

Do Authors With Stronger Bibliographic Coupling Ties Cite Each Other More Often?

Ali Gazni^{1,2} and Fereshteh Didegah³

¹*ali.gazni@isc.gov.ir*

¹Islamic World Science Citation Center, Shiraz (Iran)

²Regional Information Center for Science and Technology, Shiraz (Iran)

³*fdidgah@gmail.com*

Statistical Cybermetrics Research Group, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY, (UK)

Abstract

Author bibliographic coupling is extended from bibliographic coupling concept and holds the view that two authors with more common references are more related and have more similar research interests. This study aims to examine the association between author bibliographic coupling strength and citation exchange in Information Science & Library Science and more specifically, in imetrics. The results show that there is a positive and significant association between these two factors in Information Science & Library Science and also in imetrics; however, the correlation is more significant among imetricians. This confirms the Merton's norm of universalism versus constructivists' particularism. A closer investigation of bibliographic coupling and citation networks among thirty highly cited imetricians shows that Thelwall, M. is in strong bibliographic coupling and citation relationships with the majority of authors in the network. He and Bar-Ilan have the strongest ABC and citation relationships in the network. Rousseau, R., Glänzel, W., Bornmann, L., Bar-Ilan, J., and Leydesdorff, L. are also in strong ABC relations with each other as well as other authors in the network.

Conference Topic

Citation and co-citation analysis

Introduction

Bibliographic coupling (BC), first introduced by Kessler in 1963, refers to the number of common references between two articles. The more the number of common references between two articles, the more intellectually related they are.

In contrast with co-citation analysis (CA) requiring strength signals (number of citations), BC could help in research fronts detection even with weak signals (Glänzel & Czerwon, 1996). Kuusi and Meyer (2007) claimed that BC has never been used for exploring technology foresight and rare studies used it for research evaluation purposes. However, they used BC for anticipating technological breakthroughs. Yan and Ding (2012) compared different types of networks, including citation based and non-citation based networks at institutional level, and found that BC and AC networks have high similarity and also found that AC has a high similarity with citation networks. Boyack, Börner and Klavans (2009) applied BC to mapping the structure and evolution of research publications in Chemistry. Soó (2014) proposed age-sensitive BC, so if two documents share recent references, they are more related than those sharing older references. Hence, not only the number of common references, but also their age, influences the extent of relatedness between two research works. Van Raan (2005) also reported that intellectual relatedness between two documents could be better obtained through using common

references that are more recent. BC is an effective way for science mapping, research fronts detection and information retrieval (See Jarneving, 2007, 2005; Morris, Yen, Wu, & Tesfaye, 2003; Qiu, 2007). Peters, Braam and van Raan (1995) investigated chemical engineering publications and found that publications with common citations to highly cited papers are more related. White et al. (2004) claim that intellectual ties based on shared references could serve as a better predictor for citations between authors than social ties.

Author bibliographic coupling (ABC), first proposed by Zhao and Strotmann (2008), is extended from BC concept and holds the view that two authors with more common references have more similar research interests. They mentioned that BC is fixed when two articles are published but ABC is constantly evolving over time as the two authors' oeuvre grows. Ma (2012) stated that ABC has an advantage in providing a more comprehensive and concrete map of intellectual structure of the fields and detecting their research fronts in comparison to author co-citation analysis (ACA). The very few studies on ABC did only an author coupling analysis of intellectual structure of few subject fields. For example, using a combination of ACA and author bibliographic coupling analysis (ABCA), Zhao and Strotmann (2014) sought to predict future research trends in information science (IS). They studied research fronts and knowledge bases of IS and also the structural evolution of IS between two 5-year periods (2001-2005 and 2006-2010). They found ABCA an appropriate method to investigate authors' specific research interests in IS and suggested using ACA and ABCA together to better investigate intellectual structure of a subject domain. The same combined method was used in Byun and Chung (2012) to study the research trends of authors in social welfare science; they also suggested using both ACA and ABCA together to investigate traditional and future research trends of a specific domain.

The extent to which two authors are coupled through common references is measured by ABC strength which has different methods to calculate it: Simple, minimum and combined methods (Ma, 2012). Rousseau (2010) also proposed a simple method for calculating the relative ABC by dividing the number of common references between two authors by the total number of their references. Frequency of common references was simply used to measure ABC strength in this study.

