Journal of Informetrics 15 (2021) 101167

Contents lists available at ScienceDirect

Journal of Informetrics

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

A novel approach to measuring science-technology linkage: From the perspective of knowledge network coupling

Zhichao Ba[a], Zhentao Liang[b,*]

[a] School of Economics and Management, Nanjing University of Science and Technology, Nanjing, China
[b] School of Information Management, Wuhan University, Wuhan, China

ARTICLE INFO

Keywords:
Science-technology linkage
Knowledge network
Network coupling
Structural-coupling strength
Energy conservation

ABSTRACT

Identifying and measuring science-technology linkage is important for understanding interactions between science and technology (S&T). Previous studies have focused mainly on knowledge linkages of knowledge systems between S&T but have ignored their structural linkages. To this end, we propose a novel knowledge network coupling approach to gauge network linkage between S&T by integrating knowledge linkages and structural linkages. Four network construction strategies were first adopted to determine appropriate knowledge networks of S&T, and then their coupling strengths over time were calculated based on similarities of coupling nodes' degree distribution and similarities of coupling edges' weight distribution. An experimental study in the field of energy conservation confirms that our approach was indeed successful in revealing interactions between S&T. The proposed approach enriches the current methodology for measuring S&T linkages and provides references for policymakers to conduct policy adjustments, by identifying the lead-lag relationship between S&T.

1. Introduction

Science and technology (S&T), the two foremost forces that jointly determine the progress of innovations, are viewed as closely connected, interacting, and interdependent systems (Wang & Guan, 2011). It has been proved in practice that a plethora of breakthroughs or growth points have emerged from dynamic interactions between S&T (Mukherjee, Romero, Jones, & Uzzi, 2017). According to the linear model of science, the relationship between S&T is a one-way flow that starts with basic scientific research and moves to applied research, inventions, and products. However, non-linear relationships, by which technological advances drive scientific progress may also exist (Ahmadpoor & Jones, 2017). To understand and promote advances in S&T, extensive attention has been given by researchers and institutions to the study of S&T linkages. The term "S&T linkages" refers to the internal relationships of knowledge elements (i.e., terms, authors, and references) between the scientific knowledge system and the technological knowledge system (Xu, Winnink, Yue, Liu, & Yuan, 2020). The purpose of detecting and measuring these linkages is to determine their direction, strength, and structure and to reveal the processes of interaction between S&T. These linkages record the histories of knowledge diffusion and transfer between S&T (Hu & Rousseau, 2018), which can be tracked to discover technological opportunities (Albert, 2016), understand university-industry-government interaction (Leydesdoref & Meyer, 2007), measure innovation (Du, Li, Guo, & Tang, 2019), and construct high-quality roadmaps (Boyack, Klavans, Small, & Ungar, 2014; Hu & Rousseau, 2018). Thus, it is important to explore how these linkages are measured based on knowledge associations between S&T (Ke, 2020; Nguyen, Liu, Khor, Nanetti, & Cheong, 2019; Xu et al., 2019).

* Corresponding author.
E-mail addresses: bazhichaoty@njust.edu.cn (Z. Ba), liangzht5@whu.edu.cn (Z. Liang).

https://doi.org/10.1016/j.joi.2021.101167
Received 25 November 2020; Received in revised form 11 April 2021; Accepted 19 April 2021
Available online 11 May 2021
1751-1577/

Scientific publications and patents are widely considered as effective proxies of basic scientific research and technological advances, respectively (Xu et al., 2019). Previous research on the linkages between S&T has focused mainly on the analysis of patent-literature references (Huang, Yang, & Chen, 2015; Nguyen et al., 2019; Roach & Cohen, 2013) and author-inventor links (Boyach & Klavans, 2008; Breschi & Catalini, 2010). One of the main problems with the research of this sort is that it is difficult to reveal the semantic associations between S&T on the micro-content level (Bassecoulard & Zitt, 2004). Therefore, lexical- or topic-based approaches have been developed based on textual analysis to identify the content linkages between S&T, to detect hidden interactive and exclusive relations, and to analyze S&T interactions from multiple perspectives (Shibata, Kajikawa, & Sakata, 2010; Xu et al., 2020). However, these approaches are still unable to identify structural linkages between S&T knowledge systems.

