Journal of Informetrics 11 (2017) 407–422

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Regular article

The impact of collaboration and knowledge networks on citations

Jiancheng Guan[a,b,*,1], Yan Yan[a,c,1], Jing Jing Zhang[a,1]

[a] School of Economics & Management, University of Chinese Academy of Sciences, 100190 Beijing, PR China
[b] School of Business Administration, South China University of Technology, 510640 Guangzhou, PR China
[c] Innovation Management and Strategy, University of Groningen, 9747 AE Groningen, The Netherlands

ARTICLE INFO

Article history:
Received 20 May 2016
Received in revised form 6 February 2017
Accepted 19 February 2017
Available online 2 March 2017

Keywords:
Paper citations
Knowledge network
Collaboration network
Keywords
Wind energy

ABSTRACT

Research papers not only involve author collaboration networks but also relate to knowledge networks. Previous research claims that a paper's citations are related to the node attributes of its authors in collaboration networks. We further propose that a paper's citations can also be affected by the node attributes of its knowledge elements in knowledge networks. In this study, we develop a new method to construct the knowledge network using article keywords. Further, we explore the antecedents of paper citations from both the collaboration and knowledge network perspectives. Using wind energy paper data (16,351 records) collected from WoS (Web of Science) and JCR (Journal Citation Reports) database, we construct two distinct networks and empirically examine the hypotheses of the relationships between node attributes of two networks and the paper's citations, which fill the gap in prior studies and will inspire related studies. We have the following key findings: in the collaboration network, the structural holes of authors have positive but non-significant effects on the paper's citations, while the authors' centrality has inverted U effects on the paper's citations; in the knowledge network, the structural holes of knowledge elements are positively related to the paper's citations, and the knowledge elements centrality has an inverted U relationship with the paper's citations.

1. Introduction

The research impact of a paper is the extent to which it is useful to other researchers (Rousseau, García-Zorita, & Sanz-Casado, 2013; Garfield, 1979; Leydesdorff and Bornmann, 2011). To assess the research impact of papers, forward citations (named citations below) count is usually used (Garfield, 1972; Leydesdorff and Opthof, 2010). Citations received by a paper means the number of times its content (method, ideas, and so on) is formally used in subsequent papers (Lozano, Larivière, & Gingras, 2012; Uzzi, Mukherjee, Stringer, & Jones, 2013). Previous research has found a large variance of citations received by papers, as some papers are highly cited thousands of times, while nearly 20% papers have never been cited at all (Mingers and Burrell, 2006; Redner, 1998). This may raise the following question: What factors affect paper citations? To answer the question, our research will focus on the publication level rather than the researcher or institute level.

* Corresponding author at: School of Economics & Management, University of Chinese Academy of Sciences, Beijing, 100190, PR China.
E-mail address: guanjianch@ucas.ac.cn (J. Guan).

[1] The authors' names are alphabetically ordered and they contribute equally to this paper.

http://dx.doi.org/10.1016/j.joi.2017.02.007
1751-1577/

There are many reasons why papers are cited by researchers. Previous research combined different frameworks, views, and methods to study the influence factors of publication citations. For instance, Bornmann and Daniel (2008) claimed that researchers are motivated to cite a paper not only to acknowledge academic importance of this paper, but also for some non-academic reasons. Tahamtan, Safipour Afshar, and Ahamdzadeh (2016) summarized all the factors related to the number of paper citations: paper related factors (e.g. abstract), journal related factors (e.g. impact factor) and authors related factors (e.g. author number). In addition, some scholars began to realize the importance of social-based networks in influencing paper citations (Abbasi and Jaafari, 2013).

Collaboration network, as a typical social-based network in research, has received much attention (McFadyen and Cannella, 2004; Guan, Zuo, Chen, & Yam, 2016). The positions of authors in the collaboration network play important roles in taking advantage of diverse resources, thereby significantly influencing their publication citations (Abbasi, Hossain, & Leydesdorff, 2012; Newman, 2004; Li, Liao, & Yen, 2013). For example, Li et al. (2013) examined the positive effects of authors' centrality on their citation count. Abbasi, Altmann, and Hossain (2011) found that scholars' degree centrality and structural holes positively affect the citation-based performance. Based on previous research, we argue that the citation of a paper is affected by its authors in the collaboration network. Recently, Abbasi and Jaafari (2013) examined the positive effects of the geographically diverse collaboration on citations at the publication level.

However, we argue that a paper involves not only the collaboration network (social-based) but also the knowledge network (knowledge-based). A knowledge network is comprised of combinations between components or elements of scientific or technological knowledge. Knowledge elements indicate the dimensions and categories of a knowledge area. For instance, patents are usually categorized into several classes to express their technological features (Guan and Liu, 2016). Thus, a patent's technological classes have been widely accepted as valid proxies for knowledge elements (Carnabuci & Operti, 2013). Similarly, scientific papers usually include multiple keywords to indicate their knowledge elements (Muñoz-Leiva, Viedma-del-Jesús, Sánchez-Fernández, & López-Herrera, 2012; Cobo, López-Herrera, Herrera-Viedma, & Herrera, 2011; Su and Lee, 2010). According to the extant bibliometric literature, publication keywords can be considered as knowledge elements. For example, related studies used keywords in depicting knowledge structure maps (Su and Lee, 2010; Assefa and Rorissa, 2013; Yang, Han, Wolfram, & Zhao, 2016), knowledge hotspot detections (Chen, 2006) and research trend analyses (Xie, Zhang, & Ho, 2008). Following these previous approaches, in this study, we use keywords of a paper to indicate its knowledge elements. We then apply a co-keyword network to represent the knowledge network. Knowledge elements are linked through their co-occurrence in previous publications. Over time, these elements are interwoven into a network that records their combinatorial or co-occurrence histories. Knowledge elements can be replenished in use, combined or recombined in the processes of generating innovation (Garud & Kumaraswamy, 2005). In such innovative processes, knowledge elements form associative or combinational relationships with each other, leading to the formation and development of knowledge networks. In this study, we propose that the structural attributes of knowledge elements in knowledge networks will influence the knowledge elements' combination opportunities and efficiency. For example, a central knowledge element is more likely to be searched and novelly combined with other elements because it has more contents and experiences of element couplings. As such, the node attributes of knowledge elements may affect the citations of the paper involving these elements. We want to fill the gap left by the lack of studies that have built the knowledge network and investigated the influence of this network on paper citations.

Our research is at the paper level rather than researcher or institute level, as we consider the average of authors or keywords' measures for each publication. This study focuses on how the network attributes of authors and knowledge elements influence a publication's citation. This study has several contributions: (1) we develop a new method to construct the knowledge network using article keywords, which fill the gap left by prior studies and will inspire further related studies; (2) this study highlights the importance of knowledge and collaboration networks in citations. Specifically, we firstly consider node attributes (e.g., degree centrality and structural holes) of authors in collaboration networks and knowledge elements in knowledge networks as influence factors of paper citations; (3) we focus on the paper level citations. Most studies aggregated citations to the author level, organizational level or journal level. We argue that the same author sometimes has distinct citations in different papers, thus the fine-grained analysis of a paper's citations is needed.

Our empirical analysis focuses on wind energy research. Wind energy research has been experiencing unprecedented growth and is now worldwide booming. Moreover, as technologies develop, the wind energy field involves more and more researchers and yields a wealth of distinct knowledge elements. Thus, this field is a typical example of a high velocity field with easily observed collaboration and knowledge networks. Based on the retrieval strategy of wind energy, we collected a sample of 16,351 papers published from 2002 to 2015.

An objective of our research is to explore how collaboration and knowledge networks affect a paper's citations. We applied ordinary least squares (OLS) model with robust standard errors to test the above-mentioned relationships. The remainder of our paper is arranged as follows: Section 2 introduces our hypotheses. Section 3 describes our data collection and variable measurement, followed by the descriptions of wind energy research, descriptions of networks and empirical tests in Section 4. Conclusions are given in Section 5.

2. Theory and hypotheses

Different node attributes indicate different chances of accessing new information that is important in creating high quality research. In this study, we mainly study two node attributes-degree centrality and structural holes-in two kinds of networks. We choose degree centrality and structural holes for three reasons. First, innovation is full of risk and uncertainty and consequently individuals prefer to conduct local search. Thus, we focus on the local (ego-level) metrics of networks in this study. Degree centrality and structural holes are widely used local (ego-level) metrics. Second, other node attributes, such as closeness and betweenness centrality, are global level metrics. Analyzing the joint effects of multiple centralities simultaneously in an estimation model may result in the suppressor effect, which may lead to the reverse beta coefficients (Abbasi et al., 2011; Li et al., 2013). Third, degree centrality and structural holes in collaboration networks and knowledge networks have specific meanings. Herein, we reason that, within the collaboration network, the node attribute of an author indicates the degree of prestige among his egocentric research collaborators and information or knowledge access (Tsai, 2001). Structural holes of an author exist in the network where the focal author is tied to other collaborators which are not connected themselves (Burt, 2009; Zhang, Yan, & Guan, 2015). Structural holes, as a key node attribute, can capture the non-redundant and efficient access to information. Degree centrality of an author indicates the number of collaborators with which a focal author has linked (Guan, Yan, & Zhang, 2015; Otte and Rousseau, 2002). The higher an author's degree centrality, the more partners the author has. Authors tend to differ in terms of the node attributes within the collaboration network, and papers usually differ in terms of the author list. Accordingly, papers vary in terms of the average node attributes of their authors within the collaboration network.

