Identifying Interdisciplinary Relevant Knowledge Combinations via Weak Citation Relationships: A Case Study in Information Science (Postprint)

Authors: Pai Yanxin, Li Changling, Xu Lu

Date: 2023-04-01T16:16:03+00:00

Abstract

[Purpose/Significance] The increasing complexity of scientific systems has made interdisciplinary collaboration a crucial paradigm and inevitable trend in modern scientific innovation research. Identifying interdisciplinary knowledge combinations with high collaborative potential has become essential for fostering interdisciplinary research cooperation and innovation. [Method/Process] First, source literature from the target discipline along with its interdisciplinary references and citing literature are selected to construct a keyword-based weak citation association network for interdisciplinary knowledge. Second, the types of knowledge media b are classified, and the weak relationship structure of target discipline knowledge node a–knowledge media b–interdisciplinary related knowledge c is identified. Finally, the target discipline knowledge node influence index AI, knowledge media influence index BI, interdisciplinary knowledge relevance index CI, and interdisciplinary knowledge combination a–c potential collaboration index P are defined to identify interdisciplinary related knowledge combinations with high collaboration potential values. [Results/Conclusion] An empirical study was conducted using articles from 9 CSSCI journals in the field of information science from 2015–2019 and their interdisciplinary reference and citing literature as samples, thereby verifying the effectiveness and feasibility of the interdisciplinary related knowledge combination discovery method based on weak citation relationships, and identifying high collaboration potential interdisciplinary related knowledge combinations for the information science discipline, such as "scientific collaboration"–"knowledge flow"–"population dynamics model".

Full Text

Abstract

[**Purpose/Significance**] As scientific systems become increasingly complex, interdisciplinary collaboration has emerged as a crucial paradigm and inevitable trend in modern scientific innovation research. Identifying interdisciplinary knowledge combinations with high collaborative potential represents a key factor in promoting interdisciplinary scientific cooperation and innovation. [**Method/Process**] This study first selects source literature from a target discipline along with its interdisciplinary references and citing literature to construct a keyword-based interdisciplinary knowledge weak citation association network. Next, it classifies the types of knowledge medium *b* and identifies the weak relational structure of target discipline knowledge node *a*—knowledge medium *b*—interdisciplinary knowledge *c*. Finally, it defines the Target Discipline Knowledge Node Influence Index *AI*, Knowledge Medium Influence Index *BI*, Interdisciplinary Knowledge Correlation Index *CI*, and Interdisciplinary Knowledge Combination *a-c* Potential Cooperation Index *P* to identify interdisciplinary knowledge combinations with high collaborative potential. [**Result/Conclusion**] An empirical study was conducted using papers published in nine CSSCI journals in the field of information science from 2015–2019, along with their interdisciplinary references and citing literature, to verify the effectiveness and feasibility of the proposed method. The analysis identified high-potential interdisciplinary knowledge combinations such as "scientific research cooperation"—"knowledge flow"—"population dynamics model."

Keywords: strength of weak ties; citation analysis; *a-b-c* weak connection; interdisciplinary related knowledge combination; potential cooperation index *P*

Classification Number: G250
DOI: 10.13266/j.issn.0252-3116.2020.21.014

---

1 Introduction

Disciplines represent systematic classifications of human knowledge. The increasing complexity of scientific systems means that many social problems and research questions cannot be solved by single-discipline knowledge alone. Consequently, scientific knowledge exchange and collaboration across disciplinary boundaries have become more frequent, making interdisciplinary research an indispensable model for modern scientific innovation development [1]. In this context, to address disciplinary research challenges, break through scientific bottlenecks, or achieve research innovation, researchers need to continuously acquire and dynamically integrate relevant concepts, theories, methods, and technologies from other disciplines through interdisciplinary collaborative research. However, in practice, researchers generally have a relatively good understanding of their own discipline but lack a clear grasp of which interdisciplinary knowledge can be used for collaborative research. Therefore, the identification of interdisciplinary knowledge combinations has become critical for interdisciplinary collaboration.

The concept of "strength of weak ties" originates from sociology, defined as brief social contacts between two actors [2]. Its proponent, Professor M. Granovetter [2], noted that while strong ties create close and stable connections within organizations, weak ties provide important pathways for information exchange between different groups and organizations, enabling isolated subgroups to establish connections. As these weak ties strengthen, the scope of information exchange expands further, accelerating the dissemination, integration, development, and innovation of information. Subsequent scholars such as J.P. Onnela et al. [3], E. Baksy et al. [4], J. Zhao et al. [5], and E. David et al. [6] have similarly discussed this perspective. S.K. Genius [7] found that weak ties can transmit more potential, diverse, and non-redundant knowledge resources compared to strong connections. A. Abbasi et al. [8], M. Bettoni et al. [9], and L.Y. Yang et al. [10] discovered that weak connection relationships make it easier for individuals to establish broader associations with others in different subnetworks of the knowledge network, playing a positive role in promoting scientific research cooperation.

