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

Potential influential topics in information science based on citation structural variation approach

Jianhua Hou[a], Xuran Tian[b*], Yang Zhang[a]

a. School of Information Management, Sun Yat–sen University, Guangzhou, China
b. School of Information Management, Central China Normal University, Wuhan, China

ABSTRACT

This study aims to identify potential influential literature and potential influential topics in the field of information science (IS). Using citation structural variation approach, provided by CiteSpace, we discerned the potential influential theses in IS. Then, we proposed an analyzing framework based on"citation network-citation structural variation-citation cluster," which is used to discern the potential influential topics in IS. Using the structural variation approach, we analyzed the contents and features of potential influential theses. The potential influential topics, discerned by the "citation network-citation structural variation-citation cluster" framework, corroborated the existing knowledge base to some extent and seemed to be divergent and intermingled. Such potential impact literature and potential impact research topics are of certain reference value for future frontier research.

KEYWORDS

Structural Variation Model; Potential Influential Theses; Potential Influential Topics; CiteSpace

1 Introduction

The identification of the frontiers of scientific knowledge and the predictive analysis of its development trends are pressing issues in the field of scientometrics and information science (IS). For researchers in related scientific knowledge fields, research has crucial practical significance to track academic frontiers and monitor future development directions. With the continuous progress of literature information visualization technology, the analysis assisted by computer programs provides critical reference values for researchers to make subjective research judgments on the research field. The existing research (Tahamtan et al., 2016) generally claims that the literature on the research topic with potential influence is reflected by the literature citation, which is affected by three main factors, namely, literature-related factors, journal-related factors, and authors (Bjarnason & Sigfusdottir, 2002). Of these, related factors, the literature of people with high academic influence, is more likely to garner widespread attention, and its subsequent research topics are more likely to become potentially influential topics.

In 2012, Chaomei Chen, a Chinese scholar at Drexel University in the United States (Chen, 2012), projected a theoretical and computational model, called the structural variation mod-

*Corresponding author: tianxuran@mails.ccnu.edu.cn

el, to predict the potential of the cited literature to alter the degree of structure of the original knowledge base network. The primary focus is to introduce new citation links to the knowledge-based cyberspace by focusing on new literature, estimate the potential influence of newly introduced literature on the future development of the knowledge domain by the boundary crossing effect of citation links, and analyze the reasons for the literature to gain high citations from both intrinsic and extrinsic aspects. Regarding more dominant intrinsic factors, three metrics of structural variation that can determine potentially influential literature are proposed, namely, modularity change rate (ΔM), cluster linkage (ΔC_{lw})and centrality divergence (ΔC_{kl}), and the predictive effects of the three indicators in different domains are validated.

Accordingly, this study examines potentially influential literature in the IS based on the structural variation model, as well as combines the three metrics of structural variation to perform a cluster analysis of the literature, based on the frame of "citation network– citation structural variation– citation cluster." The major contributions of this research include the following: (i) propose an analysis framework to determine potentially influential research topics based on the "citation network– citation structural variation– citation cluster" framework; (ii) using structure variation models, an empirical analysis of potentially influential research themes in IS was performed.

2 Methods

As a data source, a two-step approach was used to select the influential literature in the field of IS in recent years. Using the structural variation model, we estimated the potential influence research topics in this field. Based on the research framework of "citation network– citation structural variation– citation cluster," we analyzed potential influential literature and research topics in the field of IS.

2.1 Data Retrieval and Processing

In this study, we primarily adopted a two-step method for the selection and processing of research data. The first step involved selecting core journals that can represent a high influence in the field of IS. We used *JASIST, Scientometrics, and Journal of Informetrics* as basic data of journals. The second step involved using the journal co-citation analysis (JCA) function of the CiteSpace to select journals with a higher co-citation frequency as the data source of journals. In addition, we excluded comprehensive and non-IS journals (e.g., *Science and Nature*), and finally identified nine journals as the data sources (Table 1).

Table 1 The nine most frequently co-cited journals

co–cited	publisher
1665	SCIENTOMETRICS
1409	JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY
1067	JOURNAL OF INFORMETRICS
685	JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY
667	RESEARCH POLICY
625	JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE
550	INFORMATION PROCESSING & MANAGEMENT

co–cited	publisher
394	JOURNAL OF DOCUMENTATION
391	RESEARCH EVALUATION
278	ANNUAL REVIEW OF INFORMATION SCIENCE AND TECHNOLOGY
228	JOURNAL OF INFORMATION SCIENCE

Note: The renames of journals in *JAIST* are listed separately: The *Journal of the American Society for Information Science* has been changed to bimonthly since 1970, and was renamed the *Journal of the American Society for Information Science and Technology* in January 2001, which then in 2014 was renamed to the current *Journal of the Association for Information Science and Technology (JAIST)* .

A total of nine journals were included in the Web of Science Core Collection database with a total of 12,225 documents, of which we selected only two types of journal articles and reviews, and removed 983 documents of other types, leaving 11,242 documents. Figure 1 shows the year-by-year distribution of the collected data. The trend in Figure 1 illustrates that the paper output in the field of IS shows an overall increasing trend, except for a marginal decline in the number of literature in 2017 and 2018, which basically demonstrates an increasing trend year-by-year, as observed in the field of IS, the amount of each highly co-cited journal exhibits an overall increasing trend, and it is evident that the whole IS is developing at a faster rate.

