2	Theory of policy process	5	3	0	15	1	15.000
3	Random Forest	7	3	3	21	4	5.250
4	Conditional random field model	16	7	24	112	25	4.480
5	OPTICS algorithm	2	2	0	4	1	4.000
6	Social learning theory	2	2	0	4	1	4.000
7	Neural network model	24	15	133	360	134	2.687
8	Maximum entropy model	6	4	9	24	10	2.400
9	Self-efficacy theory	9	8	30	72	31	2.323

No.	Interdisciplinary Reference Keywords	C_i	N_i	K_i	$C_i \times N_i$	K_i+1	IDN
10	S–CAD method	2	1	0	2	1	2.000
11	Exponential random graph model	2	1	0	2	1	2.000
12	SOVA algorithm	2	1	0	2	1	2.000
	ACP method	2	2	1	4	2	2.000
……	……	……	……	……	……	……	……

We have published the detail of identifying interdisciplinary knowledge based on reference research (Du et al., 2020).

Discussion

On the one hand, through scientific literature and online media data, choosing appropriate research methods, we have identified key knowledge nodes of disciplines such as hotspots, frontiers, and interdisciplinary keywords. On the other hand, through the weak relationship of the co-word network and citation network, we found these key knowledge nodes' interdisciplinary knowledge. So, which of these knowledge pairs can become a potential interdisciplinary knowledge growth point? This requires a combination of quantitative and qualitative research methods.

(1) Building models to identify potential interdisciplinary knowledge growth points. Research on interdisciplinary knowledge fusion, or forming a new discipline: Discipline growth point, such as physical chemistry, biophysics, etc. Generating new knowledge: interdisciplinary knowledge growth point, such as cooperation network, co-word network, etc. What are the background, causes, and influencing factors of these existing growth points, especially interdisciplinary knowledge growth points? The answer to this question can be used for reference in identifying potential interdisciplinary knowledge growth points. Therefore, the focus of follow-up research is: Based on scientific literature data, using time series analysis, system dynamics, link prediction, information entropy, and other methods to analyze the key knowledge of the existing interdisciplinary knowledge and its interdisciplinary related knowledge. Furthermore, build potential interdisciplinary knowledge growth points model based on their connection strength over time, the characteristics of different evolution stages, the key factors of the formation process and their relevance, etc. Apply the model to analyze the matching results of key knowledge nodes of the subject and interdisciplinary related knowledge, and identify potential interdisciplinary knowledge growth points with greater cooperation and innovation potential.

(2) Using the Delphi method and other methods to evaluate and predict potential interdisciplinary knowledge growth points. The qualitative analysis method based on experts' wisdom refers to a method that researchers use experts' knowledge and experience to evaluate the research results to verify the research results' reliability (Bengisu & Nekhili, 2006). The interview method was used to verify the findings of Upham, which shows that three-quarters of their findings are meaningful and important research directions (Upham & Small, 2010). Therefore, qualitative methods such as interdisciplinary expert consultation and questionnaire surveys can evaluate and verify the matching results of key discipline knowledge nodes and interdisciplinary related knowledge, or the potential interdisciplinary knowledge growth point identified by the predictive model, to improve the identification results reliability and validity.

Whether it is quantitative identification or qualitative evaluation, it is not a direct verification of the findings. Whether a potential interdisciplinary knowledge growth point can finally become an interdisciplinary knowledge growth point depends on the researcher's cognition and understanding of the research target field's characteristics.