Currently, some scholars in scientometrics have used weak ties in co-word networks for related research. Wei Ling et al. [11] constructed a high-frequency term weak co-occurrence network based on weak tie theory to analyze interdisciplinary patterns at the micro-level of information science. Li Changling et al. [12] and Liu Xiaohui et al. [13] used open and closed non-related knowledge discovery methods to identify potential interdisciplinary research topics between information science and computer science. However, M. Song et al. [14] found that citation analysis can identify more relationship pairs than co-occurrence analysis, and these pairs exhibit greater uniqueness and diversity.

Citations most directly reveal knowledge diffusion, association, and evolutionary paths between different disciplines, effectively promoting the synergy, intersection, integration, development, and innovation of interdisciplinary knowledge [15]. Therefore, interdisciplinary citation analysis is an effective tool for identifying interdisciplinary knowledge flow [16]. Weak knowledge associations exist in citation networks. According to the definition of weak ties, they can be categorized into several types: (1) low-threshold relationship nodes—connections where relationship strength in the network is below a threshold, representing weak relationships opposed to strong ones [11], such as citation relationships with small connection strength between source literature and its references or citing literature; (2) inter-subnetwork relationship nodes—relatively sparse connections between nodes from different subnetworks, such as citation relationships between different disciplinary knowledge in interdisciplinary citation networks; and (3) indirect relationship nodes—non-related knowledge connected indirectly through other nodes, such as co-citation or bibliographic coupling relationships, or knowledge nodes with good citation relationships but no co-occurrence relationships. Currently, few scholars have used weak ties in citation networks for knowledge mining and discovery research, and even fewer have addressed interdisciplinary research. Du Dehui et al. [18] used information science source literature keywords and interdisciplinary reference keywords as association data to build an interdisciplinary knowledge citation network, removing keyword co-occurrence relationships to form a weak citation relationship network for identifying interdisciplinary knowledge. Among the three types of weak ties in citation networks—low-threshold, inter-subnetwork, and indirect relationships—indirect relationship nodes between subnetworks without direct connections hold more potential collaborative value and are more suitable for interdisciplinary knowledge discovery [14].

Therefore, this study extracts keywords from disciplinary source literature, its interdisciplinary references, and interdisciplinary citing literature to construct an interdisciplinary knowledge association network. It identifies indirectly connected knowledge nodes between subnetworks as weak relationship association data to explore methods for discovering interdisciplinary knowledge combinations. Using information science as an example, this research aims to provide important decision-making references for future interdisciplinary collaborative research and targeted scientific innovation in this discipline.

---

2 Research Steps and Methods

Based on weak citation relationships between indirectly connected knowledge nodes in interdisciplinary subnetworks, this study identifies interdisciplinary knowledge combinations. The process consists of three stages: (1) constructing an interdisciplinary knowledge citation/cited weak association network; (2) identifying the weak relational structure of target discipline knowledge node *a*—knowledge medium *b*—interdisciplinary knowledge *c* in the interdisciplinary citation network; and (3) constructing an evaluation and identification model for interdisciplinary knowledge combination *a*-*c*. The specific steps and methods are shown in Figure 1 [Figure 1: see original paper].

2.1 Constructing an Interdisciplinary Knowledge Citation/Cited Weak Association Network

The discipline under study is the target discipline. The references and citing literature of target discipline source literature come from two sources: the target discipline itself (intra-disciplinary) and non-target disciplines (interdisciplinary). Knowledge uses literature as a carrier, and through citation and cited relationships between documents, knowledge flows and integrates across different disciplines, forming a dynamic knowledge exchange and association network [19]. This network is an interactive network composed of several strong and weak connection relationships [20]. Among them, citations between target discipline documents mostly represent internal knowledge exchange with close knowledge association and strong sharing capabilities, enabling deep-level communication and constituting strong relationship connections. In contrast, interdisciplinary references and citing literature have fewer exchange frequencies and lower association degrees with target discipline knowledge, representing weak connection relationships. Using keywords from the target discipline and interdisciplinary literature, a citation knowledge association network G is constructed, as shown in Figure 2 [Figure 2: see original paper].

2.2 Identifying the *a*-*b*-*c* Weak Relational Structure in Interdisciplinary Citation Networks

Weak tie theory posits that in social networks, if a and c have a common friend b, the probability of a and c becoming friends increases, thereby establishing some connection [2]. This principle also applies to interdisciplinary knowledge citation networks, as shown in Figure 3 [Figure 3: see original paper].

The a-b-c structure is a weak relational structure where node b acts as a "bridge" [2] connecting a and c through weak ties—that is, the knowledge medium. In interdisciplinary citation networks, knowledge media are often interdisciplinary intersection nodes.