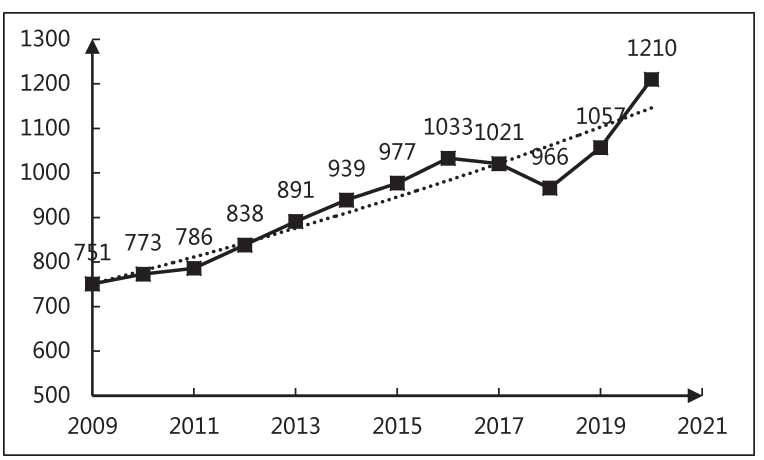

Figure 1 Literature volume line chart 2009– 2020.

2.2 Methods and Analysis Tools

CiteSpace, a freely available Java application to visualize and analyze trends and patterns in scientific literature, is designed as a tool for progressive knowledge domain visualization. The specific analysis of the literature data involves using the citation structure variation model in the CiteSpace 5.7 software, which predominantly relies on three metrics (i.e., ΔM, ΔC_{lw} and ΔC_{kl}) to examine the influence of the existing co-citation network on the citing literature. The ?M characterizes the upsurge in connections of the knowledge-based network due to citing documents ; these increased connections could appear in the same cluster or between different clusters. The higher the absolute value of ΔM, the greater the influence of the cited literature on the structural variation of the network, that is, the greater the potential influ-

ence of the cited literature on the development of the discipline, and the more likely such cited literature is to be potentially influential in this field. The ΔC_{lw} mainly reflects the span of the fundamental network node connection between different clusters due to citing documents. As mentioned before, a higher value of ΔC_{lw} indicates that the citation links of the cited literature span more between different clusters, suggesting that the cited literature captivates a multidisciplinary subject knowledge base, and the stronger the cross-cutting properties of such a cited literature. Besides, the more likely it is to be a potential force resulting in changes in the underlying network structure. Moreover, the more likely it is to be the potentially influential literature with the greatest degree of innovation. Furthermore, ΔC_{kl} measures the degree of variation in the mediated centrality distribution of nodes in the underlying network due to the cited literature. The higher the value of ΔC_{kl}, the greater the influence of the cited literature on the divergence of centrality of the original nodes in the underlying network, that is, the greater the influence on the change in the structure of the underlying network.

Using the structural variation model, the prediction of potentially influential literature was based on the three metrics mentioned above, and the prediction of potentially influential topics was based on the theme words extracted from the frame of "citation network– citation structural variation– citation cluster."

3 Results

3.1 The Knowledge Structure of Research Topics in IS

Using the CiteSpace5.7 software and the method of document co-citation analysis (DCA), we set LRF = 3, LBY = 10, Maximum Links Per Node = 5.0, and Nodes Labeled = 1.0% on the Project interface. The citation network finally presents 11 clusters, of which the first eight clusters were selected in this study. The size of these eight clusters accounted for 96.05% of all clusters, which can reflect the overall situation of research in this field.

Table 2 The eight largest clusters under the literature co-citation network

Cluster ID	Size	Silhouette	Label (LSI)	Label(LLR)	Topics
#0	118	0.627	comprehensive analysis;japanese national doctoral survey;knowledge diffusion	research performance (7180.5, 1.0E –4);citation pattern (6246.07, 1.0E –4); long –term influence (6246.07, 1.0E–4)	research performance
#1	107	0.735	interactive overlays; scopus;aggregated journal –journal citation relations	bibliometric mapping (6808.2, 1.0E–4);innovation studies (6485.44, 1.0E –4);hirsch index (5148.44, 1.0E–4)	bibliometric mapping
#2	106	0.763	agent –based model; garfield readers; gender gap	mendeley reader (11176.94, 1.0E–4);later citation count (9074.68, 1.0E –4);social science research network (7765.65, 1.0E–4)	altmetrics
#3	97	0.62	aggregated journal –journal citation relations;interactive overlays;related impact measures	new crown indicator (9135.66, 1.0E –4);comparing set (6924.24, 1.0E –4); fractional counting (6322.97, 1.0E–4)	bibliometric indicators