We further propose that, within the knowledge network, the structural attributes of a knowledge element illustrate its combinatorial opportunities and knowledge flow advantages. Structural holes of a knowledge element mean that the degree of disconnectedness among other elements is linked to the focal element; this indicates its information control advantages and the extent to which combinatorial opportunities were unexploited in the focal element's ego-network. Degree centrality of a knowledge element refers to the volume of linkages directed to the focal element in the knowledge network, which illustrates the combinatorial potential of the element with other elements. Knowledge elements differ in terms of node attributes within the knowledge network, and papers usually involve different knowledge elements within their portfolios. Consequently, papers vary in terms of the average node attributes of their knowledge elements within the knowledge network.

2.1. The collaboration network

Here, the collaboration network refers to the scientist co-authorship network, where nodes are scientists and ties mean cooperation in prior papers. An author spans structural holes if the author connects with those authors who are not connected with one another (Burt, 1997; Ahuja, 2000a). Structural holes illustrate the degree to which a focal author's partners are connected with each other. A structural hole is considered to be a lack of linkage between any pair of nodes in the network (Abbasi, Wigand, & Hossain, 2014). For instance, author A is bridging the structural holes among authors B, C and D when author A has three direct partners (authors B, C and D), who are not tied to each other. That is, author A occupies several gaps between authors B, C and D (from author A's perspective). On the contrary, author A lacks structural holes if author A has three direct ties to authors B, C and D, who are all linked to each other. Spanning structural holes can bring authors the maximum unique but diverse information from ties. Moreover, spanning structural holes illustrates non-redundancy among ties (Guan and Liu, 2016; Zaheer and Soda, 2009). An author occupying rich structural holes can receive more unique and heterogeneous information spillovers than an author lacking structural holes. Therefore, spanning structural holes can be more efficient for an author to obtain, sift and store information spillovers. As an author spans more structural holes, he attains more non-redundant information and control benefits (Rodan and Galunic, 2004). Some critical information might flow through the co-authorship ties. The author situated in structural holes is likely to gain potential opportunities for controlling information across unconnected authors (Burt, 2009). He can receive unique information from these non-redundant relationships. Thus, information control economically benefits the author regarding the amount of ties needed to access new information, which may save time and energy to improve research. Meanwhile, control benefits make the author act as intermediary between disconnected researchers (Rowley, Behrens, & Krackhardt, 2000), who rely on the author to facilitate information exchange flow across the collaboration network. The key information received may make the scholar's work more attractive; thereby positively affecting his research impact. Thus,

Hypothesis 1a. For a paper, its authors' average structural holes in the collaboration network are positively related to its citation count.

If an author occupies a central location in a collaboration network, the author has many ties with others and is likely to access desired information resources (Ahuja, 2000b; Abbasi et al., 2012). These resources can fuel the author's research by providing the external information and new ideas. Meanwhile, the central authors can communicate with more authors in the collaboration network than anyone else (Ahuja, Galletta, & Carley, 2003; Phelps, Heidl, & Wadhwa, 2012). Therefore, they promote their own knowledge more efficiently. However, authors with centrality exceeding a certain degree will fast access a large amount of relevant information, which is out of bounded rationality (Paruchuri, 2010). Excess information will overload and overwhelm these authors' research ability, eventually resulting in their knowledge quality reduction. Furthermore, such central authors promote their knowledge inefficiently because of information overload. Therefore, the centrality of authors will negatively affect citations beyond a certain centrality threshold.

Hypothesis 1b. For a paper, its authors' average centrality in the collaboration network has inverted U effects on its citation count.

2.2. The knowledge network

In the knowledge network, nodes denote knowledge elements (indicated by keywords in this paper) and ties mean co-occurrence of different elements. If two knowledge elements are jointly used in previous research, they are called "combined knowledge elements", otherwise "uncombined knowledge elements". A knowledge element spans the structural holes if it is combined with those knowledge elements that are not combined (Burt, 2009). The uncombined knowledge elements are linked to the same knowledge element. Thus they may be related and have more combinatorial opportunities (Wang, Rodan, Fruin, & Xu, 2014; Alavi and Tiwana, 2002). The knowledge search is often local and associative (Cyert and March, 1963). If a knowledge element spans rich structural holes, it will be non-redundant for researchers to search via it and find combinatorial opportunities between uncombined knowledge elements (Guan and Liu, 2016), thus, the papers involving this element will receive more citations. Meanwhile, structural holes in the knowledge network have advantages over knowledge flow by controlling the spread of its own information as researchers who study these uncombined elements are likely to use information in this element (Burt, 2009). On the contrary, if the related elements are often combined in previous papers, the creative combination of them may be nearly exhausted and the control advantages of structural holes disappear. Therefore, elements occupying rich structural holes in knowledge networks can provide combinatorial opportunities and the advantage of controlling knowledge flow and then lead to a higher citation count of the papers involved.

Hypothesis 2a. For a paper, its knowledge elements' average structural holes in the knowledge network are positively related to its citation count.

In a knowledge network, the centrality of a knowledge element illustrates the element's combinatorial opportunities with other elements. Knowledge elements are interdependent and constitute a larger knowledge system (Hansen, 1999; Guan and Liu, 2016). As a knowledge element's centrality increases, its combinatorial opportunities tend to rise; there are two main reasons for this. Firstly, the increasing centrality of a knowledge element illustrates a signal of its feasible and desirable combinations, which enhance other authors' beliefs in combining this knowledge element with others (Kuhn, 2012). More researchers will allocate attention to investigating combinatorial chances around this knowledge element. Secondly, a knowledge element with increasing centrality can provide a better understanding for researchers in combining this element with others (Fleming, 2001; Wang et al., 2014). Therefore, the increasing centrality of knowledge elements in a paper can enhance its citations. However, after a point of centrality in the knowledge network, the knowledge element's combinatorial opportunities will decrease. To be specific, after the amount of combinations, the knowledge element has exhausted combination values (e.g., scientific or technological) (Fleming, 2001) and cannot be perceived as fruitful. Thus, citations to the papers with high knowledge centrality tend to decrease.

Hypothesis 2b. For a paper, its knowledge elements' average centrality in the knowledge network has inverted U effects on its citation count.

Fig. 1 presents our research framework. As we can see from this Figure, we propose that a paper involves several authors (e.g., 1, 2, 3 and 4) in the collaboration network and knowledge elements (e.g., a, b, c, d and e) in the knowledge network. This study explores how node attributes of authors and knowledge elements in two networks affect the paper's qualities (i.e., citations).

3. Methods

3.1. Data collection

In this study, the data used was acquired from WoS (Web of Science) and JCR (Journal Citation Reports) databases. To retrieve wind energy paper records in WoS, we adopted the search terms "TS = ('wind power' OR "wind turbine*" OR "wind energy*" OR "wind farm*" OR "wind generation" OR "wind systems") and PY = (2002–2015)" used in previous research (Guan et al., 2015a; Sanz-Casado, Garcia-Zorita, Serrano-López, Larsen, & Ingwersen, 2013). Search term "TS" means using all search terms in four fields of a paper: title, abstract, author keyword and keywords plus. Finally, we collected 16,351 papers published from 2002 to 2015, including 15,611 journal articles, 7 book chapters and 733 conference papers. Each paper record from WoS included its journal, title, authors, abstract, keywords, citations, publication year and other detailed information. After extracting the list of journals that all the papers were published in, we downloaded these journals' information from JCR in each year from 2002 to 2015, and then added impact factors (5 years) of all journals annually in our dataset. For the missing journals in JCR, we manually collected their IF (impact factors) from their official websites, or supplemented IF using statistics of citations per document of each journal in each year from SJR (SCImago Journal & Country Rank). SJR provides citations per document for journals in each year, which is equivalent to journal impact factors in Thomson Reuters. As SJR only provides 2, 3 and 4 years statistics, we supplemented IF using citations per document of each journal (4 years) in each year from SJR. Totally, there are only 28 journals (4.01% of journals in our sample) that cannot be found in JCR but can be found in SJR. In the analysis part, we did robust checks without these data, which showed that the results are consistent.