2.2.1 Classification of Knowledge Medium *b* To determine whether target discipline knowledge a can establish a weak tie connection with interdisciplinary knowledge c, we must first identify knowledge medium b as a "bridge-builder." The citation and cited relationships of interdisciplinary literature essentially involve the free combination of knowledge genes from different disciplines, forming cross-disciplinary knowledge that enters the scientific knowledge exchange system and produces interdisciplinary knowledge chains and networks [21]. Therefore, interdisciplinary research keywords serve as knowledge media connecting target discipline knowledge and interdisciplinary knowledge through weak ties, as illustrated in Figure 4 [Figure 4: see original paper].

Let S be the keyword set of target discipline literature, R be the keyword set of its interdisciplinary references, and D be the keyword set of its interdisciplinary citing literature. The interdisciplinary knowledge weak citation association network G consists of keywords from sets S, R, and D based on citation and cited relationships. Disciplinary knowledge node a is represented by keyword a_i S; interdisciplinary knowledge c is represented by keyword c_i (R D - S); and knowledge medium b is represented by keyword b_i (Set 1 Set 2 Set 3), where Set 1 = (S R - D), Set 2 = (S D - R), and Set 3 = (S R D) represent interdisciplinary research keywords in network G.

Based on knowledge flow direction, knowledge medium b_i is subdivided into three types: (1) Inflow-type knowledge medium $b\_\{i1\}$ Set 1, where interdisciplinary knowledge flows into the target discipline through references; (2) Outflow-type knowledge medium $b\_\{i2\}$ Set 2, where the target discipline exports knowledge through interdisciplinary citing literature; and (3) Flow-type knowledge medium b_{i3} Set 3, where knowledge is transmitted between disciplines through both citation and cited relationships.

2.2.2 Identifying *a-b-c* Weak Relational Flow Paths Based on Different Medium Types Taking disciplinary knowledge node a_i as an example, it can connect with interdisciplinary knowledge c_i through knowledge media $b1_i$, $b2_i$, and $b3_i$ to form *a-b-c* weak relational structures. In Figure 4, unidirectional arrows represent reference or citation behavior (knowledge inflow or outflow), while bidirectional arrows represent mutual citation behavior (both knowledge inflow and outflow). There are seven different weak connection path types: (1) two paths based on "inflow-type" knowledge medium b_{i1}: $a_i \leftarrow b_{i1} \leftarrow c_{i1}$ and $a_i \leftarrow b_{i1}$ c_{i3}; (2) two paths based on "outflow-type" knowledge medium b_{i2}: $a_i \rightarrow b_{i2}$ c_{i3} and $a_i \rightarrow b_{i2} \rightarrow c_{i2}$; and (3) three paths based on "flow-type" knowledge medium b_{i3}: a_i $b_{i3} \leftarrow c_{i1}$, a_i b_{i3} c_{i3}, and a_i $b_{i3} \rightarrow c_{i2}$. A program was written to traverse all relationships in network G and extract all *a-b-c* weak relational structures.

2.3 Constructing an Interdisciplinary Knowledge Combination Identification Model

The activity level of disciplinary knowledge node a_i, the mediating capability of knowledge medium b_i, and the connection strength between nodes are all influential factors in analyzing the potential cooperation possibility between a and c in the *a-b-c* weak relational structure. Therefore, this study defines the Target Discipline Knowledge Node Influence Index AI, Knowledge Medium Influence Index BI, and Interdisciplinary Knowledge Correlation Index CI to quantitatively describe the characteristics and correlations of nodes in the *a-b-c* weak relationship. Based on these, a Potential Cooperation Index P model for interdisciplinary knowledge combinations is constructed to scientifically and reasonably identify interdisciplinary knowledge combinations.

2.3.1 Target Discipline Knowledge Node Influence Index *AI* Influenced by the hot topic emergence effect, knowledge innovation is more likely to occur in research processes showing clear upward trends [22]. Therefore, disciplinary knowledge nodes with clearly rising research trends exhibit high activity levels and greater potential for cooperation with interdisciplinary knowledge. Trend analysis is a classic quantitative forecasting method that regresses keyword frequencies across different time points on a temporal scale, using least squares to fit a line to historical data and analyzing the rate of change to predict future development trends [23]. Scholars have used this algorithm to determine the development trends of academic terms [19, 24]. This study employs trend analysis to define the Target Discipline Knowledge Node Influence Index AI, which describes the activity level of disciplinary knowledge nodes by judging their 热度变化趋势:

$$AI = \frac{\sum_{y=1}^{Y}(y \times F_y) - \sum_{y=1}^{Y} y \times \sum_{y=1}^{Y} F_y}{\sum_{y=1}^{Y} y^2 - (\sum_{y=1}^{Y} y)^2} \times \frac{F_1 + F_2 + \cdots + F_N}{\sqrt{F_1^2 + F_2^2 + \cdots + F_N^2}}$$

In Equation (1), Y represents the time span of the data sample (in years), and F_y is the research frequency of keyword a_i in the target discipline in year y. AI is the slope (rate of change) of the fitted line. If $AI > 0$, the keyword's research 热度 shows an upward trend, and larger AI values indicate greater rates of change and higher activity levels.