Cluster ID	Size	Silhouette	Label (LSI)	Label(LLR)	Topics
#4	82	0.875	h–index;elite marketing scholars;exploring bibliometric properties	hirsch index (35176.95,1.0E –4); broad review (22005.55,1.0E –4); new bibliometric indicator(22005.55, 1.0E–4)	hirsch index
#5	42	0.99	innovating knowledge communities;group collaboration;wine sectors	reassessing academic entrepreneurship (4848.05, 1.0E –4); inventor ownership (4743.49, 1.0E–4); research–derived entrepreneurship(4743.49, 1.0E–4)	innovation management
#6	19	0.914	publication patterns; senior authors;computer science journals	social science (2235.35,1.0E –4); CRIS data (933.15, 1.0E–4);registered output (933.15, 1.0E–4)	current research information system(CRIS)
#7	13	0.972	case study;journal impact factor;two –stage publication process	national research assessment exercises (2250.29, 1.0E–4);evaluating research (847.33, 1.0E –4);education system (596.87, 1.0E–4)	research assessment exercises

Combining the subject identifiers extracted from Table 2, we summarized eight research topics in the document co-cited network, namely, #0 research performance, #1 bibliometric mapping, #2 Altmetrics, #3 bibliometric indicators, #4 Hirsch index, #5 innovation management, #6 current research information system (CRIS, and #7 research assessment exercises. Of note, the clusters of the first three research topics are all over 100.

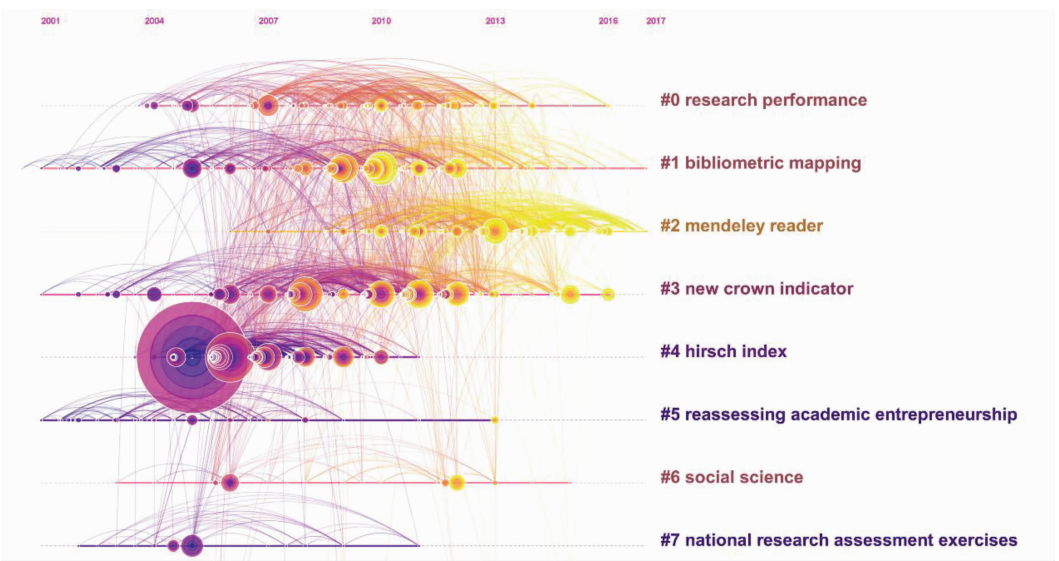

Figure 2 Timeline view of the co-citation network for the eight largest clusters.

From the timeline view (Fig. 2) of the running results of CiteSpace5.7, the four clusters (i.e., #0 research performance, #1 bibliometric mapping, #2 Altmetrics, and #3 bibliometric indicators) formed later. Of these, #2 Altmetrics was the latest to form and continues to this day. The recent development speed of Altmetrics is fast, and the research content of related networks is very extensive, which is the frontier research topic in this field. Lines in Fig. 2 depict

lines of co-cited documents; while the horizontal lines indicate the citation relationships between the same topic clusters, the vertical lines indicate the citation relationships between different topic clusters. During the theme development, the five clusters (i.e., #0 research performance, #1 bibliometric mapping, #2 Altmetrics, #3 bibliometric indicators and #4 Hirsch index) developed more self-citations, implying that these themes needed their subject knowledge as the basis for its research and development. Notably, the vertical line denotes the connection between different clusters. The connection between two different clusters is formed through the co-citation relationship of papers under different themes, which has the characteristics of marginal topics. In addition, #5 innovation management, #6 CRIS, and #7 the clusters of the three themes of scientific research evaluation have more connections with other themes, of which #4 Hirsch index has more horizontal and vertical lines. Notably, #4 Hirsch index was formed earlier, and its node is the largest in the graph, suggesting that #4 Hirsch index exerts the greatest influence on these topic clusters, the highest degree of relevance, and is the easiest to promote the development of other topic clusters. For example, #4 Hirsch index and #0 research performance; the vertical line between #4 Hirsch index and #2 Altmetrics. Furthermore, the development of these two themes is partly driven by research related to #4 Hirsch index.