Concluding Remarks

The scientific system's complexity makes it impossible for many social issues and scientific research to be completed independently in a single discipline, and interdisciplinary research has become an indispensable research mode in modern science (Gates et al., 2019). However, the knowledge of a certain discipline can be co-innovated with what knowledge of which disciplines, which is an obstacle to researchers' interdisciplinary research. First, based on the scientific growth point, this article defines the concepts of "growth point of knowledge" and "interdisciplinary knowledge growth point", and analyzes its generation principle and related research. Finally, this paper discusses the methods and research ideas of identifying key knowledge nodes and interdisciplinary related knowledge and generating potential interdisciplinary knowledge growth point by combining the two. It makes scientific research attempts to draw relevant conclusions. In future research: On the one hand, the characteristic attributes and evaluation indicators of growth points of knowledge are continuously improved and perfected while applying; on the other hand, put the research ideas into scientific research activities, establish a scientific and effective potential interdisciplinary knowledge growth point identification system to help researchers to identify interdisciplinary related knowledge, find the research gaps of their disciplines, and promote interdisciplinary knowledge integration and innovation.

Acknowledgments

The researches described in this paper is supported by the National Social Science Foundation of China, Research on Identification of Interdisciplinary Potential Knowledge Growth Point and Innovation Trend Forecast (No. 19ATQ006). The authors are grateful to Prof. Qiu Junping from Hangzhou Dianzi University. Thanks are also due to the editors and reviewers for their constructive suggestions.

References:

Abbasi, A., Altmann, J. R., & Hossain, L. (2011). Identifying the effects of co–authorship networks on the per–formance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics, 5* (4), 594–607.

Adamcsek, B., Palla, G., Farkas, I. J., Derényi, I., & Vicsek, A. T. (2006). CFinder: Locating cliques and over–lapping modules in biological networks. *BIOINFORMATICS*

Altmetric.com. (2020). *How is the Altmetric Attention Score calculated?*. Retreved 2020 from https://help.altmet–ric.com/support/solutions/articles/6000060969–how–is–the–altmetric–score–calculated–/

Baidu Encyclopedia. (2020). *Baidu Encyclopedia – scientific growth point*. Retreved 2020–12–24 from https://baike.baidu.com/item/%E7%A7%91%E5%AD%A6%E7%94%9F%E9%95%BF%E7%82%B9/4556600?fr=al–addin

Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). *The Role of Social Networks in Information Diffusion.* Paper presented at the Annual conference on world wide web.

Bengisu, M., & Nekhili, R. (2006). Forecasting emerging technologies with the aid of science and technology databases. *Technological Forecasting & Social Change, 73* (7), 835–844.

Bettoni, M., Schiller, G., & Bernhard, W.(2008). Weak ties cooperation in the CoRe knowledge network

Chen, C.(2005). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology, 3* (57), 359–377.

David, E., & Jon, K.(2012). *Networks, crowds, and markets: Reasoning about a highly connected world.* Cambridge University Press.

Deng, Q., & Wang, X.(2018). A method of identifying key research directions. *Library and Information Service, 17* (62), 60–67.

Du, D. H., Li, C. L., Xiang, F. Z., & Pai, Y. X.(2020). Discussion on the method of interdisciplinary related knowledge discovery based on citation keywords. *Journal of Intelligence, 9* (39), 189–194.

Duan, Q. F., & Pan, X. H.(2017). Identification of emerging topics in Science Using Social Media. *Journal of the China Society for Scientific and Technical Information, 12* (36), 1216–1223.

Elkin, N., & Anke, P.(2015). Topics in the literature of transition economies and emerging markets. *Scientometrics, 3* (102), 2107–2130.

Fajardo–Ortiz, D., Lopez–Cervantes, M., Duran, L., Dumontier, M., & Castano, V. M.(2017). The emergence and evolution of the research fronts in HIV/AIDS research. *PLoS One, 12* (5), e178293.

Gates, A. J., Ke, Q., Varol, O., & Barabάsi, A.(2019). Nature´s reach: narrow work has broad impact. *Nature, 575* (7781), 32–34.

Genuis, S. K.(2005). Published literature and diffusion of medical innovation: Exploring innovation generation. *Canadian Journal of Information & Library Science, 29* (1), 27–54.

Granovetter, M.(1973). The Strength of Weak Ties. *The American Journal of Sociology, 6* (78), 1360–1380.

Huang, S. Q., & Ma, J. L.(2009). Empirical research on noninteractive literature–based discovery of Chinese social sciences literature: With example of agricultural economics literature. *Journal of Library Science in China, 4* (35), 31–38.