No research on the association between ABC strength of two authors and number of citations exchanging between them is found, so this study seeks to examine this relationship in Information Science & Library Science (IS&LS) and more specifically, in imetrics. Therefore it aims to examine the correlation between ABC strength measured by the number of common references between two authors and the number of citations exchanged between them.

Research questions

According to the normative theory of citation, citations are indicators of the cognitive or intellectual influence of a scientific work (Merton, 1973). In a scientific paper, citations can be concept markers (Small, 1978), however, and can transfer knowledge and help with its enlargement (Merton, 1988). As a result, methods like CA have been used for mapping intellectual structure in science (Small, 2004), where BC is used for the same purpose. Hence, common references between pairs of documents, authors, journals or institutions show the extent to which they are related. For instance, two authors who

share a larger number of common references are likely to do research on a narrow area and exchange a high number of citations. Counting citations between two authors with different BC strengths, not only could support Robert K. Merton's norm of universalism versus constructivists' particularism, but also shows any possible difference by the number of common references as a measure of relatedness and types of authors (i.e. highly cited vs. less cited authors).

The theories of citation, normative view vs. social constructivist view, will be examined through answering these questions. The normative theory of citation holds that citations reflect the scientific quality and merits of research outputs because citers use them to reward the works of their colleagues (Small, 2004; White, 2004; MacRoberts & MacRoberts, 1987; Merton, 1973) whereas the social constructivist theory holds that authors use the references to support their own claims and points made. This latter theory emphasises factors affecting citations other than the quality and content of the cited article (White, 2004; Baldi, 1998; Gilbert, 1977).

Given that BC shows relatedness, a positive association between the number of common references and number of citations between two authors will confirm that citations are made for the matter of 'relatedness' and are not perfunctory.

To reach the research goals, this study seeks to answer these questions:

1. Do two authors with a higher number of common references cite each other more often?
2. Is the above association stronger for highly cited authors than other authors?

Methodology

Data collection:

Documents published during 1990-2012 in the journals of Information Science & Library Science (IS&LS) were extracted from Thomson Reuters Web of Science (WoS). This time period is current and consists of a reasonable number of years for investigating the relationship between number of common references and citations exchanged between authors. WoS indexes the mainstream of research and the most prestigious journals in different fields of science; however, a large number of journals in WoS come from a small number of international publishers (Didegah & Gazni, 2011).

Author names disambiguation:

The author names were disambiguated by improving Gazni & Thelwall (2014) method, resulting in 98.2% precision and 92.7% recall. The co-authorship network of authors was used for the improvement. For example, A is a disambiguated author and B is his/her co-author. The papers written by both A and B as co-authors were appended to A's articles. Author names' disambiguation will improve the accuracy of research on author level analysis by distinguishing one name that belongs to several different people and conflating the name variants of a single person.

Calculations:

To make the processing manageable, a random sample of 385 authors with any properties out of all authors who have at least one paper in the journals of IS&LS during 1990-2012

was chosen. The number of common references between these 385 authors and all other authors in the field were counted, where the joint papers were eliminated either for counting the number of common references or for counting the number of citations made and received between each pair of authors. Only citations made and received from the journals in the field were processed for either counting the number of citations between authors or counting the number of common references among them. A list of authors who have at least one common reference with the authors in the sample, and also exchanged citations with them, was created for each author in the sample. For a closer investigation of the association between the number of common references and citations between pairs of authors and also of ABC networks, a sample of highly cited authors in imetrics was taken into account. For this purpose, thirty highly cited imetricians introduced in Abrizah and colleagues (2014) were selected for further analysis. The main reason for taking this sample into account is that these are prolific authors in a specific domain, publishing for a long time and have an excellent knowledge of the domain, its publications and researchers. This is while in the sample of authors from IS&LS, there may be less prolific authors, such as students who publish for a short period of time and then disappear from the research area, and their unfamiliarity with the area will affect their reference and citation behaviours. Therefore, a sample of thirty highly cited imetricians is a consistent sample for showing the association between ABC strength and citation exchange between pairs of authors.