2.3.2 Knowledge Medium Influence Index *BI* According to weak tie theory, connections between interdisciplinary knowledge are more difficult to establish than those within a discipline. As a key bridge-builder, knowledge medium b_i that is associated with more disciplines can access more diverse information, increasing the likelihood of stimulating knowledge innovation. Therefore, the interdisciplinary degree of a knowledge medium is an important manifestation of its mediating capability. This study draws on the disciplinary distribution diversity measurement index proposed by A.L. Porter et al. [16] to define the Knowledge Medium Influence Index BI:

$$BI = \frac{(\sum_{n=1}^{N} F_n)^2}{\sum_{n=1}^{N} F_n^2}$$

In Equation (2), if knowledge medium b_i appears in N disciplines, F_n represents the number of academic papers studying b_i in the n-th discipline. The sum of $F\_1, F\_2, \ldots, F\_N$ is the total number of papers on b_i across all N disciplines. $BI \geq 1$, with larger values indicating higher cross-disciplinary diversity and greater influence. When b_i appears only in literature from one discipline, $BI = 1$.

2.3.3 Interdisciplinary Knowledge Correlation Index *CI* This index measures the strength of the weak tie connection between interdisciplinary knowledge c_i and disciplinary knowledge node a_i through knowledge medium b_i, reflecting the degree of correlation between a_i and c_i. Let $I\_\{ab\}$ be the citation/cited frequency between a_i and b_i, and $I\_\{bc\}$ be that between b_i and c_i. The correlation degree between nodes c_i and a_i is positively correlated with both $I\_\{ab\}$ and $I\_\{bc\}$. Based on this, we define the Interdisciplinary Knowledge Correlation Index CI according to the feasible connection strength defined in previous research [13]:

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

To ensure Equation (3) is meaningful, we introduce β: when $I\_\{ab\} = I\_\{bc\}$, $\beta = 1$; when $I\_\{ab\} \neq I\_\{bc\}$, $\beta = 0$.

2.3.4 Interdisciplinary Knowledge Combination Potential Cooperation Index *P* In the 17th century, Newton's law of universal gravitation provided a method for measuring spatial interactions [25]:

$$I_{ij} = K \times \frac{M_i \times M_j}{d_{ij}^2}$$

where $I\_\{ij\}$ represents the gravitational force between points i and j; M_i and M_j are the masses of points i and j; $d\_\{ij\}$ is the shortest distance between i and j; and K is the gravitational coefficient.

In interdisciplinary knowledge combinations, the possibility of cooperation between disciplinary knowledge node a_i and interdisciplinary knowledge c_i through knowledge medium b_i can be viewed as the gravitational attraction between nodes a_i and c_i in the entire citation association network, influenced by *AI*, *BI*, and *CI*. Therefore, based on the gravity model, we define the Interdisciplinary Knowledge Combination Potential Cooperation Index P as:

$$P = AI \times BI \times CI = \frac{\sum_{y=1}^{Y}(y \times F_y) - \sum_{y=1}^{Y} y \times \sum_{y=1}^{Y} F_y}{\sum_{y=1}^{Y} y^2 - (\sum_{y=1}^{Y} y)^2} \times \frac{(\sum_{n=1}^{N} F_n)^2}{\sum_{n=1}^{N} F_n^2} \times \frac{(I_{ab} \times I_{bc})^2}{|I_{ab} - I_{bc}| + \beta}$$

3 Empirical Analysis: Interdisciplinary Knowledge in Information Science

3.1 Data Sources and Preprocessing

This study selected papers published in nine CSSCI-indexed information science journals as empirical samples to verify the feasibility and effectiveness of the interdisciplinary knowledge combination identification method based on weak citation relationships. The nine journals include: *Journal of the China Society for Scientific and Technical Information*, *Information and Documentation Services*, *Library and Information Service*, *Information Studies: Theory & Application*, *Library and Information Knowledge*, *Library and Information*, *Information Science*, *Data Analysis and Knowledge Discovery*, and *Journal of Intelligence*. The time span covers 2015–2019.

1. **Source literature data download**: Using the CNKI database's professional search function, we retrieved 18,052 valid papers published in these nine journals over five years. Bibliographic information including title, keywords, abstract, publication date, and other fields was downloaded in batch as text files and imported into a MySQL relational database.

2. **Reference/citation literature data acquisition**: A Python crawler program was coded to obtain bibliographic data for references and citing literature of the 18,052 papers, including title, keywords, abstract, publication date, and journal name.

3. **Interdisciplinary reference/citation literature matching and preprocessing**: According to the journal-discipline classification in the *Chinese S&T Journal Citation Reports*, we used the method from literature [26] to read the "journal name" field in bibliographic data and match it to determine the discipline affiliation. Invalid literature with citation format errors or missing information, as well as English literature, were removed. Data from interdisciplinary Chinese journal references and citing literature were screened, similarly downloaded as text files, imported into MySQL, and matched with their target discipline data. This yielded 45,086 interdisciplinary references and 40,103 interdisciplinary citing literature.