Kuang, D. H., & Zhang, L. B.(2019). Altmetrics characteristics analysis of top Academic Papers: An Empirical Study Based on "Nature" Weekly. *Information Studies: Theory & Application, 2* (42), 80–86.

Lai, J. Y., Wei, S. Y., & Qin, Y.(2017). Subject cognition mining based on Zhihu hotspots: Taking informatics as an example. *Library and Information Service, 24* (61), 96–104.

Li, C. L., Feng, Z. G., Liu, Y. M., & Liu, X. H.(2018). Identification of potential interdisciplinary collaborators based on citation network: Taking library and information science as an example. *Information and Documentation Services, 3*, 93–98.

Li, C. L., Gao, F., & Pai, Y. X.(2020). Research on potential interdisciplinary knowledge growth points and its identification methods. *Studies in Science of Science*, 1–14.

Li, C. L., Liu, F. F., & Guo, F. J.(2013). Analysis on interdisciplinary research topics with cinder of overlapping communities visualization software—Taking the information science and computer science for example. *Library and Information Service, 7* (57)

Li, C. L., Liu, X. H., Liu, Y. M., & Feng, Z. G.(2018). Identifying potential disciplinary collaboration research topics by open literature–based discovery: Taking information science and computer science as examples. *Information Studies: Theory & Application, 2* (41), 100–104, 137.

Li, C. L., Pai, Y. X., Xiang, F. Z., & Du, D. H.(2019). Discussion on the method of identifying subjects research hot spots of highly cited using an improved Z–Index. *Information Studies: Theory & Application, 6* (43), 69–75, 69.

Li, H., Wang, H. N., & You, Q. F.(2018). The status and challenges of empirical research on altmetrics. *Information Studies: Theory & Application, 12*(41), 76–81.

Liu, X. H., Li, C. L., Cui, B., & Liu, T.(2017). Research topics identification of potential interdisciplinary collaboration based on closed and irrelevant knowledge discovery. *Information Studies: Theory & Application, 9* (40), 71–76.

Ma, F., & Liu, M.(2014). Structure, evolution and hot spot of Chinese economics´ knowledge network—Statistics Analysis of Data from CSSCI(1998–2011). *Information Science, 7* (32), 3–8.

Meyer, R. M.(2010). Diversity and network coherence as indicators of interdisciplinarity: Case studies in bio–

nanoscience. *Scientometrics.*

Onnela, J. P., Saramki, J., Hyvnen, J., Szabó, G., & Barabási, A. L.(2007). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences, 104* (18), 7332–7336.

Pai, Y. X., Li, C. L., & Liu, Y. M.(2019). AAS high concern subject research topic identification based on Z–index. *Information and Documentation Services, 6* (40), 30–37.

Pai, Y. X., Li, C. L., & Xu, L.(2020). Discussion on the method of interdisciplinary related knowledge combination identification on the perspective of weak citation relationship—Taking information science for example. *Library and Information Service*

Palla, G., Derényi, I., Farkas, I., & Vicsek, T.(2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature, 435* (7043), 814–818.

Porter, A. L., Cohen, A. S., Roessner, J. D., & Perreault, M.(2007). Measuring researcher interdisciplinarity. *Scientometrics, 72* (1), 117–147.

Prathap, & Gangan.(2017). A three–dimensional bibliometric evaluation of recent research in India. *Scientometrics, 110* (3), 1–13.

Prathap, G.(2014a). A three–class, three–dimensional bibliometric performance indicator. *Journal of the American Society for Information Science & Technology, 65* (7), 1506–1508.

Prathap, G.(2014b). A three–dimensional bibliometric evaluation of research in polymer solar cells. *Scientometrics.*

Prathap, G.(2014c). The zynergy–index and the formula for the h–index. *Journal of the Association for Information Science & Technology, 65* (2), 426–427.