Results

The association between number of common references (BC strength) and number of exchanged citations between pairs of authors in IS&LS

Spearman correlation was tested for the association between the number of common references and the number of citations exchanging between pairs of authors. The results show positive significant correlations between the number of times two authors cited each other and the number of common references between them. The correlation was tested for different groups of pairs of authors with one to 300 common references; it is stronger for the groups of authors with 300 common references than those with a single common reference (Table 1). Therefore, as the number of common references between two authors increases, the number of citations between them also increases. Table 1 shows the increase trend; however, the correlation fluctuated as the number of common references increases but tends to increase. To put it in another way, when the bibliographic coupling strength is stronger between two authors, they tend to cite each other more often. Author bibliographic coupling strength shows how strongly two authors are intellectually related. So, more intellectually related authors cite each other more often. This result confirms the normative theory of citation holding the view that authors cite relevant works, and citations reflect scientific merit and quality.

Table 1. Spearman correlation between ABC strength and number of citations in IS&LS.

No of common refs	Spearman correlation
1	0.31
10	0.36
20	0.35
30	0.38
40	0.37
50	0.37
60	0.39
70	0.38
80	0.36
90	0.39
100	0.4
150	0.46
200	0.47
250	0.58
300	0.61

ABC strength and citation relationship among thirty highly cited authors in imetrics

Thirty highly cited authors in imetrics identified in Abrizah and colleagues (2014) were chosen for a closer investigation of research goals. The main research question on the association between ABC strength and number of exchanged citations was also examined for this group of highly cited authors. Spearman correlation test shows a strong positive association between the number of common references and the number of citations between the authors (Spearman's $\rho = 0.771$, $p\text{-value} < 0.001$), once more confirming the significance of content relevance in citation behavior and normative view of citations.

Moreover, all ABC relations are mapped between each pair of highly cited authors (See Fig. 1). Based on the results, all thirty authors are in BC relationships with all or some of other authors in the network except for Griffith, BC. During 1990-2012, he has published 4 papers in imetrics and has no common references with any of the highly cited authors.

Thelwall, M. is in strong BC relationships with all other authors except with Vanleeuwen, T.N. (only one common reference) and VanRaen, A.F.J. (three common references). He and Bar-Ilan, J have shared the highest common references in the network (4,527 common references) and they have exchanged a large number of citations in the network (118 citations). Thelwall, M. has more than 100 common references with 18 authors in the network. He is also in a strong BC relationship with Vaughan, L. (2,725 common references). Thelwall, M. has also exchanged the highest number of citations in the network with Vaughan, L. (195 citations). He has also strong BC ties with seven others, Leydesdorff, L., Ingwersen, P., Rousseau, R., Cronin, B., Glänzel, W., and Egghe, L., respectively.