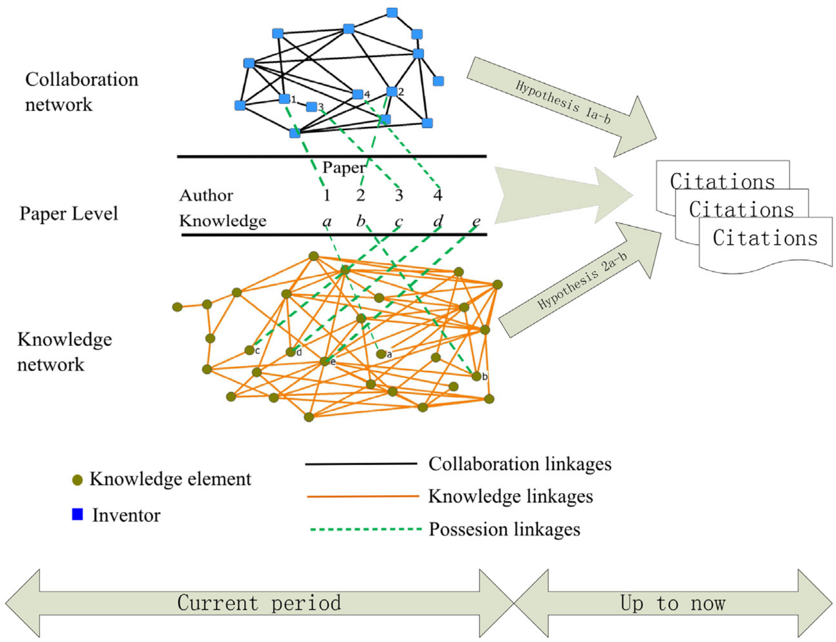

Fig. 1. Research framework.

3.2. Measurement

We used a longitudinal design in this study. To be specific, we created knowledge and collaboration networks based on moving five-year windows (i.e., 2002–2006, 2003–2007,. . .2006–2010), resulting in five snapshots of both networks. For example, when we analyzed an article published in 2006, we measured the network structures of its authors and knowledge elements in 2002–2006 and normalized citations it received until 2015. This approach was consistent with previous research (Wang, 2016).

3.2.1. Dependent variable

The dependent variable is the normalized citations of each paper in our sample. To calculate it, we first collected the total citations received by each paper. Our analysis sample stops at 2010, although we collected wind energy papers published until 2015. Articles take an average of five years to be steadily cited (Van Raan, 2006; Wang, 2016). Therefore, we may not have enough time periods to calculate the citations of the papers published after 2010. Following previous research (Cannella and McFadyen, 2013), we standardized the citation numbers using the mean and standard deviation of all citations received by the papers published in the same year. This method removes the positive citation bias in older papers (Cannella and McFadyen, 2013). If a paper is published in the year *t*, its normalized citations will be calculated as follows:

$$Normalizedcitations_{it} = \frac{citations_i - citationsmean_{allpapersint}}{citationss\tan darddeviation_{allpapersint}} \quad (1)$$

3.2.2. Independent variables

Our independent variables are centrality and structural holes in the knowledge and collaboration network. To measure them, we took the following three steps.

3.2.2.1. First step: construction of the knowledge and collaboration networks. We developed a new method using article keywords to construct the knowledge network. In this article, we used both keywords given by authors and suggested by WoS (i.e., so-called KeyWords Plus) as proxies for knowledge elements (Cobo et al., 2011; Muñoz-Leiva et al., 2012). We used two keywords in this study for two reasons: some journals in our sample, such as Plos One and Wind Engineering, do not have author-provided keywords; the other reason is that KeyWords Plus are extracted from the article's references, and may include important terms not listed among the author keywords (Garfield, 1990; Thomson, 2005). We considered the existence of a tie between two elements if they co-occur in a same paper. Secondly, we notice there are subject categories (n = 251) and research direction classifications (n = 156) given in WoS's full record. However, these classes are based on journals not on the content of articles. In contrast, keywords are considered as the basic elements of knowledge structure as they are anchored on subject matters in articles (Su and Lee, 2010). We used Science of Science (Sci2) software for data cleaning, which was specifically designed by Team (2009) for scientometric research. For example, we used Sci2 to detect similar keywords in 90% level and merged duplicate keywords manually, including the plural and singular forms and different expression forms, such as "Doubly Fed Induction Generators" and "Doubly Fed Induction Generator", "Oxide Fuel-cell" and "Oxide Fuel-cells". We also merged some acronyms with the full keywords, like "photovoltaic" and "PV". We constructed knowledge networks in Sci2 software.

The collaboration network has authors as nodes and coauthor relationship as ties (Guan and Yan 2016). Specifically, we considered there is a collaboration connection between two researchers if they coauthor papers together. To recognize author names, we utilized their full name and affiliations to capture each unique author in the collaborative network. Specifically, two authors who have same names and affiliations are considered same. We used Sci2 software to construct collaboration networks in this study.

3.2.2.2. Second step: Measurements of centrality and structural holes. This study has two sets of independent variables: knowledge network measures and collaboration network measures. Next, we explain the constructions of knowledge network measures and then describe collaboration network measures. This study used Pajek software to calculate these network measures.

3.2.2.2.1. Centrality in a knowledge network. Centrality captures the extent to which a particular knowledge element occupies central positions. Three different centrality measures are often used: degree, closeness, and betweenness, which have different meanings. According to previous studies, degree centrality is most widely used to measure a node's involvement in a local network (Opsahl, Agneessens, & Skvoretz, 2010; Otte and Rousseau, 2002), which is consistent with our purpose of indicating an element's combination with others. Thus, we chose degree centrality to indicate the amount of direct ties which a knowledge element has with other knowledge elements. To eliminate the size effect of networks, this study used normalized degree centrality proposed by Freeman (1978), which was used in previous research (Abbasi et al., 2011; Wang et al., 2014).

$$Centrality_i = \frac{\sum_{j=1}^{g} a(i,j)}{g-1} \tag{2}$$

where $a(i, j)$ is a binary variable, indicating if i connects with j or not. $a(i, j) = 1$ if i and j are connected, 0 otherwise. g denotes the number of knowledge elements in the knowledge network. For example, if i connects with five knowledge elements and there are 2981 knowledge elements in a network, the centrality of knowledge element i is 5/(2981-1) = 0.0017.

3.2.2.2.2. Structural holes in a knowledge network. A knowledge element occupies structural holes in a knowledge network if the focal knowledge element will tie to others who are not connected amongst themselves. We first adopted Burt's constraint measure to calculate network constraint C_i, which captures how strongly i can be constrained by its neighbors (Burt, 2009, 2004; Forti, Franzoni, & Sobrero, 2013). Secondly, based on previous research (Wang et al., 2014; Guan, Zhang, & Yan, 2015), we subtracted the aggregate constraint measure C_i from 2 to indicate i's control advantages created by spanning structural holes.

$$Structural \quad holes_i = 2 - C_i = 2 - \sum_j \Big(p_{ij} + \sum_{k, k \neq i, k \neq j} p_{ik} p_{kj}\Big)^2 \tag{3}$$

where i is the focal element and p_{ij} denotes the proportion at which an element j accounts for node i's contacts. For example, if i connects with j and five other elements, then p_{ij} is 1/6. K: the third element connected with both i and j. If element i ties to more elements, i has a lower p_{ij} and p_{ik}, and is thereby less constraint. Meanwhile, if k has more ties with other elements, k has a lower p_{kj}, thereby decreasing the constraint on i.

3.2.2.2.3. Centrality in a collaboration network. After constructing a collaboration network among authors, we measured the normalized degree centrality of authors in the collaboration network, which is the simplest and most straightforward method to quantify author centrality. Degree centrality can illustrate how well connected the author is in his local network. According to Eq. (2), we calculated this variable using the number of the focal author's collaborators and all authors in a network.

3.2.2.2.4. Structural holes in a collaboration network. Structural holes refer to the disconnections between a focal author's collaborators. To calculate the extent of spanning structural holes in a collaboration network, we also used the two steps mentioned above. After building a collaboration network, we first calculated the network constraint of every author in the network. This was followed by, according to Eq. (3), subtracting the constraint measure from 2.

3.2.2.3. Third step: aggregating all measures into the paper level. Since our unit of analysis is a paper and since a paper usually has multiple authors and multiple knowledge elements, we averaged the centrality and structural holes values of all of that paper's authors and elements as its value in these independent variables. For example, if a paper has three authors, whose structural holes are 1.21, 1.13, 1.14, separately. Then the structural holes value of this paper will be (1.21 + 1.13 + 1.14)/3 = 1.16, signifying the extent of network control advantages of the authors in this paper. We used CN to indicate the collaboration network and KN to denote the knowledge network.

Table 1
Definitions of control variables.