3.2 Constructing the Information Science Interdisciplinary Knowledge Weak Citation Association Network

Following the method described in Section 2.1, we constructed a knowledge association network linking information science literature keywords, interdisciplinary reference keywords, and interdisciplinary citing literature keywords from the preprocessed sample data. The operational steps were:

1. **Keyword extraction**: Since the number of author-provided keywords is limited and may not fully reflect document content, we extracted title and abstract field data for target discipline literature and its corresponding interdisciplinary references and citing literature from the MySQL database to form a corpus. Using MySQL's text mining options and Python with the Jieba segmentation package, we performed Chinese word segmentation, stop word removal, term frequency statistics, and synonym merging, using discipline-specific standard terms from the Chinese Standard Terminology Database as a segmentation dictionary. Keywords representing each document's core knowledge points were extracted and saved with their corresponding literature, regardless of frequency. This resulted in a set of 21,023 information science keywords (S), 52,805 interdisciplinary reference keywords (R), and 51,917 interdisciplinary citing literature keywords (D).

2. **Citation relationship network construction**: Using MySQL's relational matching methods, we converted document-level citation relationships into keyword-level citation relationships. A Python program was written to traverse all keywords and their citation relationships in sets S, R, and D to construct a keyword-based interdisciplinary knowledge citation network G. This directed weighted network reflects citation relationships between keywords, where direction indicates citation or being cited, and weight represents the frequency of citation/cited behavior.

3.3 Identifying *a-b-c* Weak Relational Structures in the Information Science Interdisciplinary Citation Network

Based on the methodology in Section 2.2, we identified knowledge media b and a-b-c weak relational structures in the information science interdisciplinary citation network G:

1. **Knowledge medium *b* identification**: Interdisciplinary research keywords are important knowledge media for establishing a-b-c weak relationships. We first extracted these by performing co-occurrence analysis on keywords from sets S, R, and D, yielding 4,817 inflow-type knowledge media in Set 1, 2,805 outflow-type knowledge media in Set 2, and 218 flow-type knowledge media in Set 3.
2. **A-b-c weak relational structure identification**: Using each keyword in set S as a starting point, we employed the Apriori association wandering algorithm and coded a Python program to traverse the MySQL database and all association data. This identified all knowledge media b for disciplinary knowledge node a in network G and all interdisciplinary knowledge c across seven connection path types, removing direct citation relationships between a and c. The final result was 657 a-b-c weak relational structures.

3.4 Calculation and Identification of Interdisciplinary Knowledge Combination Indices in Information Science

1. **Index calculation**: For all keywords a representing information science knowledge nodes, we extracted publication year information from the MySQL database, counted annual occurrence frequencies from 2015–2019, and calculated the Knowledge Node Influence Index AI using Equation (1). Results are shown in Column 6 of Table 1 . For all keywords representing knowledge media b, we obtained their complete literature records from MySQL, determined which disciplines each keyword belonged to based on journal-discipline classification, and counted distribution across each discipline. Equation (2) was applied to obtain the Knowledge Medium Influence Index BI, shown in Column 7 of Table 1. A self-coded Python program traversed network G to obtain connection frequencies between b and a ($I\_\{ab\}$) and between b and c ($I\_\{bc\}$). Equation (3) was used to calculate the Interdisciplinary Knowledge Correlation Index CI, shown in Column 8 of Table 1. Finally, Equation (4) was applied to calculate the Potential Cooperation Index P for a and c in each a-b-c weak connection, shown in Column 9 of Table 1.
2. **Interdisciplinary knowledge combination identification results**: Ranked by Potential Cooperation Index P values from high to low, the top 10 results are presented in Table 1.

Table 1 Partial identification results for interdisciplinary knowledge combinations in information science

Rank	Target Discipline Knowledge Node a	Knowledge Medium b	Interdisciplinary Knowledge c	Flow Path	AI	BI	CI	P
1	Scientific research cooperation	Knowledge flow	Population dynamics model	$a\ b \leftarrow c$	2.60	2.73	5.33	37.35
2	Citation network	Multi-data fusion	Finite element analysis method	$a \leftarrow b \leftarrow c$	2.45	2.93	4.99	35.85
3	Altmetrics	Big data	Entity similarity	$a \leftarrow b \leftarrow c$	2.88	2.24	5.04	32.48
4	Think tank	Big data	Multi-label learning	$a \leftarrow b\ c$	2.88	2.24	4.79	30.87
5	Interdisciplinary	Big data	Interpersonal intelligence network	$a \leftarrow b \leftarrow c$	2.88	2.24	4.62	29.82
6	Knowledge graph	MongoDB database	Hyperstring gravity model	$a\ b \rightarrow c$	2.45	2.93	4.07	29.25
7	High-cited literature	Random forest algorithm	Sleeping beauty literature	$a\ b \leftarrow c$	2.45	2.93	3.73	26.77
8	Link prediction	Random forest algorithm	Entity similarity	$a\ b \rightarrow c$	2.45	2.93	2.87	20.60
9	Knowledge flow	Association rules	Multi-label learning	$a \leftarrow b \rightarrow c$	2.60	2.73	2.58	18.37
10	Scientific research cooperation	Knowledge flow	Population genetics	$a \rightarrow b \leftarrow c$	2.60	2.73	2.25	16.02