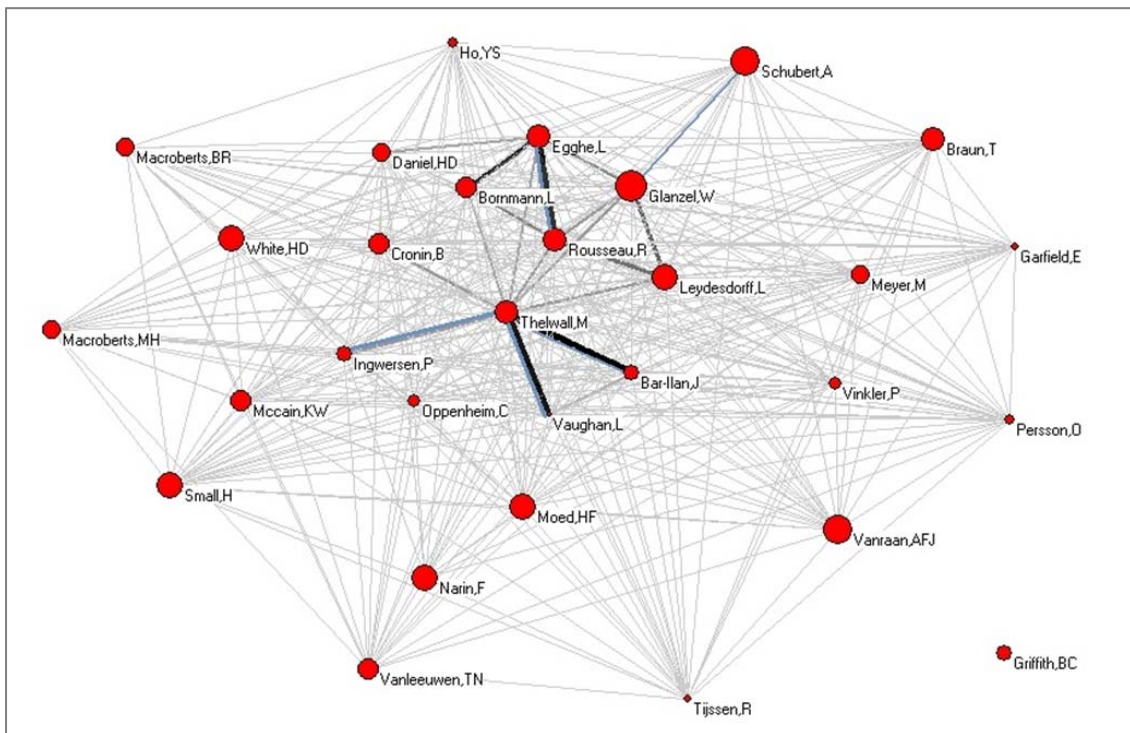


Figure 1. ABC among highly cited authors in imetrics; the black lines show ABC relations and the width of the lines shows ABC strength between pairs of authors; the blue lines show the strongest citation relations in the network and the width of the lines shows the number of citations exchanged between pairs of authors; the size of vertices shows the number of other highly cited authors in the network that each author is in an ABC relation with.

Another strong ABC relationship, and also citation relationship, is seen between Rousseau, R and Egghe, L. (2,270 common references and 175 exchanged citations). Rousseau, R is also in strong BC relationships with other authors in the network. He has strong BC ties with Leydesdorff, L., Bornmann, L., Glänzel, W., and Thelwall, M., respectively.

Glänzel, W., Bornmann, L., Bar-Ilan, J., and Leydesdorff, L. are also in strong BC relationships with other authors in the network. They also have strong citation relationships with each other as well as other highly cited authors.

The correlation between ABC strength and citation exchange in imetrics in comparison with IS&LS

The correlation between the number of common references and the number of citations for top thirty imetricians was examined first amongst themselves and then between them and all other authors in IS&LS with whom they are in BC or citation relationships. As shown in Figure 2, a stronger relationship exists between the authors in the first group than in the second one and regarding the top thirty imetricians, the correlation varies from one author to another one.

For each highly cited imetrician, the proportions of common references with each in-group authors was estimated. Fig. 3 shows that each highly cited author is in a BC relationship with 27 other in-group authors. For example, about 24% of references of

each author are common with one other author. The author distribution of the number of common references with other authors demonstrates a core-scatter shape.

Core references in imetrics

We tried to go further than author couples for common references and identified a number of common references between three and more authors. The thirty highly cited authors in imetrics were examined for this purpose.