Linkages between the scientific knowledge system and the technological knowledge system consist of knowledge linkages and structural linkages. Knowledge linkages reflect the correspondence between the knowledge elements or nodes of the two systems, and structural linkages represent the consistency of knowledge linkages or edges. Structural linkages may be even more important than knowledge linkages, because innovation opportunities and new elements can be researched based on their ties (Guan and Liu, 2016); therefore, it promises to be beneficial to measure S&T linkages in terms of both structural linkages and knowledge linkages. Furthermore, knowledge networks, formed by the coupling of knowledge elements, can be used as a proxy for scientific or technological knowledge systems (Larsen, 2008; Ma'ayan, 2011). Transforming the linear model into a network model should be more advantageous to quantifying and modeling S&T linkages (Murray, 2002). Therefore, we propose a novel network coupling approach to measuring S&T linkages by identifying the coupling relationships between scientific knowledge networks and technological knowledge networks. In particular, we will try to answer the following questions:

1) How can we construct appropriate knowledge networks for science and technology to explore their linkages?
2) How can we measure the coupling of knowledge networks between S&T by considering their structural and knowledge linkages?
3) What are the interaction patterns of S&T network linkages?

To verify the effectiveness and robustness of our proposed approach, we use a case study involving energy conservation, a field that urgently requires a deep integration of S&T to innovate (Croucher, 2011). The analytical methodology employed in this paper can be generally applied to explorations of S&T knowledge linkages. By providing an actionable approach through network modeling, our study contributes a new method for measuring S&T linkages in terms of knowledge content and knowledge structure, which can be used as a tool for discovering knowledge flows, linked evolution paths, and new knowledge growth points in the dynamic evolution of S&T.

2. Literature review

2.1. Measurement of the linkages between science and technology

To measure the relationships between S&T, several quantitative approaches have been introduced in the literature from the viewpoint of patent-literature knowledge linkages, such as citation relations, author-inventor links, international patent classification (IPC)-journal classification mapping, and topic linkages. These approaches generally regard scientific publications and patents as proxies for the science and technology systems, respectively.

2.1.1. Using citation relations to identify science-technology linkages

Citation-based linkages, such as nonpatent references (NPRs) and patents cited in the literature, have been extensively studied to measure the increasing interdependencies and interactions between S&T (Du et al., 2019; Roach & Cohen, 2013; Verhoeven, Bakker, & Veugelers, 2016). As a direct and straightforward form of linkage, citations confirm and illustrate the social character of knowledge diffusion and transfer between S&T (Li, Azoulay, & Sampat, 2017). Hu and Rousseau (2018) utilized an NPR-based approach to detect knowledge transitions from discovery- and application-oriented science to technology and identified different states of knowledge utilization in the evolution of S&T. Popp (2017) found that extensively cited academic literature was valuable to the creation of applied technology. Based on a large-scale patent dataset and NPRs, Ahmadpoor and Jones (2017) analyzed citation patterns between S&T from multiple perspectives. They found that a majority of patents could be linked to prior scientific advances, at distances of 2 to 4 citations. The characteristics of documents, scientific/technological fields, and institutions close to the border of S&T were also studied. On the other hand, Glanzel and Meyer (2003) and Kwon, Liu, Porter, and Youtie (2019) explored the impact of technology on basic science by analyzing patents cited in the literature and concluded that research addressing emerging technological ideas had a greater scientific impact.

The above one-way analyses of S&T citation linkages offer limited insight to reveal the interactions of S&T from a perspective of mutuality (Xu et al., 2020). Specifically, they can only observe patent-to-paper citations or paper-to-patent citations, and they are unable to investigate two-way relationships between S&T. To this end, some scholars propose a cross-citation analytical method for investigating mutual relations between them. Gao, Ding, Teng, and Pang (2012) implemented hybrid documents co-citation analysis (HDCCA) to describe the interplay between S&T and deduced that science and technology undertook different functions and predominated during different periods of technology diffusion. Huang et al. (2015) conducted a cross-citation experiment to understand the characteristics, speed, and dynamic changes of S&T interactions in fuel cells and discovered that papers citing patents were more time-sensitive than patents citing papers. Although citation analysis for uncovering S&T linkages is objective, available, and standardized, some researchers questioned the reliability of the approach, due to the complexity and ambiguities of citation practices. It is suspected whether citations ultimately decided by the patent examiner can truly represent the actual references of inventors (Criscuolo & Verspagen, 2008). More important, citation analysis failed to identify content linkages between S&T that could unveil some implicit interactive and exclusive relations based on textual analysis.