3.2 Analysis of Potentially Influential Papers in IS

We used the structural variation approach of CiteSpace5.7 software to predict potentially influential documents. The higher the value of ΔM, ΔC_{lw} and ΔC_{kl}, the higher the potential influence of the cited literature. Table 3 presents the 10 cited documents with the highest values of each of the three metrics in this co-citation network.

Table 3 The top 10 citing documents with the largest three index values

ΔM	literature	ΔC_{lw}	literature	ΔC_{kl}	literature
22.33	EGGHE L, 2010 (a)	0.29	NORRIS M, 2010 (a)	0.84	KURTZ MJ, 2010
20.51	NORRIS M, 2010 (a)	0.25	ALONSO S, 2010	0.56	SCHUBERT A, 2010
8.57	ALONSO S, 2010	0.12	FRANCESCHINI F, 2010	0.46	LEYDESDORFF L, 2010
8.11	FRANCESCHINI F, 2010	0.11	EGGHE L, 2010 (a)	0.46	OPTHOF T, 2010
7.68	MOUSSA S, 2010	0.08	NORRIS M, 2010 (b)	0.37	VIEIRA PC, 2010
3.83	NORRIS M, 2010 (b)	0.07	MOUSSA S, 2010	0.32	ALONSO S, 2010
3.71	WALTMAN L, 2016	0.06	LAZARIDIS T, 2010	0.3	SCHMOCH U, 2010
3.26	ZHANG L, 2011	0.05	AHLGREN P, 2010	0.28	BRESCHI S, 2010
3.08	AHLGREN P, 2010	0.05	EGGHE L, 2010 (b)	0.24	TONTA Y, 2010
3.03	BORNMANN L, 2014	0.05	LEYDESDORFF L, 2013	0.23	PERC M, 2010

Based on the potentially influential literature, we further analyzed the research topics related to these potentially influential literature per the structural variation network. As shown in Figs. 3 and 4, solid pink lines denote the existing connections in the citation network, while red dashed lines reflect the new connections in the citation network.

As ΔM and ΔC_{lw} share some similarity in the calculation methods, seven of the top 10 literature of ΔC_{lw} were the same as the top 10 citations of ΔM. Moreover, only one of the top 10 literature of ΔC_{kl} is repeated with the previous one, which is ALONSO S, 2010.

The two citing documents with the highest Δ *M* were EGGHE L, 2010(a) and NORRIS M, 2010(a), with Δ *M* reaching 22.33 and 20.51, respectively, far exceeding the Δ *M* of other citing documents. The emergence of two high Δ *M* EGGHE L, 2010(a) and NORRIS M, 2010(a) further connects the two originally loosely connected topic clusters and account for a large change to the original citation network.

Egghe L, 2010(a) offered a comprehensive summary of the application of Hirsch index. For example, Hirsch index has been applied from the initial assessment of a scholar' s citation influence and prediction of future influence to the influence factors of journals and other different aspects. In the discussion, Egghe (2010) accentuated that while examining the derivative index of Hirsch index (including the increase of time dimension of Hirsch index), in-depth substantive research should be conducted on the measurement of influencing factors. In addition, Norris M, 2010, reviewed the literature on Hirsch index and highlighted that Hirsch index offers the advantage of simple understanding and has a larger application space compared with other derived indexes (Norris & Oppenheim, 2010).

The two literature mentioned above represent the research on science evaluation issues; thus, it can be viewed that discussion on the construction of scientific and objective quantitative science evaluation index system is a relatively crucial content at present. Viewing the new connections made by EGGHE L, 2010(a), it newly connects #1 bibliometric mapping to #4 Hirsch index, while NORRIS M, 2010, newly connects #7 research assessment to #4 Hirsch index; both citations have new connections within the #4 Hirsch index theme. Hence, both the content and the position of the literature in the structure reflects the theme of the literature.

Besides the two cited literature with a Δ *M* of>20, three literature had a Δ *M* of>5, namely, Alonso S, 2010, Franceschini F, 2010, and Moussa S, 2010. Other than the newly connected topic cluster, the correlation between #3 bibliometric indicators and #4 Hirsch index was added in Alonso S, 2010. In addition, FRANCESCHINI F, 2010 newly connected three new clusters of #3 bibliometric indicators with #4 Hirsch index; #7 scientific research assessment with #4 Hirsch index; and #3 bibliometric indicators with #7 scientific research assessment. Furthermore, Moussa S, 2010 only strengthened the connection between #4 Hirsch index cluster itself.

The Δ *M* of the other five pieces of literature was all small (range: 3– 4). The cluster topic of the new link was the same as the previous literature. Except for WALTMAN L, 2016, a comprehensive analysis at the top 10 citations revealed that most connected new thematic clusters while constructing new links within the #4 Hirsch index itself. WALTMAN L, 2016, evolved links to multiple other clusters based on the links between #3 bibliometric indicators topic clusters, the most across clusters in this 10-article literature. Perhaps, this is one of the reasons why this paper is one of the most recently published paper with potential impact.

The mining of the literature content revealed that the part of literature with a higher value of Δ *M* was mostly review, which always received more attention (Biscaro & Giupponi, 2014), and in the structural variation model, review tended to have higher values, as it tends to synthesize more fields and, thus, has more connections across boundaries (Chen, 2012).