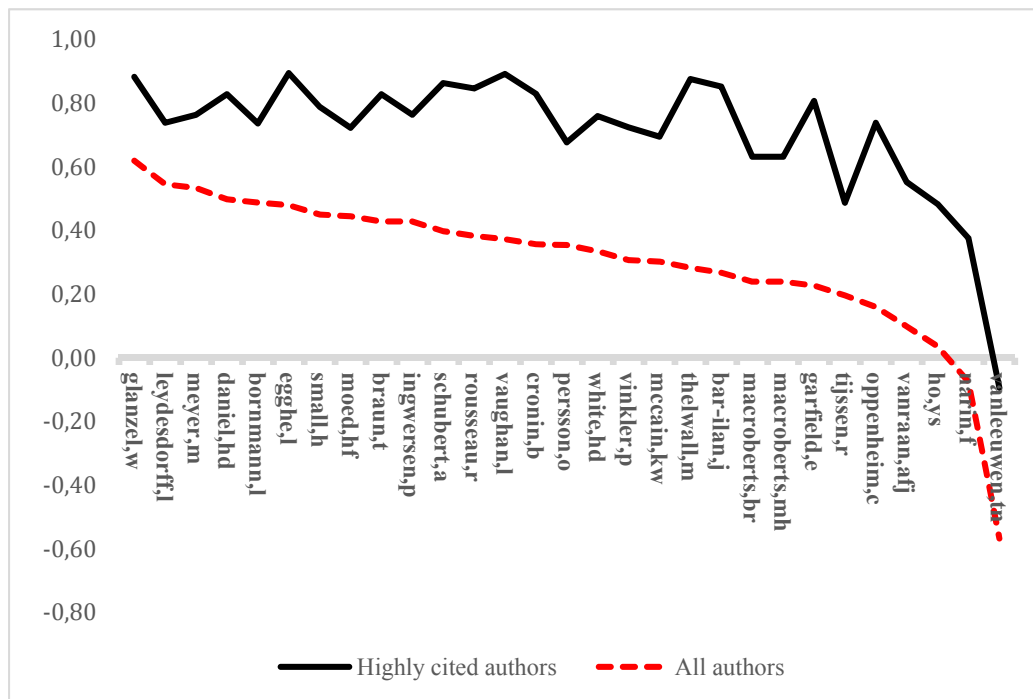


Figure 2. ABC strength and citation correlation between highly cited authors and all authors in IS&LS.

The interesting result is that seventeen highly cited imetricians have one reference in common. The common reference is Hirsch's paper on H-index (Hirsch, J.E. (2005): An index to quantify an individual's scientific research output. *Proceedings of the national academy of sciences of the United States of America*, 102 (46)). Egghe, L., Rousseau, R., and Bornmann, L. have cited this paper more than thirty times in their publications showing that the H-index is one of their common research interests. It is interesting to note that Egghe, L. and Rousseau, R. also have the strongest citation relationship with each other in the network (seventeen5 citations have been exchanged between them) and these two imetricians are also in a strong citation relationship with Bornmann, L. with Bornmann, L. being the fourth top author in citation relationships with both Egghe, L. and Rousseau, R. The strong citation relationships between these authors are mainly due to their similar research interests, one of which is H-index. Twelve highly cited authors have simultaneously five references in common which are listed in Table 2. Eleven authors have nine references and ten authors have eleven references in common.

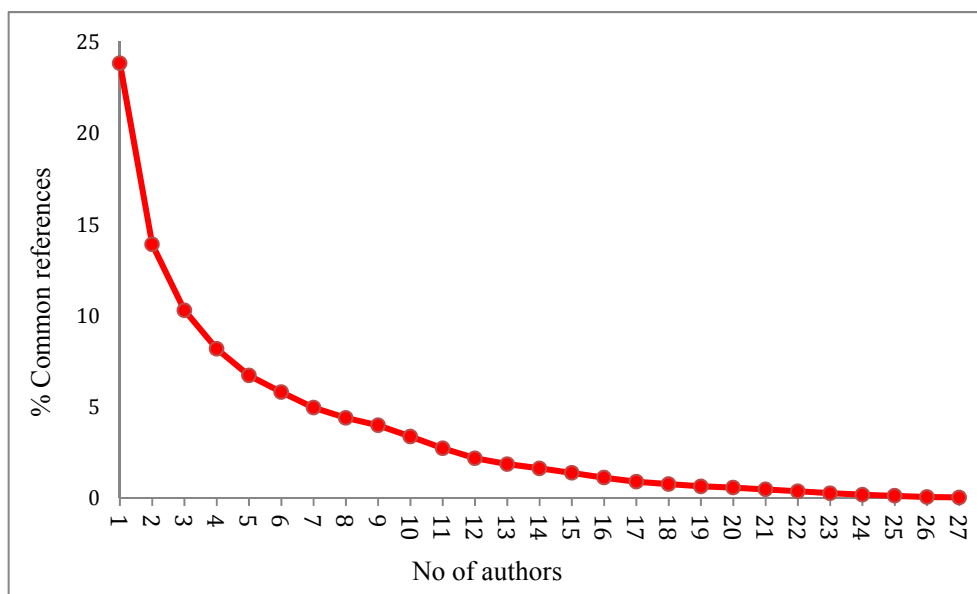


Figure 3. The proportion of common references between each of thirty highly cited imetricians and other in-group authors.

Table 2. Five common references between twelve highly cited imetricians.

VanRaen, A.F.J. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. <i>Scientometrics</i> , 67(3).
Meho, L. & Cronin, B. (2006). Using the h-index to rank influential information scientists. <i>JASIS&T</i> , 57(9).
Glänzel, W., Thijs, B., & Schlemmer, B. (2003). Better late than never? On the chance to become highly cited only beyond the standard bibliometric time horizon. <i>Scientometrics</i> , 58(3).
Macroberts, B.R. & Macroberts, M.H. (1996). Problems of citation analysis. <i>Scientometrics</i> , 36(3).
Moed, H.F., Vanleeuwen, T.N., & Debruin, R.E. (1995). New bibliometrics tools for the assessment of national research performance- database description, overview of indicators and first applications. <i>Scientometrics</i> , 33(3).