		Description	Theoretical base
Author(s) related factors	Author numbers	The number of authors in the focal paper was controlled.	More authors may have more opportunities to get higher citations (Batista, Campiteli, & Kinouchi, 2006).
	Funding	This variable means the number of funding records.	Financial supports can affect the quality of articles (Boyack and Börner, 2003)
	Institutions	The number of institutions that authors in the focal paper belongs to.	The number of institutions in a paper can improve number of citations (Bjarnason & Sigfusdottir, 2002).
	International	This variable is 1 if authors in the focal paper come from different countries, is 0 if they come from the same country.	International cooperation has significant impact on a paper's citations (Inzelt, Schubert, & Schubert, 2009).
Paper related factors	Abstract length	The abstract length refers to the number of sentences in the abstract. A sentence is identified when it stops as a space: ".".	Characteristics of a paper's abstract are related to its citation (Didegah and Thelwall, 2013; Letchford, Preis, & Moat, 2016)
	Acknowledgements	The number of words in the acknowledgement part in the focal paper.	Research peer supports can affect the quality of articles (Wang and Shapira, 2011).
Journal related factors	WoS categories	Every journal in WoS is assigned with one or more subject categories. This variable indicates the number of Web of Science categories of the journal the paper published on.	A paper's citations are associated with its subjects or field (Skilton, 2006).
	Journal IF	Journal impact factor (5 years) was included to control journal effects.	Journal influence is an important factor of paper citation performance (Bensman, 2008)

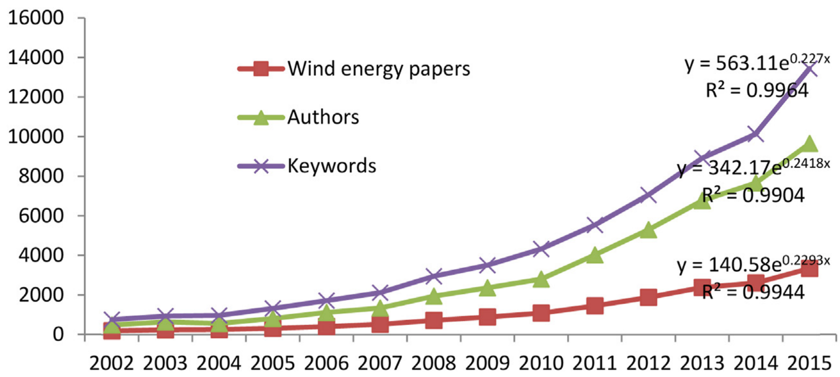

Fig. 2. Papers, authors and keywords in the wind energy field per year during 2002–2015.

3.2.3. Control variables

In the statistical analysis of all hypotheses proposed, eight variables were controlled, which included *abstract length, acknowledgements, journal IF (impact factor), WoS categories, authors number, funding, Institution numbers, International collaboration.* Table 1 summarizes definitions of control variables.

4. Analysis and results

4.1. Description of the wind energy field

Wind energy is considered a promising energy because it is environmental friendly, omnipresent and freely available. Wind energy field, one of the most crucial renewable energy fields, is a multidisciplinary area both within this field and with other fields. To be specific, within the wind energy field, it yields a wealth of distinct research topics or knowledge elements (63,784 in our analysis period). With other fields, wind energy, as a sub-field of renewable energy, is usually connected to other fields (Guan et al., 2015a). The range and depth of the research subject can generate a great deal of data about interdisciplinary structures (Tijssen, 1992). Thus, the wind energy field is an appropriate context to test the effects of knowledge networks. Interdisciplinary is much invoked in the energy fields (Owens & Driffill, 2008). As shown in Fig. 2, the number of wind energy papers, authors and keywords has, typically, exponential growth patterns of an emerging scientific field. Thus, wind energy research can be considered as an emerging scientific field.

Figs. 3 and 4 depict the main component of the knowledge and collaboration network (2011–2015), separately. The main components are composed of knowledge elements connected, such that any element can be reached from others by walking a way of intermediate elements (Schilling and Phelps, 2007). Both of the knowledge and collaboration networks have several clusters. The results show that knowledge and collaboration networks differ in the degree of cohesion and density.

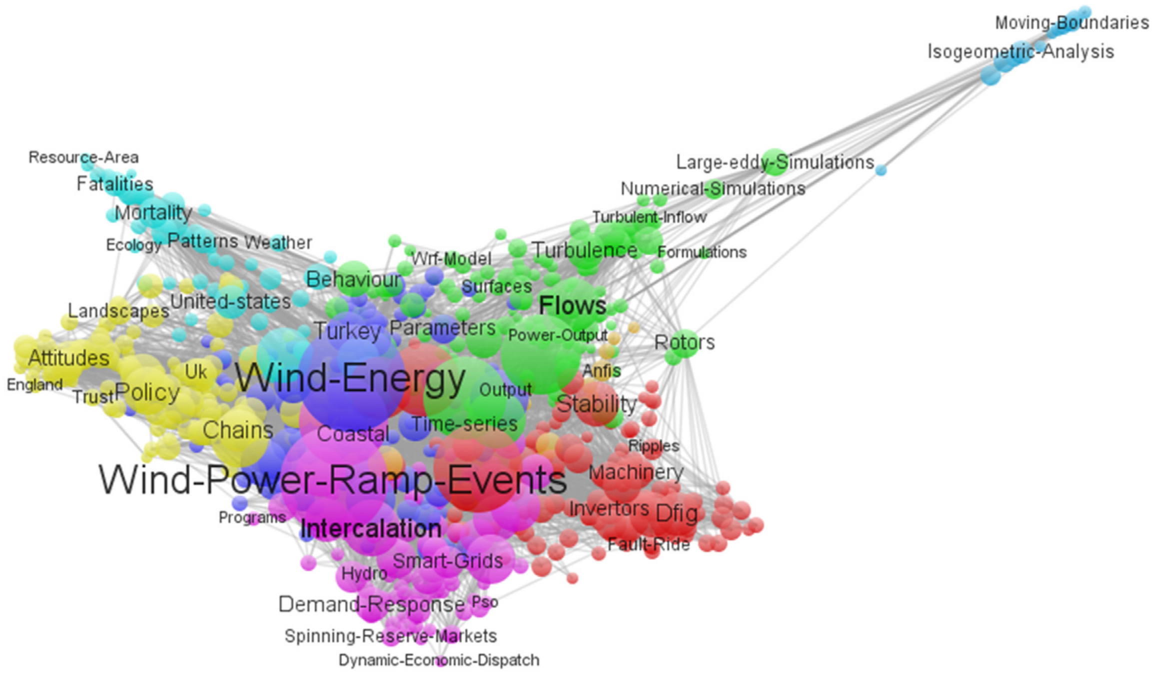

Fig. 3. The main component of knowledge network (2011–2015).
Note: The figure only displays the elements whose degree is more than 80. According to the relatedness, the elements are clustered into several clusters. Different clusters are indicated by different colors. The bigger the circle is, the more frequency it has. VOSviewer is used for this visualization (Van Eck and Waltman, 2010; Leydesdorff et al., 2013).

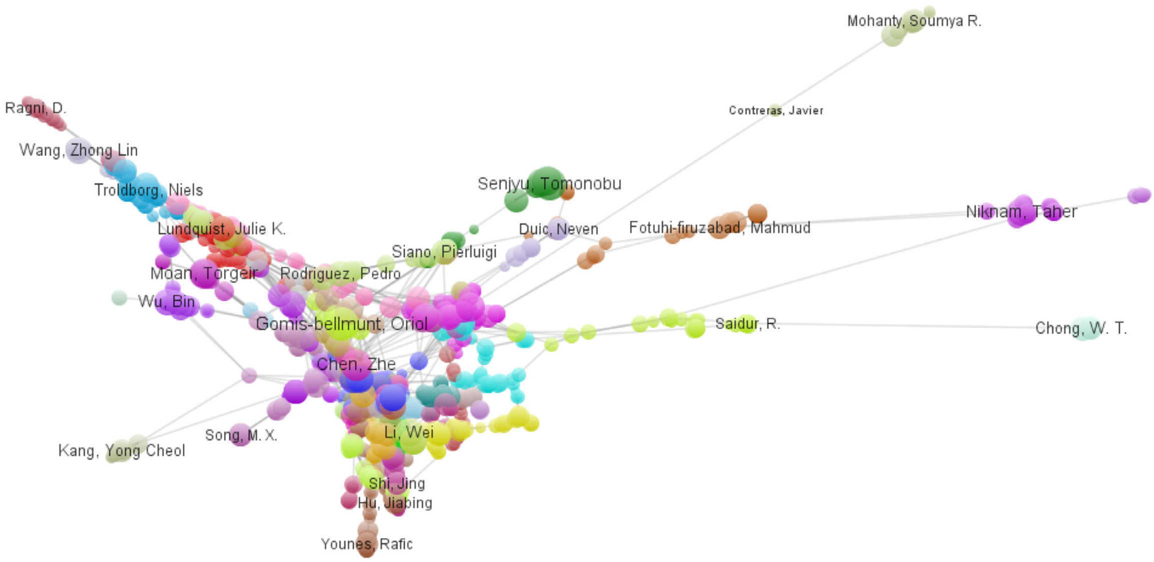

Fig. 4. The main component of collaboration network (2011–2015).
Note: The nodes represent authors. The bigger the node is, the more publications it has (VOSviewer is used for this visualization).