KURTZ MJ, 2010, was the highest citation with a Δ C_{kl} value of 0.84, newly connecting #1 bibliometric mapping and #3 bibliometric indicators thematic clusters, and cluster #1 bibliometric mapping and #3 bibliometric indicators themselves were reinforced. The content, SCHUBERT A, 2010, was the same as the review type of literature. In addition, Kurtz and

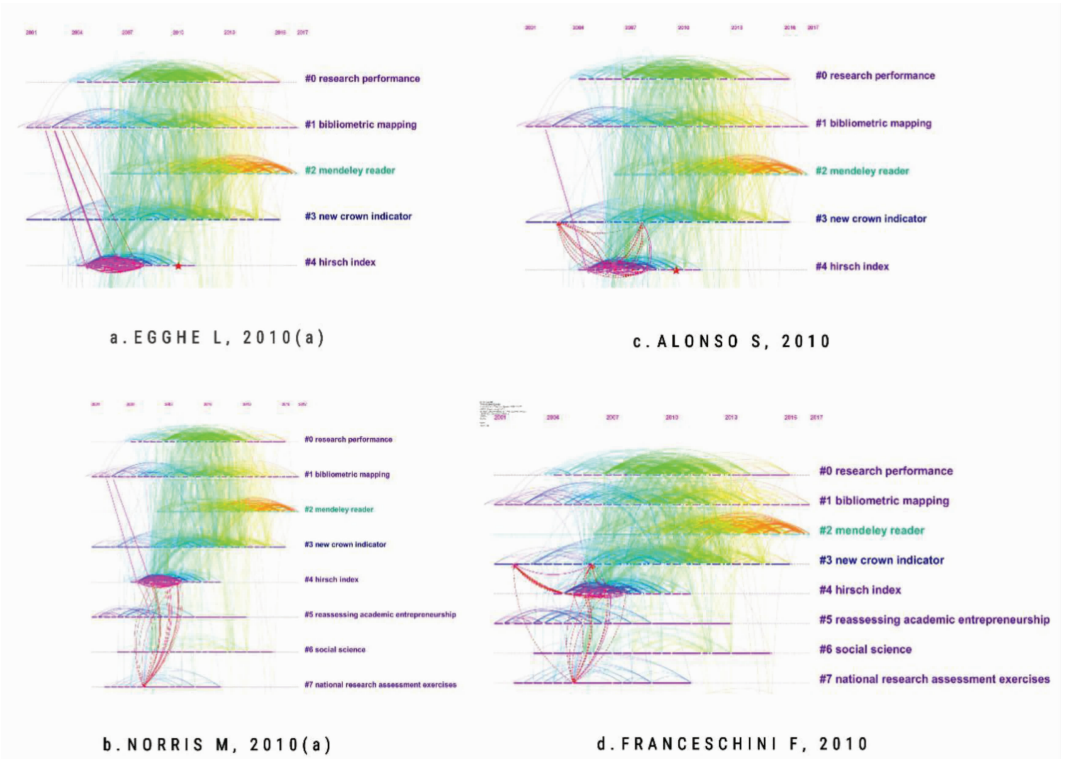

Figure 3 Diagram of the top four papers with the largest Δ *M* values.

Bollen (2010) addressed the current inadequate assessment of the impact factor of journals, and the definition of bibliometrics warrants expansion in this context as the way of accessing and reading literature today has evolved intensely and the use of data was constantly updated. SCHUBERT A, 2010, proposed an H-similarity measure for the citation impact of journals (Schubert, 2010); this standardization enabled journals from disciplines with lower average citation level (mathematics and engineering) to get to the top of the ranking in the journal evaluation process. In a way, this remains a further modification of the application of the Hirsch index to the impact judgment of journals, further improving the limitations of the existing journal impact factors. Besides, this literature is a further link between the Hirsch index and the newly evolved index from the Hirsch index. LEYDESDORFF L, 2010, explored the differences in the graphs presented by different data sources, using data from different databases (Leydesdorff et al., 2010). Furthermore, OPTHOF T, 2010, questioned the standardization of assessment research performance advocated by Leiden University, claiming that although any indicator is widely available for use in the policy-making or management process, the indicator quality remains significant (Opthof & Leydesdorff, 2010). Furthermore, VIEIRA PC, 2010, explored differences between the three subject areas of finance, management, and marketing using JCA, which is essential for the classification of subjects and the scientific evaluation of higher education institutions and subject rankings (Vieira & Teixeira, 2010).