4 Analysis of Identification Results

4.1 Effectiveness Analysis of Identification Results

4.1.1 Target Discipline Knowledge Node Influence Index *AI* Effectively Identifies Rising Disciplinary Hotspots In Table 1, Column 2 shows keywords representing target discipline knowledge nodes a. Keywords such as "Altmetrics," "think tank," "scientific research cooperation," and "interdisciplinary" have relatively high AI values in Column 6, indicating that research literature on these topics in information science showed an increasing year-by-year trend from 2015–2019, making them disciplinary research hotspots.

These results align with part of the "rising trend" research hotspots identified in literature [27]. "Sleeping beauty literature" and "high-cited literature" are also keywords with continuously growing research 热度 and attention in recent information science research [28-30]. This demonstrates that the AI index based on trend analysis is feasible and effective for analyzing rising research hotspots, with identification results having greater potential to introduce concepts, theories, methods, and technologies from other disciplines for interdisciplinary collaboration.

4.1.2 Knowledge Medium Influence Index *BI* Identifies Strongly Interdisciplinary Knowledge Media Examining Columns 3 and 7 in Table 1, the knowledge medium "big data" has the highest BI value of 3.24, indicating its obvious interdisciplinary nature, diverse disciplinary involvement, extensive related knowledge, and strong mediating capability as a powerful "bridge-builder." Research shows that big data, as a product of the in-depth development of the information society, has become a research hotspot in the scientific community, spanning information science, social science, mathematics, education, psychology, economics, and many other fields, demonstrating typical interdisciplinarity [31-32]. This aligns with our research conclusions. Therefore, the BI index can be used to calculate the interdisciplinarity of keywords and identify strong "bridge-builders."

4.1.3 Interdisciplinary Knowledge Correlation Index *CI* Effectively Identifies Interdisciplinary Knowledge Highly Relevant to the Target Discipline In Table 1, Columns 4 and 8 show that keywords such as "finite element analysis method," "random forest algorithm," "multi-label learning," and "population dynamics model" have relatively high CI values. This indicates they establish strong weak-tie connections with information science knowledge nodes through knowledge media "multi-data fusion," "link prediction," "association rules," and "knowledge flow," respectively. For example, in the chain "citation network" ← "multi-data fusion" ← "finite element analysis method" with flow path $a \leftarrow b \leftarrow c$, "citation network" frequently cites "multi-data fusion," which in turn frequently cites "finite element analysis method," suggesting that "finite element analysis method" can be applied to "citation network" research. Literature analysis confirms their relevance (see Section 4.2). Therefore, the CI index can effectively identify interdisciplinary knowledge highly relevant to the target discipline.

4.1.4 Potential Cooperation Index *P* Effectively Combines Characteristics of *a-b-c* Weak Connections to Identify High-Potential Interdisciplinary Knowledge Combinations In Table 1, the a-b-c weak relationship "scientific research cooperation" "knowledge flow" ← "population dynamics model" ranks first in P value. Its $AI = 2.60$, $BI = 2.73$, and $CI = 5.33$ rank 4th, 3rd, and 2nd, respectively. Although not the highest in target discipline knowledge node activity, knowledge medium diversity, or interdisciplinary knowledge correlation, all three values are relatively high. Consequently, the attraction between "scientific research cooperation" and "population dynamics model" is strongest, with the closest potential connection and the greatest weak-tie connection potential ($P = 37.35$), making it the highest-potential interdisciplinary knowledge combination. Thus, the P index can effectively combine node attributes and relationships in a-b-c weak ties to identify interdisciplinary knowledge combinations with high collaborative potential.

4.2 Application Analysis of Identification Results

Table 1 presents interdisciplinary knowledge combinations with high collaborative potential values. To explore the effectiveness of these results and analyze technical solutions for applying interdisciplinary knowledge to solve information science problems, thereby promoting interdisciplinary research innovation, we analyze the top two results:

1. **"Scientific research cooperation" "knowledge flow"** ← **"population dynamics model."** Scientific research cooperation enables the integration of different knowledge and facilitates knowledge collisions among researchers, organizations, or institutions with different knowledge backgrounds, making it an important way to accelerate knowledge diffusion [33]. From the perspective of research cooperation, collaborative parties acquire complementary knowledge resources based on knowledge potential differences to achieve knowledge flow. The knowledge flow process drives knowledge dissemination, diffusion, integration, and innovation, forming knowledge chains and networks [19]. The population dynamics model is a behavioral dynamics model that studies interactions between populations and between populations and uncertain environments [34]. Knowledge flow has certain associations with population movement and is embedded in the scientific research cooperation process. Therefore, we can attempt to apply the population dynamics model to the scientific research cooperation process to simulate and analyze knowledge diffusion patterns and evolutionary paths of knowledge integration, 挖掘 research cooperation patterns, relationship structures, and status changes to reveal underlying laws of scientific research cooperation development.