The knowledge network is characterized by a high degree of cohesion and density, which provides more opportunities for knowledge search and recombination. Conversely, the collaboration network is sparse and fragmented, resulting in fewer chances for knowledge search and joint problem solving.

4.2. Empirical analysis and results

Table 2 shows the key variables' descriptive statistics and bivariate correlations. OLS regression has several classical assumptions, such as no multicollinearity, no autocorrelation and homoscedasticity. Multicollinearity indicates high corre-

Table 2
Correlations and descriptive statistics.

Variables	Mean	SD	VIF	1	2	3	4	5	6	7	8	9	10	11	12
1 Normalized citations	0	1	–	1											
2 Author numbers	3.02	1.72	2.05	0.11**	1										
3 Journal IF	2.11	1.83	1.06	0.39**	0.13**	1									
4 Abstract length	8.12	4.14	1.04	0.04**	0.05**	.07**	1								
5 WoS categories	1.76	0.87	1.01	0.03***	−0.04**	0.02	0.02	1							
6 Funding	0.46	1.04	2.17	0.05**	0.14**	0.06**	0.05**	.03***	1						
7 Acknowledgements	77.4	171.5	2.20	0.04*	0.14**	0.05**	0.10**	0.03	0.43**	1					
8 Institutions	1.61	1.01	2.93	0.07**	0.52**	0.13**	0.05**	−0.01	0.21**	0.24**	1				
9 International	0.39	0.49	2.41	0.03*	0.34**	0.05**	0.03**	0.001	0.21**	0.22**	0.65**	1			
10 CN centrality	0.002	0.001	3.28	0.13**	0.57**	0.05**	0.004	−0.04*	−0.01	.09**	0.18**	−0.01	1		
11 CN structural holes	1.22	0.26	2.80	0.14**	0.47**	0.06**	−0.02	−0.03***	0.09**	.20**	.26**	.08**	0.48**	1	
12 KN centrality	0.02	0.02	1.25	0.04**	−0.07**	0.09**	0.003	0.02	−0.002	0.05**	−0.01	−0.02	−0.04**	0.03	1
13 KN structural holes	1.71	0.48	1.31	0.14**	0.09**	0.16**	0.15**	.06**	0.09**	0.12*	.09**	.09**	0.08**	0.12**	0.42**

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.1$.

Table 3
OLS regression results (with robust standard errors).

Models	1	2	3	4
Variables	Dependent variable: Normalized citations			
Author numbers	0.032* (0.014)	−0.017 (0.017)	0.032* (0.014)	−0.015 (0.017)
Journal IF	0.220** (0.034)	0.222** (0.034)	0.219** (0.035)	0.221** (0.035)
Abstract length	0.001 (0.004)	0.002 (0.004)	0.001 (0.004)	0.002 (0.004)
WoS categories	0.019 (0.022)	0.022 (0.022)	0.021 (0.022)	0.023 (0.022)
Funding	0.03 (0.028)	0.031 (0.027)	0.028 (0.028)	0.03 (0.027)
Acknowledgements	−0.00006 (.00009)	−0.00004 (.00009)	−0.00010 (.00009)	−0.00007 (0.00009)
Institutions	−0.015 (0.053)	−0.005 (0.052)	−0.015 (0.052)	−0.004 (0.052)
International	−0.003 (0.085)	−0.006 (0.085)	−0.022 (0.086)	−0.021 (0.085)
CN centrality		213.679** (81.054)		228.926** (79.674)
CN centrality^2		−30247.33* (15085.11)		−32395.73* (14808.21)
CN structural holes		0.192*** (0.107)		0.123 (0.104)
KN centrality			0.846 (2.068)	0.606 (2.055)
KN centrality^2			−53.279** (21.584)	−47.381* (21.385)
KN structural holes			0.625** (0.116)	0.58** (0.114)
cons	−0.575** (0.113)	−0.831** (0.147)	−1.682** (0.217)	−1.786** (0.245)
F value	6.56**	8.63**	27.39**	22.86**
Adjust R squared	0.1547	0.1670	0.1673	0.1776

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.1$

lations among the independent variables, which may result in adverse results on estimated coefficients. In Table 2, all VIF (variance inflation factor) values are far less than 5, meaning there is no significant multicollinearity. Autocorrelation refers to the error terms' cross-correlation in contiguous time and usually arises in using time series data, which is not a concern in this research. Finally, homoscedasticity means that random error terms have the same variance for all predictor variables. The violation of this assumption (i.e. heteroscedasticity) may result in the overestimation of goodness of fit in regression. To avoid this issue, we used white general test for heteroscedasticity. We implemented the post-estimation command "imtest, white" after OLS regression in Stata. Results ($chi2(88) = 395.70$, $p < 0.001$) indicate that heteroskedasticity is detected. To ensure the accuracy of our hypothesis testing, we used OLS regression with robust standard errors, which can correct the problem of heteroscedasticity (Wooldridge 2015).

We tested four hypotheses by entering all control variables in model 1, independent variables separately in model 2 and 3, and all variables in the full model 4. Table 3 presents the testing results for the number of normalized citations. As shown in models 2 and 4, the coefficient for CN structural holes is not stably significant, indicating that Hypothesis 1a is not fully supported. We use r to indicate coefficient below. The coefficient for CN centrality square is negative and significant ($r = -30679.17$, $p < 0.05$ in model 2; $r = -33072.97$, $p < 0.05$ in model 4), indicating that CN centrality has an inverted U relationship with normalized citations. Hence, Hypothesis 2a is supported. As shown in models 3 and 4, the coefficient for KN structural holes is positive and significant ($r = 0.612$, $p < 0.001$ in model 3; $r = 0.573$, $p < 0.001$ in model 4), indicating that Hypothesis 2a is supported. Models 3 and 4 indicate support for Hypothesis 2b, where the normalized citations of a paper are curvilinearly related to the level of KN centrality ($r = -55.279$, $p < 0.01$ in model 3; $r = -48.539$, $p < 0.01$ in model 4). ΔR^2 between model 2 and model 4 is 0.016 ($p < 0.001$). In other words, adding the paper's knowledge network centrality and structural holes significantly improved the model. The paper's centrality and structural holes in the knowledge network are

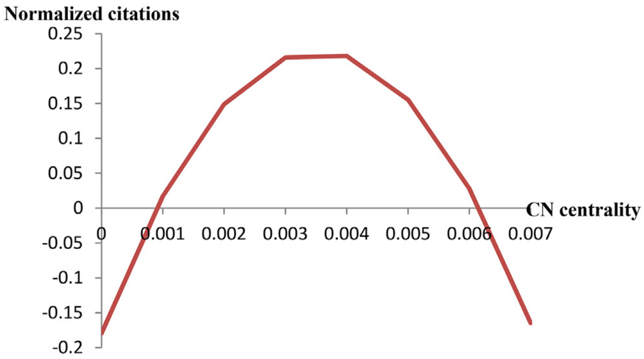

Fig. 5. The relationship between CN centrality and normalized citations.

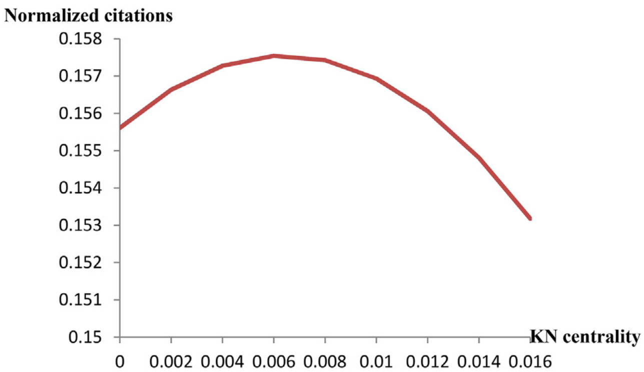

Fig. 6. The relationship between KN centrality and normalized citations.

important, influential factors that can explain 1.6% of the variance in the normalized citations of papers. Further, the Journal Impact Factors of JCR and Scopus are calculated differently and are based on different databases. Thus, to test the robustness of our results, we only reserved the journals which have JCR impact factors and re-estimate the models. The results are consistent with prior results in Table 3.