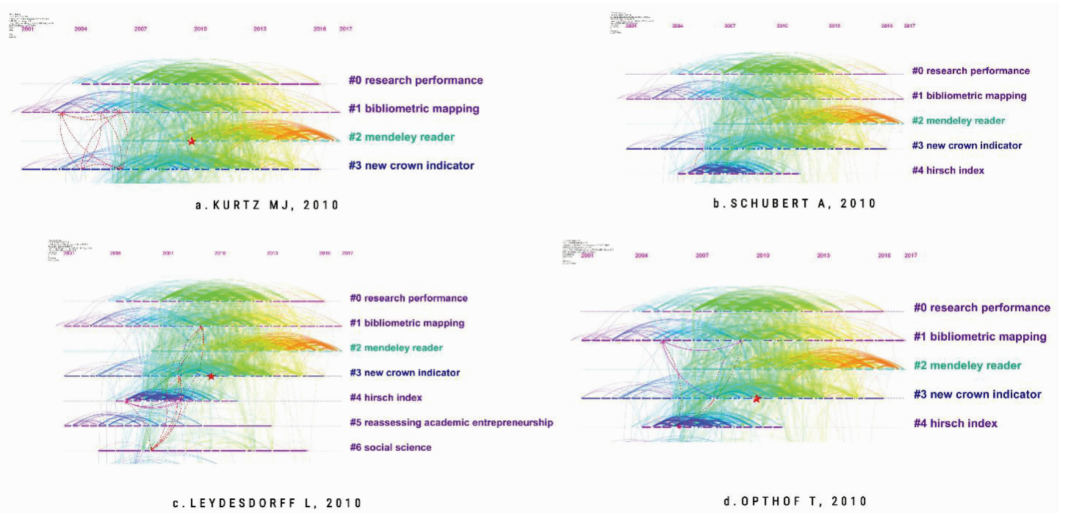

Figure 4 Diagram of the top four papers with the largest Δ C_{kl} values.

Synthesizing the above literature, the study of high potential impact literature primarily involved bibliometric indicators, journal evaluation, institutional knowledge base, policy formulation and management process, scientific evaluation of universities, and research output.

3.3 Analysis of Potential Impact Research Themes

Using potentially influential literature as the basis for potential research topic prediction, together with the index of structural variation model, can make the predicted potentially influential research topics more objective and effective. Of these, for the selection of potentially influential literature, we mainly selected the literature with the top 20% of the absolute value of Δ *M*. The reasons mainly included, (i) the selection of the indicator of Δ *M*, which has mostly included the literature with two indicators; (ii) the new inclusion of the literature with Δ *M*<0, but whose absolute value is also in the top 20%, which can make a more comprehensive analysis and prediction.

Table 4 Literature with the top 20% absolute rate of modularity change rate (Δ *M*)

Δ*M*	year	literature
22.326	2010	EGGHE L 2010 ANNU REV INFORM SCI V44 P65 DOI 10.1002/aris.2010.1440440109
20.5062	2010	NORRIS M 2010 J DOC V66 P681 DOI 10.1108/00220411011066790
8.572	2010	ALONSO S 2010 SCIENTOMETRICS V82 P391 DOI 10.1007/s11192–009–0047–5
8.1133	2010	FRANCESCHINI F 2010 SCIENTOMETRICS V85 P203 DOI 10.1007/s11192–010–0165–0
7.679	2010	MOUSSA S 2010 J INFORMETR V4 P107 DOI 10.1016/j.joi.2009.10.001
3.8336	2010	NORRIS M 2010 J INFORMETR V4 P221 DOI 10.1016/j.joi.2009.11.001
3.7066	2016	WALTMAN L 2016 J INFORMETR V10 P365 DOI 10.1016/j.joi.2016.02.007
3.2618	2011	ZHANG L 2011 J INFORMETR V5 P583 DOI 10.1016/j.joi.2011.05.004
3.082	2010	AHLGREN P 2010 J AM SOC INF SCI TEC V61 P1424 DOI 10.1002/asi.21333
3.0319	2014	BORNMANN L 2014 SCIENTOMETRICS V98 P487 DOI 10.1007/s11192–013–1161–y

ΔM	year	literature
…	…	…
−0.9139	2010	ZHANG L 2010 SCIENTOMETRICS V82 P687 DOI 10.1007/s11192−010−0180−1
−1.1179	2013	LEYDESDORFF L 2013 J AM SOC INF SCI TEC V64 P96 DOI 10.1002/asi.22765
−1.1524	2011	YAN EJ 2011 J AM SOC INF SCI TEC V62 P1498 DOI 10.1002/asi.21556
−1.1574	2012	LEYDESDORFF L 2012 SCIENTOMETRICS V92 P355 DOI 10.1007/s11192−012−0660−6
−1.2978	2010	TAKEDA Y 2010 SCIENTOMETRICS V83 P783 DOI 10.1007/s11192−010−0158−z
−1.2983	2010	RAFOLS I 2010 J AM SOC INF SCI TEC V61 P1871 DOI 10.1002/asi.21368
−1.3538	2010	KURTZ MJ 2010 ANNU REV INFORM SCI V44 P3
−2.0849	2010	LEYDESDORFF L 2010 J AM SOC INF SCI TEC V61 P352 DOI 10.1002/asi.21250
−2.1457	2010	VAN ECK NJ 2010 SCIENTOMETRICS V84 P523 DOI 10.1007/s11192−009−0146−3
−2.2121	2010	GONZALEZ−PEREIRA B 2010 J INFORMETR V4 P379 DOI 10.1016/j.joi.2010.03.002

Among them, the number of literature with $\Delta M>0$ was 654, and the number of literature < 0 was 291 (Table 4). The top 20% of the literature with > 0 were 131, and the top 20% of the literature with<0 were 58, and the sum of ΔM of these 20% of the literature also accounted for 82.47% and 76.43% of the total literature. Finally, these top 20% of the literature were subjected to DCA, citation clustering was performed, and clustered subject terms were extracted (Table 5).