2. **"Citation network"** ← **"multi-data fusion"** ← **"finite element analysis method."** A citation network is a large-scale knowledge network composed of research literature containing massive citation relationships and text attributes. "Multi-data fusion" refers to using data fusion algorithms to efficiently integrate multiple types of association data, thereby extracting and identifying more accurate potential semantic relationships between knowledge units through richer information [35]. The "finite element analysis method" divides the solution domain into many small interconnected subdomains called finite elements, assumes an appropriate approximate solution for each element, and then derives the solution satisfying overall conditions to solve the problem [36]. Citation

networks contain various complex networks. Following the finite element analysis method's 思想, these complex networks can be decomposed into multiple simple subnetworks, with multi-data fusion effectively integrated for each subnetwork to simulate and achieve the overall network. Therefore, we can attempt to apply the finite element analysis method to fuse multi-source data in citation networks, improving the efficiency of citation network data mining and knowledge discovery.

5 Conclusion

This study proposes a method for discovering interdisciplinary knowledge combinations based on weak citation relationships. First, it constructs a keyword-based interdisciplinary knowledge weak citation association network. Second, it identifies a-b-c weak relational structures in the citation network and defines the Knowledge Node Influence Index AI, Knowledge Medium Influence Index BI, and Interdisciplinary Knowledge Correlation Index CI to construct the Potential Cooperation Index P model. Finally, it conducts an empirical study using papers from nine leading information science journals (2015–2019) and their interdisciplinary references/citations. Effectiveness analysis demonstrates the rationality and validity of the proposed method, while application analysis shows that the identification results have significance for interdisciplinary collaboration.

The empirical sample in this study comes from Chinese literature in information science. Future research should verify whether this method applies to other disciplines or literature in other languages. Additionally, the development of informal academic communication methods such as social media provides new research perspectives and data sources for interdisciplinary knowledge exchange and collaboration. Therefore, future studies can utilize cross-language literature and social media data to further investigate interdisciplinary knowledge combination identification for more comprehensive insights.

References

[1] GATES A J, KE Q, VAROL O, et al. Nature's reach: narrow work has broad impact[J]. Nature, 2019, 575(7781): 32-34.

[2] GRANOVETTER M. The strength of weak ties[J]. The American journal of sociology, 1973, 78(6): 1360-1380.

[3] ONNELA J P, SARAMAKI J, HYYONEN J, et al. Structure and tie strengths in mobile communication networks[J]. Proceedings of the National Academy of Sciences of the United States of America, 2007, 104(18): 7332-7336.

[4] BAKSHY E, ROSENN I, MARLOW C, et al. The role of social networks in information diffusion[C]//Proceedings of the 21st annual conference on World

Wide Web. Lyon: Annual conference on World Wide Web, 2012: 519-528.

[5] ZHAO J, WU J, XU K. Weak ties: subtle role of information diffusion in online social networks[J]. Physical review E, 2010, 82(1): 16105.

[6] DAVID E, JON K. Networks, crowds, and markets: reasoning about a highly connected world[M]. United Kingdom: Cambridge University Press, 2012.

[7] GENIUS S K. Published literature and diffusion of medical innovation: exploring innovation generation[J]. The Canadian journal of information and library science, 2005, 29(1): 27-54.

[8] ABBASI A, ALTMANN J, HOSSAIN L. Identifying the effects of co-authorship networks on the performance of scholars: a correlation and regression analysis of performance measure and social network analysis measures[J]. Journal of informetrics, 2011, 5(4): 594-607.

[9] BETTONI M, SCHILLER G, BERNHARD W. Weak ties cooperation in the CoRe knowledge network[C]//The 9th European conference on knowledge management. Southampton: Southampton Solent University, 2008: 1-7.

[10] YANG L Y, MORRIS S. Mapping institutions and their weak ties in a specialty: a case study of cystic fibrosis body composition research[J]. Scientometrics, 2009, 79(2): 421-434.

[11] WEI L, XU H Y, GUO T, et al. Interdisciplinary pattern analysis of information science based on weak co-occurrence and burst detection[J]. Journal of Intelligence, 2018, 37(9): 179-184.

[12] LI C L, LIU X H, LIU Y M, et al. Potential interdisciplinary research topic identification between information science and computer science based on open non-related knowledge discovery[J]. Information Studies: Theory & Application, 2018, 41(2): 100-104.