Shang, X. L.(2018). The pattern of knowledge combination in cross–disciplines based on multi–mode topic network––An example of digital library. *Information Science, 3* (36), 130–137, 150.

Sheng, Y.(2012). Subject hotspots discovery, tracking and analysis based on Microblog—Take the field of data mining as an example. *Library and Information Service, 8* (56), 32–37.

Small H, & Sweeney E.(1985). Clustering the science citation index using co–citations. *II.Mapping science, 8*, 321–340.

Song, M., Kang, K., & An, J. Y.(2018). Investigating drug–disease interactions in drug–symptom–disease triples via citation relations. *Journal of the Association for Information Science and Technology, 69* (11)

Swanson, D. R.(1986). Fish oil, Raynaud´s syndrome, and undiscovered public knowledge. *Perspectives in Biology & Medicine, 30* (1), 7–18.

Takahashi, T., Tomioka, R., & Yamanishi, K.(2014). Discovering Emerging Topics in Social Streams via Link–Anomaly Detection. *IEEE Transactions on knowledge and data engineering, 1* (26), 120–130.

The Biology Editorial Committee of EOC.(1991). *Encyclopedia of China: Biology*. China Encyclopedia press.

Upham, S. P., & Small, H.(2010). Emerging research fronts in science and technology: patterns of new knowledge development. *SCIENTOMETRICS, 83* (1), 15–38.

Vlachý J.(1984a). Research fronts in physics 1983. *Czechoslovak Journal of Physics, 2* (34), 171–174.

Vlachý J.(1984b). Priority choice and research front specialties in physics. *Czechoslovak Journal of Physics, 1* (34), 95–98.

Wang, X. W., Mao, W. L., & Wang, Z.(2014). Detecting and tracing emerging research trends real–timely using web data. *Science of Science and Management of S.&.T.(Monthly), 4* (35), 3–9.

Wang, X., & Li, R.(2018). Interpretation of citation phenomena from the perspective of knowledge ecology. *Journal of Intelligence, 9* (37), 179–184.

Wei, L., Xu, H. Y., Guo, T., & Fang, S.(2015). Study on the interdisciplinary topics of information science based on weak co–occurrence and burst detecting. *Library and Information Service, 21* (59), 105–114.

Wu, L., Tian, R. Y., & Zhang, X. F.(2017). Interdisciplinary topic detection method and empirical research based on topic correlation analysis: A case sudy of animal resource and breeding. *Library and Information Service, 1*, 72–79.

Xu, H. Y., Guo, T., Yue, Z. H., & Fang, S.(2015). Study on the interdisciplinary topics of information science based on TI index series. *Journal of the China Society for Scientific and Technical Information, 10* (34), 1067–1078.

Xu, K., & Bi, Q.(2019). Keywords classification and hotspots feature analysis. *Information Studies:Theory & Ap–plication, 4* (42), 96–100.

Yang, L., Morris, S. A., & Barden, E. M.(2009). Mapping institutions and their weak ties in a specialty: A case study of cystic fibrosis body composition research. *Scientometrics, 79* (2), 421–434.

Ye, X.(2007). Hotspots analysis on highly cited papers of library and information science in China. *Library and Information Service, 12* (51), 138–141.

Yu, L. P., & Guo, Q. H.(2019). Modified study of cited peak paradox and impact factor: Time impact factor. In–formation Studies: *Theory & Application, 7* (42), 54–58.

Zhang, D. M.(1986). On the scientific growth point. *Studies in Philosophy of Science and Technology* (01), 1–7.

Zhao, J., Wu, J., & Xu, K. (2010). Weak ties: Subtle role of information diffusion in online social networks. *Physical Review E, 82* (1 Pt 2), 16105.

Zhao, Y. X., & Yang, Z. P.(2016). Detecting research hot–spots using altmetrics: A case study on information & computing science. *Journal of Intelligence, 11* (35), 39–44.

Zhu, Q., & Leng, F.(2014). Topic identification of highly cited papers based on citation content analysis. *Journal of Library Science in China, 01* (40), 39–49.