Table 5 Subject terms of the literature with an absolute value of modularity change rate(ΔM)>0

ΔM	Subject terms
Greater than 0	hirsch index；new bibliometric indicator；network analysis;information science; introducing metaknowledge；computational research;twitter count；twitter index； author −level bibliometric indicator；south africa；bibliometric indexe；national research founda tion；vanclays criticism；journal citation reports databases； mendeley readership count；large −scale analysis；environmental science；google scholar；knowledge network centrality；research performance；interaction effect； top −cited paper；citation analysis；the −art report；citizen bibliometrics；developing field−independent index；indirect citations paradigm；normalized indicator；revisiting country research profile；universal distribution；inferring frontier research；brazilian researcher；bibliometric indicator；large data；ranking marketing journal；scholar − based hg−index；citation−based metrics；journal assessment；scientometrics law； co−author core
Less than 0	computer program；bibliometric mapping；scopus data；journal citation report； journal map；previous mapping approaches；information science；bests performing region；alternative；library management；research policy；journal ranking； interdisciplinary research；scientific prestige；sjr indicator；fractional count

The part of the literature with positive and large values of ΔM exerts a crucial influence on the structure of the entire network and is the most potentially influential; hence, the subject terms in this part plays a crucial role in the prediction of potential influential topics. Among them, the highest frequency of subject terms was explored in the index category, including Hirsch index, hg-index, and other metrics. Hg-index, which is based on both

h-index and g-index, characterize the scientific output of researchers and try to keep the advantages of both measures as well as to minimize their disadvantages (Alonso et al., 2010). After 2010, influenced by the rapid progress of the Internet and social networks, scientometrics, which focuses on academic influence evaluation research, has once again undergone a major transition in research, from academic influence assessment in the scientific field to academic influence evaluation all aspects of society. The impact assessment in the scientific field has shifted to the impact assessment in all aspects of society, and alternative measurement studies have emerged as a new way of academic impact evaluation rather than the traditional "citation" evaluation. This brings the likelihood of replacing citation as a method of measuring impact (Mingers & Leydesdorff, 2015). Moreover, journal assessment, knowledge network centrality, and scientometrics law are crucial research topics.

In addition, the literature with a negative value of Δ *M* has a reference value for prediction research of potential influential themes, except for the themes mentioned, library management, research policy, and other themes also have greater influence. For example, text mining of non-English language documents, computer programs for further reform of library management, knowledge flow measurement studies, cross-international cooperation of university institutions, cross-disciplinary cooperation, and other research themes also merit our attention.

4 Discussion and Conclusions

Using the CiteSpace 5.7 software, the three metrics based on the structural variation model were used to estimate and identify potentially influential literature in the field of IS. Based on the analysis framework of "citation network– citation structural variation– citation cluster," the potential research topics were estimated by extracting the subject terms of potentially influential literature. The potential topics estimated by this analysis framework can have a higher confidence level. Based on the prediction results, some of the potentially influential literature has become highly cited literature in the field, such as EGGHE L, 2010(a), which confirms the validity of the structural transformation prediction model.

This study has some limitations. The first is the data selection, the two-step method was used, three journals with high influence in the field of IS were selected as the source journals, and then the journals with high frequency were selected as the database through the method of journal co-citation. This method can effectively reflect the current real-time influential data, but from the selection of the final journals and extracted subject terms, the selected journal topics were all biased toward information measurement research and cannot completely represent the IS research field. Second, it is influenced by some of the review-type literature, and the quality of the cited literature is unevenly distributed from the cited literature predicted by the structural transformation. The literature with high indicator values, some of which are review-type literature, is less helpful in estimating the development trend of the research topic, and the literature with lower indicator values, which are too numerous, have more challenges in identifying the quality. In addition, the top 20% of the literature has less number and, thus, fewer extracted subject terms, which renders the final theme prediction influential. Subsequently, more effective methods would be sought to differentiate the predictions, making the prediction levels richer, and the prediction results clearer for the development trend of marginal topics and the development trend of topics themselves.