[13] LIU X H, LI C L, CUI B, et al. Potential interdisciplinary research topic identification between information science and computer science based on closed non-related knowledge discovery[J]. Library and Information Service, 2015, 59(21): 199-213.

[14] SONG M, KANG K, YOUNG A. Investigating drug-disease interactions via citation relations[J]. Journal of the Association for Information Science and Technology, 2018, 69(11): 1355-1368.

[15] LIANG Z T, BA Z C, XU J. Development path analysis of interdisciplinary fields based on citations: a case study of eye-tracking research[J]. Library and Information Service, 2019, 63(23): 65-78.

[16] PORTER A L, COHEN A S, ROESSNER D J, et al. Measuring researcher interdisciplinarity[J]. Scientometrics, 2007, 72(1): 117-147.

[17] LIN X Y, LUO H Y, LI X C. Mechanism of enterprises absorbing heterogeneous knowledge from social media networks: from the perspective of weak

tie relationships[J]. Information Studies: Theory & Application, 2019, 42(3): 65-71.

[18] DU D H, LI C L, XIANG F Z, et al. Discussion on interdisciplinary knowledge discovery method based on citation keywords[J/OL]. Journal of Intelligence: 1-6[2020-09-21]. http://kns.cnki.net/kcms/detail/61.1167.G3.20200721.1527.004.html.

[19] ZHANG R, ZHAO D X, TANG X L, et al. Research on cross-disciplinary migration and development of academic terms from the perspective of knowledge flow[J]. Information Studies: Theory & Application, 2020, 43(1): 47-55.

[20] LOUADI M, MOHAMED E L. Knowledge heterogeneity and social network analysis-towards conceptual and measurement clarifications[J]. Knowledge management research & practice, 2008, 6(3): 199-213.

[21] WANG X, LI R. Interpretation of various citation phenomena from the perspective of knowledge ecology[J]. Journal of Intelligence, 2018, 37(9): 179-184.

[22] FAJARDO-ORTIZ D, LOPEZ-CERVANTES S, DURAN L, et al. The emergence and evolution of the research fronts in HIV/AIDS research[J]. Plos one, 2017, 12(5): e178293.

[23] ZHOU F G, HUANG Y C. Applied linear regression analysis[M]. Beijing: China Renmin University Press, 1989.

[24] LI F, LIU J Y, JIANG L Q. Development and latest trends in forecasting methods[J]. Journal of Intelligence, 2005(6): 76-77.

[25] NEWTON. Mathematical principles of natural philosophy[M]. ZHAO Z J, trans. Beijing: The Commercial Press, 2009.

[26] FENG Z G, LI C L, LIU X H, et al. Interdisciplinarity analysis of library and information science based on citation and cited literature information[J]. Information Science, 2018, 36(3): 105-114.

[27] LI C L, PAI Y X, XIANG F Z, et al. Discussion on improved z-index method for identifying highly cited disciplinary research hotspots[J]. Information Studies: Theory & Application, 2020(6): 39-45, 66.

[28] LI Y L, ZHANG X T, WANG S S, et al. The persistence and expansion of information science: a review based on 2018 ASIS&T annual conference papers[J]. Library and Information Knowledge, 2019(3): 4-16.

[29] LIU Y M, LI C L, DU D H. Construction and application of a time series perspective model for identifying PLOS ALM indicator characteristics[J]. Information and Documentation Services, 2019, 40(6): 6-13.

[30] ZONG Z J. Research progress on sleeping beauty literature identification methods[J]. Library and Information Service, 2019, 63(16): 132-142.

[31] LV X Z, WANG H, ZHOU P. Comparative study on the interdisciplinarity of big data papers between China and the United States[J]. Science Research

Management, 2019, 40(4): 1-13.

[32] LI G J, CHENG X Q. Big data research: a major strategic area for future science, technology, and economic and social development—current status and scientific thinking on big data research[J]. Bulletin of Chinese Academy of Sciences, 2012, 27(6): 647-657.

[33] LI G, BA Z C. Research on knowledge diffusion evolution model in scientific research cooperation supernetwork[J]. Journal of the China Society for Scientific and Technical Information, 2017, 36(3): 274-284.

[34] KRIVAN V, LEWIS M, BENTZ B J, et al. A dynamical model for bark beetle outbreaks[J]. Journal of theoretical biology, 2016, 407(21): 25-37.

[35] CHEN W J, XU H Y. A knowledge representation method for citation networks based on multi-data fusion[J]. Information Studies: Theory & Application, 2020, 43(1): 150-154.

[36] BREKELMANS A M, POORT H W, SLOOFF F J. A new method to analyse the mechanical behaviour of skeletal parts[J]. Acta orthopaedica scandinavica, 2009, 43(5): 301-317.

Author Contributions

牌艳欣 (Pai Yanxin): Data collection and processing, manuscript writing and revision
李长玲 (Li Changling): Research design and guidance, manuscript writing and revision
徐璐 (Xu Lu): Manuscript revision and proofreading

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