Beyond statistical effects of node attributes of two networks, their practical significance also needs emphasis. The following step was to estimate the practical significance of our results based on Model 4 in Table 3. Following the approaches of previous studies (Datta, Guthrie, Wright, 2005; Huselid, 1995), we estimated the practical significance of four node attributes of two networks by calculating the influence on a paper's unnormalized citations of increasing a variable by one standard deviation. For CN structural holes, we held all other variables at their mean values. We found that one-standard-deviation increase of CN structural holes is related to a 1.5 increase in the unnormalized citations (please see calculation procedures in Appendix A), which represents a 5.4% increase relative to the mean of citations per paper (28.5). Likewise, we found that one-standard-deviation increase of KN structural holes is related to a 13.2 increase in the unnormalized citations, which represents a 46.3% increase relative to the mean of citations per paper (28.5). Thus, KN structural holes have bigger effect size on paper's citation than CN structural holes.

However, for CN/KN centrality, the effect size of them depends on the value of CN/KN centrality. At first, the effect of CN/KN centrality on paper's citation is positive, and the effect size decreases to zero when CN/KN centrality increases to a threshold value. After that, the effect of CN/KN centrality on the paper's citation becomes negative, and their effect size increases as CN/KN centrality increases. For example, from the regression model (Model 4 in Table 3), when all variables are held at their mean values, the paper with a CN centrality score one standard deviation above its mean is related to a 2.7 decrease in the unnormalized citations, which represents a 9.5% decrease relative to the mean of citations per paper (28.5). The paper with KN centrality scores one standard deviation above its mean is related to a 1.9 decrease in the unnormalized citations, which represents a 6.7% decrease relative to the mean of citations per paper (28.5). Based on the regression results in Table 3, we set all other variables to their average values and then show an inverse U-shape relationship between CN/KN centrality and normalized citations in Figs. 5 and 6. The plots show that as the level of CN/KN centrality increase from low to moderate, normalized citations increase and as the level of CN/KN centrality increases from moderate to high, normalized citations decrease.

We drew on Wales et al. (2013) to assess the validity of an inverted-U relation between CN (or KN) centrality and citations (see Table 4). First, we applied the joint significance tests to the estimated coefficients of centrality and centrality squared (CN: 4.56, $p < 0.01$; KN: 15.01, $p < 0.01$). Then, we conducted Sasabuchi tests (H0: Monotone or U shape), indicating that

Table 4
Test of an inversely U-shaped relationship.

Dependent variable: Citations	Full model (CN centrality)	Full model (KN centrality)
Test of joint significance of the monomial and quadratic terms (F value)	4.56**	15.01**
Sasabuchi test for inverse U shape(t value)	4.85**	2.13*
Extremum point	0.0036	0.0071
95% confidence interval, Fieller method	[0.0024, 0.0148]	[−0.2304, 0.0281]
95% confidence interval, Delta method	[0.0020,0.0052]	[−0.0290, 0.0432]

+ $p < 0.1$.
* $p < 0.05$.
** $p < 0.01$.

our results are consistent with the presence of an inverted U relationship (CN: 4.85, $p < 0.001$; KN: 2.13, $p < 0.05$). Thirdly, we calculated the Fieller and Delta confidence interval and found that extremum points are included in these confidence intervals. As we can see, the thresholds of CN and KN centrality in the inverted-U relations are 0.0036 and 0.0071, separately. Therefore, the significant inverted-U relationship between CN (or KN) centrality and citations is observed.

We performed additional sensitivity analyses to assess the robustness of our findings. Firstly, the typical distribution of citations tends to be highly skewed and lognormal. As previous research suggested (Thelwall and Wilson, 2014), we used the natural logarithm of one plus citation numbers and then used OLS model with robust standard errors for regression (please see Table A1 in Appendix A). Results from this analysis were consistent with those showed in Table 3. Secondly, due to the over-dispersion in citation numbers, the negative binomial (NB) regression model is appropriate for informetric datasets (Ajiferuke and Famoye, 2015). In this study, we estimated all the models with NB regression, using robust standard errors (please see Table A2 in Appendix A). The results were consistent with those reported in Table 3. Overall, the results of various sensitivity analyses provided extra support for our findings.

5. Conclusion and discussion

Research papers involve not only the collaboration network but also the knowledge network. Previous research claimed that a paper's impact (i.e., citations) is related to the collaboration network characteristics of its authors. We propose that a paper's citations relate to the knowledge network positions of its knowledge elements. This study begins with a construction of knowledge networks using article keywords, and measures each article's knowledge network attributes. Further, we examine the relationships between collaboration networks (or knowledge networks) and paper citations, using the wind energy paper data (2002–2015) from WoS and JCR databases. In empirical parts, we use White test for heteroskedasticity, OLS regression with robust standard and Sasabuchi inverted-Utests. Our analysis results generate the following findings.

First, we found that knowledge and collaboration networks differ in the extent of cohesion and integration. Knowledge network is more cohesive and extensive. In addition, based on OLS regression results, we found that adding the node attributes of a paper' authors in collaboration networks and its knowledge elements in knowledge networks can significantly improve the model fits. Thus, a paper's centrality and structural holes in collaboration networks and knowledge networks are important, influential factors. However, we found that several control variables are not statistically significant in our regression results. It is possible that these factors such as the abstract length, WoS categories, funding, acknowledgements, institution numbers and international collaborations are weakly related to a paper's citation, while journal impact factor, node attributes of collaboration networks and knowledge networks are more strongly related to a paper's citation.

Second, based on the regression and Sasabuchi test results on the paper level, we found the structural holes of authors in the collaboration network have no significant effects on paper citation count, and the centrality of authors in the collaboration network has inverted U effects on the paper citation count, which filled the gap in previous linear centrality-citation relationship research. Specifically, previous research found a positive relationship between author centrality and his research impact, ignoring the negative influence of author centrality. Meanwhile, previous research focused on the individual level, while our research firstly explored this relationship at the paper level.

Third, we firstly confirmed that the structural holes of knowledge elements are positively related to paper citations. These findings illustrated elements occupying structural holes in the knowledge network which can provide combinatorial opportunities and knowledge flow control advantages and, in turn, lead to higher citation counts of involved papers. The knowledge elements centrality has an inverted U relationship with paper citations. A paper's knowledge elements centrality positively correlates with the citations of this paper. However, after a certain degree, because of lower combination values and opportunities, knowledge elements centrality is negatively related to the citation.

This study has several methodological and theoretical contributions. Firstly, we developed a new method in constructing knowledge networks using author keywords and WoS keywords at the paper level. According to previous research, a delineation of the network needs to be chosen (Krackhardt, 1990; Tsai and Ghoshal, 1998). This study defined the collaboration network and knowledge network in the wind energy field. Our method has much generalisability in the other fields only if they are clearly and rigorously defined. Future research should have a clear network boundary to avoid the overestimation of network characteristics (e.g. structural holes).

Secondly, this study highlights the importance of knowledge and collaboration networks in citations. We explored the antecedents of paper citations from both collaboration and knowledge network perspectives, which filled the gap from prior studies and will inspire related studies. We studied how author collaboration and knowledge network attributes are related to citations at the paper level. Finally, this study calls attention to the knowledge network research. The same node attribute has different meanings and influence in the collaboration and knowledge networks. We built a theory about node attributes in the knowledge network via this study by positing potentially important predictors of paper citations. Using network analysis methods to explore the structure features of knowledge elements provides research bases for the knowledge-based search, which may help researchers to develop research directions using a knowledge network perspective.

Our study also has some limitations. First, we use Sci2 tool to mitigate the different authors with the same name issue. However, name disambiguation problem can still exist because authors who have multi-affiliations or moved from one institute to another during the period of our study are not handled. Name disambiguation is a big challenge and cannot be completely eliminated (Li et al., 2014). Future work can combine other databases and provide the algorithm to uniquely identify authors. Second, control variables have a large influence on the results of the models. Some other variables, such as the age of the authors, are not controlled. Thus, we treat all authors as if they are equally experienced researchers. Future work can consider the age of the authors to better explain the citations of the focal paper. In addition, from the sticky information perspectives (Von Hippel, 1994), structural holes in knowledge networks could lead to fewer citations. It will be interesting to relate with centrality of a key word in knowledge networks. Future research can test the effects of interaction between centrality and structural holes in knowledge networks. Third; although we define our network boundary; there may be an overestimation problem of network characteristics (e.g. structural holes). For example; we cannot capture some connections between authors (or keywords) in other fields. Future research could select a more broad research field to avoid this issue.

Acknowledgments

This study is supported by the Grants from National Natural Science Foundation of China (Nos.71673261 and 71373254) and from The Research Team of Natural Science Foundation of Guangdong Province in China (2016A030312005). The authors are very grateful for the valuable comments and suggestions from two anonymous reviewers and Prof. Editor-in-Chief Ludo Waltman, which significantly improved the quality of the paper.

Appendix A.