Acknowledgments

This research was supported by The National Social Science Fund of China (Grant No. 20BTQ085) and The Soft Science Project of Science and Technology Program of Guangdong Province(CN), 2019B101001024

Addendum:

Literature	★	Existing inter–topic links	New inter–topic links
EGGHE L, 2010 (a)	#4	#4–#4; #1–#4	#1–#4; #4–#4
NORRIS M, 2010 (a)		#1–#4; #4–#4; #7–#4	#4–#4; #7–#4
ALONSO S, 2010	#4	#1–#4; #4–#4; #3–#4	#3–#3; #4–#4; #3–#4
FRANCESCHINI F, 2010		#3–#3; #4–#4; #7–#4	#3–#7; #4–#4; #3–#4; #4–#7
MOUSSA S, 2010		#4–#4; #1–#4	#4–#4
NORRIS M, 2010 (b)		#4–#4; #1–#4	#4–#4; #7–#4
WALTMAN L, 2016	#3	#0–#0; #0–#2; #0–#3; #0–#4; #1–#3; #2–#3; #3–#3; #3–#4; #3–#6; #6–#6	#0–#0; #0–#1; #0–#2; #0–#3; #0–#4; #0–#6; #1–#2; #1–#3; #1–#4; #1–#6; #2–#3; #2–#4; #2–#6; #3–#4; #3–#6; #4–#6
ZHANG L, 2011		#4–#4	#4–#4; #3–#4
AHLGREN P, 2010		#4–#4	#4–#4; #7–#4
BORNMANN L, 2014		#0–#4; #2–#4; #2–#3; #3–#3; #4–#4	#0–#0; #0–#2; #0–#3; #0–#4; #2–#3; #2–#4; #3–#3; #3–#4; #4–#4
NORRIS M, 2010 (a)		#1–#4; #4–#4; #7–#4	#4–#4; #7–#4
ALONSO S, 2010	#4	#1–#4; #4–#4; #3–#4	#3–#3; #4–#4; #3–#4
FRANCESCHINI F, 2010		#3–#3; #4–#4; #7–#4	#3–#7; #4–#4; #3–#4; #4–#7
EGGHE L, 2010 (a)	#4	#4–#4; #1–#4	#1–#4; #4–#4
NORRIS M, 2010 (b)		#4–#4; #1–#4	#4–#4; #7–#4
MOUSSA S, 2010		#4–#4; #1–#4	#4–#4
LAZARIDIS T, 2010		#4–#4	#4–#4; #7–#4
AHLGREN P, 2010		#4–#4	#4–#4; #7–#4
EGGHE L, 2010 (b)		#4–#4	#4–#4
LEYDESDORFF L, 2013		#0–#3; #1–#3; #1–#6; #3–#3; #3–#4	#1–#3; #1–#4; #1–#6; #3–#3; #3–#4; #3–#6; #4–#6
KURTZ MJ, 2010	#2		#1–#1; #1–#3; #3–#3
SCHUBERT A, 2010			#3–#4
LEYDESDORFF L, 2010	#3	#4–#4	#1–#3; #1–#4; #1–#6; #3–#4; #4–#6
OPTHOF T, 2010	#3	#1–#1; #1–#3	#1–#3; #1–#4; #3–#4
VIEIRA PC, 2010			#1–#7
ALONSO S, 2010	#4	#1–#4; #4–#4; #3–#4	#3–#3; #4–#4; #3–#4
SCHMOCH U, 2010			#0–#5
BRESCHI S, 2010			#0–#5
TONTA Y, 2010			#4–#1
PERC M, 2010		#4–#4	#4–#1

References

Alonso, S., Cabrerizo, F. J., Herrera–Viedma, E., & Herrera, F.(2010). Hg–index: A new index to characterize the scientific output of researchers based on the h– and g–indices. *Scientometrics, 82* (2), 391–400.

Biscaro, C., & Giupponi, C. (2014). Co–Authorship and bibliographic coupling network effects on ctations. *Plos One, 9*(6).

Bjarnason, T., & Sigfusdottir, I. D. (2002). Nordic impact: Article productivity and citation patterns in sixteen Nordic sociology departments. *Acta Sociologica, 45* (4), 253–267.

Chen, C.(2012). Predictive effects of structural variation on citation counts. *Journal of the American Society for Information Science and Technology, 63* (3), 431–449.

Egghe, L.(2010). The hirsch Index and related impact measures. *Annual Review of Information Science and Technology, 44*, 65–114.

Kurtz, M. J., & Bollen, J.(2010). Usage Bibliometrics. *Annual Review of Information Science and Technology, 44*, 3–64.

Leydesdorff, L., de Moya–Anegon, F., & Guerrero–Bote, V. P.(2010). Journal maps on the nasis of scopus da–ta: A comparison with the Journal Citation Reports of the ISI. *Journal of the American Society for Information Science and Technology, 61* (2), 352–369.

Mingers, J., & Leydesdorff, L.(2015). A review of theory and practice in scientometrics. *European Journal of Operational Research, 246* (1), 1–19.

Norris, M., & Oppenheim, C.(2010). The h–index: A broad review of a new bibliometric indicator. *Journal of Documentation, 66* (5), 681–705.

Opthof, T., & Leydesdorff, L. (2010). Caveats for the journal and field normalizations in the CWTS ("Leiden") e–valuations of research performance. *Journal of Informetrics, 4* (3), 423–430.

Schubert, A.(2010). A reference–based Hirschian similarity measure for journals. *Scientometrics, 84* (1), 133–147.

Tahamtan, I., Afshar, A. S., & Ahamdzadeh, K.(2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics, 107* (3), 1195–1225.

Vieira, P. C., & Teixeira, A. A. C.(2010). Are finance, management, and marketing autonomous fields of scientif–ic research? An analysis based on journal citations. *Scientometrics, 85* (3), 627–646.