Discussion and conclusion

This study examined the association between author bibliographic coupling strength and the number of times authors cited each other. The results of the study on authors in IS&LS showed that there is a positive and significant correlation between ABC and exchanged citations between two linked authors confirming that authors are citing related authors and relevant research works in their field (Table 1). This finding opposes the social constructivist view holding that authors cite others for some other reasons than relevance or rewarding the cited author, but it confirms the normative theory of citations. A group of thirty highly cited authors in imetrics were also examined for this purpose. The result of the association between ABC and the number of citations shows a positive strong correlation between ABC and exchanged citations between imetricians. Therefore, highly cited authors in imetrics are in strong BC relationships with whom they also have strong citation relationships.

The number of common references between pairs of authors was accepted as a measure of relatedness between them. Therefore relatively, the higher number of common

references between two authors, especially in a long-term period, could show the extent to which they are working in similar research areas; however, authors may change their research interests over time due to changes in the research fields. The higher number of citations between two authors with higher number of common references, when they are not co-authors, could probably show that they cite each other since they may work on similar research areas and also for the matter of relevancy.

ABC relations between the thirty highly cited imetricians were examined and mapped and strong relationships were determined. Thelwall, M. and Bar-Ilan, have the strongest ABC relationship in the network; they are also in a strong citation relationship. Rousseau, R., Glänzel, W., Bornmann, L., Bar-Ilan, J., and Leydesdorff, L. are also in strong ABC relations with each other as well as other authors in the network. In an investigation of the number of common references in groups of two and more imetricians, smaller groups have a larger number of references in common while larger groups have fewer numbers of common references. For example, seventeen imetricians have only one reference in common while some two-author groups have more than a thousand common references. The latter groups presumably work on narrow research areas. Larger groups with fewer number of common references suggest membership in a wider research area. The results show that a maximum of seventeen authors have one reference on H-index in common. Authors citing this single paper are also in strong citation relationship with each other.

Comparing the correlation between number of common references and number of exchanged citations for highly cited imetricians and all authors in IS&LS related to Fig. 2 shows that number of common references between imetricians increases the probability of higher citations between them more than that of IS&LS. Moreover, ABC relationship or common references with each single author may result in different number of citations with him/her.

Intuitively, considering the core-scatter distribution of citations to papers in the science network, an author probably has common references with a large number of other authors, while he/she probably has more common references with a fewer number of other authors (Fig. 3). The author presumably has more related research interests with the latter group of authors where some of them may belong to the same research community.

The number of common references and citations between pairs of authors could be also influenced by the number of papers published by the authors. For example, two authors may have five common references whilst the first author only published a single paper during his/her entire research life and the second one published more than twenty papers. The first author will have fewer common references with any other authors in the field than the second author and he/she will have less opportunity to cite other authors due to his/her short research life. So authors' research lifetime in the science network (e.g. newcomers, students, faculty members and professional researchers) does matter. Authors with a longer research life have more chances to know other researchers in similar research fields and they also have extra opportunities to focus on more specific and narrow research topics, compared to authors with a shorter research lifetime. Hence, a stronger association between the number of common references and citations exchanged between authors is found for the former group.

Science network and its attributes are continuously changing over time and a research specialty may appear or disappears after a while; authors may also change their research interests during their research lifetime. In the current study, a longer time span is used to

show that clustering authors, based on more recent common references, may be replaced by a shorter one, which could result in a stronger relationship between the bibliographic coupling network and the citation network. According to the results of current studies, authors with a longer research lifetime and more citations demonstrate a stronger relationship between their number of common references and citations. However, even weak ties in bibliographic coupling networks could also be used for research front detection purposes. Bibliographic coupling is not enough for mapping intellectual structure of science and measuring relatedness by itself. Thus, as with previous studies, it is better to be combined with other methods, such as co-citations, to realise better results.