The calculation of practical significance

Let's take the calculation of CN structural holes' practical significance as an example. Suppose the change of unnormalized citations is

$$\Delta Citations = Citations' - Citations,$$

Since

$$\frac{Citations' - Averagecitations}{StandardDeviation_{citations}} = \beta * CN'_{structralholes} + \alpha, \tag{1}$$

$$\frac{Citations - Averagecitations}{StandardDeviation_{citations}} = \beta * CN_{structralholes} + \alpha, \tag{2}$$

then (1)-(2):

$\frac{\Delta Citations}{StandardDeviation_{citations}} = \beta * \Delta CN_{structralholes}.$

Thus, $\Delta Citations = \beta^* \Delta CN_{structralholes}{}^* Standard\ Deviation_{citations} = 0.123 * 0.26 * 47.52 \approx 1.5.$

Results of sensitivity analysis

See Tables A1 and A2.

Table A1
OLS regression results (with logged citations and robust standard errors).

Models Variables	1 Dependent variable: Logged citations	2	3	4
Author numbers	0.035*	−0.064***	0.038**	−0.057**
	(0.015)	(0.02)	(0.014)	(0.019)
Journal IF	0.265***	0.269***	0.261***	0.265***
	(0.036)	(0.035)	(0.032)	(0.032)
Abstract length	0.006	0.007	0.007	0.007****
	(0.005)	(0.005)	(0.005)	(0.005)
WoS categories	0.096***	0.101***	0.099***	0.103***
	(0.023)	(0.022)	(0.022)	(0.022)
Funding	−0.004	0.004	−0.009	0.001
	(0.022)	(0.021)	(0.023)	(0.022)
Acknowledgements	0.00003	0.00009	0.00003	0.00003
	(0.00013)	(0.00012)	(0.00013)	(0.00012)
Institutions	−0.062	−0.033	−0.061	−0.029
	(0.041)	(0.042)	(0.04)	(0.04)
International	−0.051	−0.047	−0.098	−0.089
	(0.073)	(0.071)	(0.071)	(0.069)
CN centrality		723.166***		772.279***
		(115.372)		(111.686)
CN centrality^2		−119138.7***		−125553.2***
		(27985.36)		(27230.77)
CN structural holes		0.12		0.305*
		(0.13)		(0.127)
KN centrality			4.306****	4.233****
			(2.621)	(2.566)
KN centrality^2			−117.794***	−109.595***
			(31.494)	(30.675)
KN structural holes			1.332***	1.291***
			(0.132)	(0.132)
cons	1.979***	1.926***	−0.443****	−0.243
	(0.107)	(0.151)	(0.233)	(0.25)
F value	13.58***	20.59***	37.66***	39.74***
R squared	0.1632	0.1943	0.2059	0.2336

The figures in parentheses are robust standard errors.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.
**** $p < 0.1$.

Table A2
Negative binomial regression results (with robust standard errors).

Models Variables	1 Dependent variable: Citations	2	3	4
Author numbers	0.065***	−0.038*	0.061***	−0.04*
	(0.019)	(0.019)	(0.018)	(0.018)
Journal IF	0.376***	0.372***	0.358***	0.355***
	(0.029)	(0.028)	(0.029)	(0.028)
Abstract length	−0.002	0.002	−0.003	0.001
	(0.005)	(0.005)	(0.005)	(0.005)
WoS categories	0.057*	0.065*	0.07**	0.077**
	(0.027)	(0.027)	(0.027)	(0.026)
Funding	−0.009	0.008	−0.004	0.015
	(0.025)	(0.026)	(0.025)	(0.027)
Acknowledgements	0.00009	0.00003	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Institutions	−0.014	0.013	−0.017	0.014
	(0.04)	(0.04)	(0.039)	(0.038)
International	−0.171*	−0.127****	−0.2**	−0.152*
	(0.073)	(0.075)	(0.075)	(0.076)
CN centrality		695.751***		760.864***
		(110.415)		(110.542)

Table A2 (*Continued*)

Models Variables	1 Dependent variable: Citations	2	3	4
CN centrality^2		−88124.85***		−97074.96***
		(20413.02)		(20589.05)
CN structural holes		0.293*		0.463***
		(0.15)		(0.146)
KN centrality			3.15	3.591
			(3.208)	(3.09)
KN centrality^2			−117.823***	−115.451***
			(35.333)	(34.207)
KN structural holes			1.165***	1.154***
			(0.175)	(0.168)
cons	2.271***	2.372***	0.199	0.471
	(0.112)	(0.172)	(0.321)	(0.331)
Wald chi2	189.66***	257.11***	314.01***	390.11***
Log likelihood	−14616.707	−14548.111	−14553.893	−14484.337

The figures in parentheses are robust standard errors.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.
**** $p < 0.1$.

References

Abbasi, A., & Jaafari, A. (2013). Research impact and scholars' geographical diversity. *Journal of Informetrics*, *7*(3), 683–692.
Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, *5*(4), 594–607.
Abbasi, A., Hossain, L., & Leydesdorff, L. (2012). Betweenness centrality as a driver of preferential attachment in the evolution of research collaboration networks. *Journal of Informetrics*, *6*(3), 403–412.
Abbasi, A., Wigand, R. T., & Hossain, L. (2014). Measuring social capital through network analysis and its influence on individual performance? *Library & Information Science Research*, *36*(1), 66–73.
Ahuja, M. K., Galletta, D. F., & Carley, K. M. (2003). Individual centrality and performance in virtual R&D groups: An empirical study? *Management Science*, *49*(1), 21–38.
Ahuja, G. (2000a). Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly*, *45*(3), 425–455.
Ahuja, G. (2000b). The duality of collaboration: Inducements and opportunities in the formation of interfirm linkages. *Strategic Management Journal*, *21*(3), 317–343.
Ajiferuke, I., & Famoye, F. (2015). Modelling count response variables in informetric studies: Comparison among count linear, and lognormal regression models. *Journal of Informetrics*, *9*(3), 499–513.
Alavi, M., & Tiwana, A. (2002). Knowledge integration in virtual teams: The potential role of KMS. *Journal of the American Society for Information Science and Technology*, *53*(12), 1029–1037.
Assefa, S. G., & Rorissa, A. (2013). A bibliometric mapping of the structure of STEM education using co-word analysis? *Journal of the American Society for Information Science and Technology*, *64*(12), 2513–2536.
Batista, P. D., Campiteli, M. G., & Kinouchi, O. (2006). Is it possible to compare researchers with different scientific interests? *Scientometrics*, *68*(1), 179–189.
Bensman, S. J. (2008). Distributional differences of the impact factor in the sciences versus the social sciences: An analysis of the probabilistic structure of the 2005 journal citation reports. *Journal of the American Society for Information Science and Technology*, *59*(9), 1366–1382.
Bjarnason, T., & Sigfusdottir, I. D. (2002). Nordic impact: Article productivity and citation patterns in sixteen Nordic Sociology departments? *Acta Sociologica*, *45*(4), 253–267.
Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, *64*(1), 45–80.
Boyack, K. W., & Börner, K. (2003). Indicator-assisted evaluation and funding of research: Visualizing the influence of grants on the number and citation counts of research papers. *Journal of the American Society for Information Science and Technology*, *54*(5), 447–461.
Burt, R. S. (1997). The contingent value of social capital. *Administrative Science Quarterly*, *42*(2), 339–365.
Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, *110*(2), 349–399.
Burt, R. S. (2009). *Structural holes: The social structure of competition*. Cambridge: Harvard University Press.
Cannella, A. A., & McFadyen, M. A. (2013). Changing the exchange the dynamics of knowledge worker ego networks. *Journal of Management*, *0149206313511114*.
Carnabuci, G., & Operti, E. (2013). Where do firms' recombinant capabilities come from? Intraorganizational networks, knowledge, and firms' ability to innovate through technological recombination. *Strategic Management Journal*, *34*(13), 1591–1613.
Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature? *Journal of the American Society for Information Science and Technology*, *57*(3), 359–377.
Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field. *Journal of Informetrics*, *5*(1), 146–166.
Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. New Jersey: Prentice-Hall Inc.
Datta, D. K., Guthrie, J. P., & Wright, P. M. (2005). Human resource management and labor productivity: Does industry matter? *Academy of Management Journal*, *48*(1), 135–145.
Didegah, F., & Thelwall, M. (2013). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of Informetrics*, *7*(4), 861–873.
Fleming, L. (2001). Recombinant uncertainty in technological search? *Management Science*, *47*(1), 117–132.
Forti, E., Franzoni, C., & Sobrero, M. (2013). Bridges or isolates? Investigating the social networks of academic inventors. *Research Policy*, *42*(8), 1378–1388.
Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239.
Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, *178*(4040), 471–479.
Garfield, E. (1979). Is citation analysis a legitimate evaluation tool? *Scientometrics*, *1*(4), 359–375.
Garfield, E. (1990). Keywords plus-ISIís breakthrough retrieval method. Part 1. Expanding your searching power on Current Contents on Diskette. *Current Contents*, *1*(32), 5–9.

Garud, R., & Kumaraswamy, A. (2005). Vicious and virtuous circles in the management of knowledge: The case of Infosys Technologies. *MIS Quarterly*, 9–33.

Guan, J. C., & Liu, N. (2016). Exploitative and exploratory innovations in knowledge network and collaboration network: A patent analysis in the technological field of nano-energy. *Research Policy*, *45*(1), 97–112.

Guan, J. C., & Yan, Y. (2016). Technological proximity and recombinative innovation in the alternative energy field. *Research Policy*, *45*(8), 1460–1473. http://dx.doi.org/10.1016/j.respol.2016.05.002

Guan, J. C., Yan, Y., & Zhang, J. J. (2015). How do collaborative features affect scientific output? Evidences from wind power field. *Scientometrics*, *102*(1), 333–355.

Guan, J. C., Zhang, J. J., & Yan, Y. (2015). The impact of multilevel networks on innovation. *Research Policy*, *44*(3), 545–559.

Guan, J. C., Zuo, K. R., Chen, K. H., & Yam, R. C. (2016). Does country-level R&D efficiency benefit from the collaboration network structure? *Research Policy*, *45*(4), 770–784.

Hansen, M. T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*, *44*(1), 82–111.

Huselid, M. A. (1995). The impact of human resource management practices on turnover productivity, and corporate financial performance. *Academy of Management Journal*, *38*(3), 635–672.

Inzelt, A., Schubert, A., & Schubert, M. (2009). Incremental citation impact due to international co-authorship in Hungarian higher education institutions. *Scientometrics*, *78*(1), 37–43.

Krackhardt, D. (1990). Assessing the political landscape-structure, cognition, and power in organizations. *Administrative Science Quarterly*, *35*(2), 342–369.

Kuhn, T. S. (2012). *The structure of scientific revolutions*. London: University of Chicago Press, Ltd.

Letchford, A., Preis, T., & Moat, H. S. (2016). The advantage of simple paper abstracts. *Journal of Informetrics*, *10*(1), 1–8.

Leydesdorff, L., & Bornmann, L. (2011). Integrated impact indicators compared with impact factors: An alternative research design with policy implications. *Journal of the American Society for Information Science and Technology*, *62*(11), 2133–2146.

Leydesdorff, L., & Opthof, T. (2010). Scopus's source normalized impact per paper (SNIP) versus a journal impact factor based on fractional counting of citations. *Journal of the American Society for Information Science and Technology*, *61*(11), 2365–2369.

Leydesdorff, L., Carley, S., & Rafols, I. (2013). Global maps of science based on the new Web-of-Science categories? *Scientometrics*, *94*(2), 589–593.

Li, E. Y., Liao, C. H., & Yen, H. R. (2013). Co-authorship networks and research impact: A social capital perspective? *Research Policy*, *42*(9), 1515–1530.

Li, G.-C., Lai, R., D'Amour, A., Doolin, D. M., Sun, Y., Torvik, V. I., et al. (2014). Disambiguation and co-authorship networks of the US patent inventor database (1975–2010)? *Research Policy*, *43*(6), 941–955.

Lozano, G. A., Larivière, V., & Gingras, Y. (2012). The weakening relationship between the impact factor and papers' citations in the digital age. *Journal of the American Society for Information Science and Technology*, *63*(11), 2140–2145.

McFadyen, M. A., & Cannella, A. A. (2004). Social capital and knowledge creation: Diminishing returns of the number and strength of exchange relationships. *Academy of Management Journal*, *47*(5), 735–746.

Mingers, J., & Burrell, Q. L. (2006). Modeling citation behavior in management science journals. *Information Processing &Management*, *42*(6), 1451–1464.

Muñoz-Leiva, F., Viedma-del-Jesús, M. I., Sánchez-Fernández, J., & López-Herrera, A. G. (2012). An application of co-word analysis and bibliometric maps for detecting the most highlighting themes in the consumer behaviour research from a longitudinal perspective? *Quality & Quantity*, *46*(4), 1077–1095.

Newman, M. E. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences*, *101*(1), 5200–5205.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths? *Social Networks*, *32*(3), 245–251.

Otte, E., & Rousseau, R. (2002). Social network analysis: A powerful strategy, also for the information sciences. *Journal of Information Science*, *28*(6), 441–453.

Owens, S., & Driffill, L. (2008). How to change attitudes and behaviours in the context of energy. *Energy Policy*, *36*(12), 4412–4418.

Paruchuri, S. (2010). Intraorganizational networks interorganizational networks, and the impact of central inventors: A longitudinal study of pharmaceutical firms. *Organization Science*, *21*(1), 63–80.

Phelps, C., Heidl, R., & Wadhwa, A. (2012). Knowledge networks, and knowledge networks a review and research agenda. *Journal of Management*, *38*(4), 1115–1166.

Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B-Condensed Matter and Complex Systems*, *4*(2), 131–134.

Rodan, S., & Galunic, C. (2004). More than network structure: How knowledge heterogeneity influences managerial performance and innovativeness? *Strategic Management Journal*, *25*(6), 541–562.

Rousseau, R., García-Zorita, C., & Sanz-Casado, E. (2013). The h-bubble. *Journal of Informetrics*, *7*(2), 294–300.

Rowley, T., Behrens, D., & Krackhardt, D. (2000). Redundant governance structures: An analysis of structural and relational embeddedness in the steel and semiconductor industries? *Strategic Management Journal*, *21*(3), 369–386.

Sanz-Casado, E., Garcia-Zorita, J. C., Serrano-López, A. E., Larsen, B., & Ingwersen, P. (2013). Renewable energyresearch 1995–2009: A case study of wind power research in EU, Spain, Germany and Denmark. *Scientometrics*, *95*(1), 197–224.

Schilling, M. A., & Phelps, C. C. (2007). Interfirm collaboration networks: The impact of large-scale network structure on firm innovation. *Management Science*, *53*(7), 1113–1126.

Skilton, P. F. (2006). A comparative study of communal practice: Assessing the effects of taken-for-granted-ness on citation practice in scientific communities. *Scientometrics*, *68*(1), 73–96.

Su, H.-N., & Lee, P.-C. (2010). Mapping knowledge structure by keyword co-occurrence: A first look at journal papers in Technology Foresight. *Scientometrics*, *85*(1), 65–79.

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

Team, S. (2009). *Science of science (Sci2) tool*. Indiana University and SciTech Strategies [https://sci2.cns.iu.edu]

Thelwall, M., & Wilson, P. (2014). Regression for citation data: An evaluation of different methods. *Journal of Informetrics*, *8*(4), 963–971.

Thomson (2005). Glossary of Thomson scientific terminology. KeyWords Plus. Retrieved on July 28, 2016, from https://images.webofknowledge.com/WOKRS56B5/help/WOS/hp_full_record.html.

Tijssen, R. J. (1992). A quantitative assessment of interdisciplinary structures in science and technology: Co-classification analysis of energy research? *Research Policy*, *21*(1), 27–44.

Tsai, W., & Ghoshal, S. (1998). Social capital and value creation: The role of intrafirm networks? *Academy of Management Journal*, *41*(4), 464–476.

Tsai, W. (2001). Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of Management Journal*, *44*(5), 996–1004.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, *342*(6157), 468–472.

Van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538.

Van Raan, A. F. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, *67*(3), 491–502.

Von Hippel, E. (1994). Sticky information and the locus of problem solving: Implications for innovation. *Management Science*, *40*, 429–439.

Wales, W. J., Parida, V., & Patel, P. C. (2013). Too much of a good thing? Absorptive capacity, firm performance, and the moderating role of entrepreneurial orientation. *Strategic Management Journal*, *34*(5), 622–633.

Wang, J., & Shapira, P. (2011). Funding acknowledgement analysis: an enhanced tool to investigate research sponsorship impacts: The case of nanotechnology. *Scientometrics*, *87*(3), 563–586.

Wang, C. L., Rodan, S., Fruin, M., & Xu, X. (2014). Knowledge networks, collaboration networks, and exploratory innovation. *Academy of Management Journal*, *57*(2), 484–514.
Wang, J. (2016). Knowledge creation in collaboration networks: Effects of tie configuration? *Research Policy*, *45*(1), 68–80.
Wooldridge, J. (2015). *Introductory econometrics: A modern approach*. Mason: Cengage Learning.
Xie, S., Zhang, J., & Ho, Y.-S. (2008). Assessment of world aerosol research trends by bibliometric analysis? *Scientometrics*, *77*(1), 113–130.
Yang, S., Han, R., Wolfram, D., & Zhao, Y. (2016). Visualizing the intellectual structure of informationscience (2006–2015): Introducing author keyword coupling analysis. *Journal of Informetrics*, *10*(1), 132–150.
Zaheer, A., & Soda, G. (2009). Network evolution: The origins of structural holes? *Administrative Science Quarterly*, *54*(1), 1–31.
Zhang, J. J., Yan, Y., & Guan, J. C. (2015). Scientific relatedness in solar energy: A comparative study between the USA and China. *Scientometrics*, *102*(2), 1595–1613.