References

- Abrizah, A., Erfanmanesh, M., Rohani, V. A., Thelwall, M., Levitt, J. M., & Didegah, F. (2014). Sixty-four years of informetrics research: productivity, impact and collaboration. *Scientometrics*, 101(1), 569-585.
- Baldi, S. (1998). Normative versus social constructivist processes in the allocation of citations: a network-analytic model. *American Sociological Review*, 63, 829-46.
- Boyack, K. W., Börner, K., & Klavans, R. (2009). Mapping the structure and evolution of chemistry research. *Scientometrics*, 79(1), 45-60.
- Byun, J-H & Chung, E-K (2011). Domain Analysis on Electrical Engineering in Korea by Author Bibliographic Coupling Analysis. *Journal of Information Management*, 42(4), 75-94.
- Didegah, F., & Gazni, A. (2011). The extent of concentration in journal publishing. *Learned Publishing*, 24(4), thirty3-310.
- Gazni, A. & Thelwall, M. (2014). The long-term influence of collaboration on citation patterns. *Research Evaluation*, doi: 10.1093/reseval/rvu014.
- Gilbert, G. N. (1977). Referencing as persuasion. *Social Studies of Science*, 7, 113-122.
- Glänzel, W., & Czerwon, H.J. (1996). A new methodological approach to bibliographic coupling and its application to the national, regional, and institutional level. *Scientometrics*, 37, 195-221.
- Jarneving, B. (2005). Acomparision of two bibliometric methods for mapping of the research front. *Scientometrics*, 65, 245-263.
- Jarneving, B. (2007). Bibliographic coupling and its application to research front and other core documents. *Journal of Informetrics*, 1, 287-307.
- Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American documentation*, 14(1), 10-25.
- Kuusi, O., & Meyer, M. (2007). Anticipating technological breakthroughs: Using bibliographic coupling to explore the nanotubes paradigm. *Scientometrics*, 70(3), 759-777.
- Ma, R. (2012). Author bibliographic coupling analysis: A test based on a Chinese academic database. *Journal of Informetrics*, 6(4), 532-542.
- MacRoberts, M. H. & MacRoberts, B. R. (1987). Another test of the normative theory of citing. *Journal of the American Society for Information Science*, 38, 305-306.
- Merton, R. K. (1973). The normative structure of science. *The sociology of science: Theoretical and empirical investigations*, 267.
- Merton, R. K. (1988). The Matthew effect in science II: cumulative advantage and the symbolism of intellectual property. *Isis*, 79, 606-23.
- Morris, S. A., Yen, G., Wu, Z., & Tesfaye, B. (2003). Time line visualization of research fronts. *Journal of the American Society for Information Science and Technology*, 54(5), 413-422.
- Peters, H. P., Braam, R. R., & van Raan, A. F. (1995). Cognitive resemblance and citation relations in chemical engineering publications. *Journal of the American Society for Information Science*, 46(1), 9-21.
- Qiu, J. P. (2007). *Methods of citation analysis*. In Informetrics. Wuhan: Wuhan University Press.
- Rousseau, R. (2010). Bibliographic coupling and co-citation as dual notions. In L. Birger (Ed.), *A Festschrift in honour of Peter Ingwersen, special volume of the e-zine of the ISSI, June 2010*, 173-183.
- Small, H. (1978). Cited documents as concept symbols. *Social Studies of Science*, 8(3), 327-340.

- Small, H. (2004). On the shoulders of Robert Merton: towards a normative theory of citation. *Scientometrics*, 60, 71-79.
- Soós, S. (2014). Age-sensitive bibliographic coupling reflecting the history of science: The case of the Species Problem. *Scientometrics*, 98(1), 23-51.
- Van Raan, A. F. (2005). Reference-based publication networks with episodic memories. *Scientometrics*, 63(3), 549-566.
- Yan, E., & Ding, Y. (2012). Scholarly network similarities: How bibliographic coupling networks, citation networks, co-citation networks, topical networks, co-authorship networks, and co-word networks relate to each other. *Journal of the American Society for Information Science and Technology*, 63(7), 1313-1326.
- White, H. D. (2004). Reward, persuasion, and the Sokal Hoax: a study in citation identities. *Scientometrics*, 60(1), 93-120.
- Zhao, D., & Strotmann, A. (2008). Evolution of research activities and intellectual influences in information science 1996–2005: Introducing author bibliographic-coupling analysis. *Journal of the American Society for Information Science and Technology*, 59(13), 2070–2086.
- Zhao, D., & Strotmann, A. (2014). The knowledge base and research front of information science 2006–2010: An author co-citation and bibliographic coupling analysis. *Journal of the Association for Information Science and Technology*, 65(5), 995-1006.