2.1.2. Using author-inventor links to identify science-technology linkages

By observing the dual role of authors or inventors in the S&T systems, some scholars matched inventors with authors to identify points of exchange between basic science and technology development (Boyack & Klavans, 2008; Noyons, Raan, Grupp, & Schmoch, 1994; Zhang, Liu, & Wei, 2019). Such author-inventor links play a bridging role in S&T and are an appropriate proxy of their interactions and knowledge diffusion (Cassiman, Glenisson, & Looy, 2007; Forti, Sobrero, & Franzoni, 2007). Wang and Guan (2011) combined patent citations and author-inventor links to measure S&T interactions in nanotechnology and found that it was advantageous for inventors to engage in scientific research for their patent invention. Similarly, Meyer (2006) carried out an exploratory comparison of inventor-authors with their non-inventing peers in non-science and technology and asserted that inventor-authors were more productive and more highly cited than their peers who concentrate on scholarly publication. Forti et al. (2007) disclosed a similar finding that participation in technological activities positively influenced scientific research behavior.

A basic requirement for the author-inventor matching approach is that there has to be a sample consisting of a plethora of researchers who play a dual role in the S&T system (Boyack & Klavans, 2008). However, the number of author-inventors in practice is relatively small, which naturally places restrictions on the application of the approach to exploring the S&T overlap. Furthermore, the approach depends heavily on the accuracy of the name disambiguation process, which can be labor-intensive work for large data analysis (Ranaei, Suominen, & Dedehayir, 2017). In addition, the approach is also unable to recognize semantic associations between S&T at the micro-content level.

2.1.3. Using IPC-ISI journal classification mapping to identify science-technology linkages

As a set of internationally accepted classification and retrieval tools, IPC is generally used to organize and classify technologies into a series of classes or subclasses, while ISI (Institute for Scientific Information) journal classification is used to categorize science into various subject categories. Matching IPC with ISI journal classification can reveal similarities in developing trends and their directions between S&T at the domain level (Han & Magee, 2018). IPC-ISI journal classification linkage, which involves correlations at a higher aggregation level, tries to establish a subject mapping of the knowledge organization system between S&T. Verbeek, Debackere, and Luwel (2002) matched IPC 4-digit classes with ISI journal classifications to trace the IPC class of a given patent and the science-domain classification of the journal in which the "source" publication appeared. Han and Magee (2018) also integrated citation-based linkages and IPC-ISI journal classification linkages to identity science-technology relationships. However, further investigation is required to tell whether an accurate subject correspondence between S&T, based on IPC-ISI journal classification linkage, can be achieved.

2.1.4. Mining topic linkages between S&T

To identify content linkages between S&T, lexical and topic-based approaches have been developed to mine the semantic relations hidden in the textual content of scientific publications and patents. The general practice is to extract topics from publications and patents and calculate their similarities to construct topic linkages (Xu et al., 2019). For example, Ranaei et al. (2017) utilized Latent Dirichlet Allocation (LDA) to identify topics and quantified S&T overlap by calculating the distribution of S&T for different topics. Shibata et al. (2010) and Xu et al. (2019) in a similar vein adopted different topic extraction algorithms to detect science-technology topic linkages. Taking advantage of topic association, Xu et al. (2020) performed a multi-fusion analysis that integrated three types of topic linkages, including co-word linkages, co-author linkages, and co-citation linkages.

While topic-based approaches can detect some hidden semantic relations at the micro-content level (Shibata et al., 2010), they only unveil linear interactions between S&T and cannot catch structural linkages between S&T knowledge systems. Thus, more sophisticated indicators and methods are required to capture multifaceted relationships and interplay between S&T. Moving beyond linear linkages to network linkages, we propose a novel network coupling approach to understanding science-technology interactions, which sheds light on networks and crossroads of scientific and technological activity.

2.2. Knowledge networks of science and technology

Scientific and technological knowledge systems can be represented in the form of conceptual networks of knowledge elements and element relationships (Yi & Choi, 2012). Several types of knowledge networks (i.e., citation, co-author, and co-word networks) have been constructed to understand how the knowledge systems of S&T are organized and evolve. Considering that keywords are the most fundamental carriers of knowledge, co-word network analysis has been extensively applied to elucidate, map, and analyze knowledge structure, flows, and trend distribution within semantic domains (Behrouzi, Shafaeipour Sarmoor, Hajsadeghi, & Kavousi, 2020; Ding, Chowdhury, & Foo, 2001; Haunschild, Leydesdorff, Bornmann, Hellsten, & Marx, 2019), especially to compare the network structures of different knowledge systems (Chen & Xiao, 2016).

The construction of knowledge networks based on the co-word approach faces two major challenges. The first is a large computational burden. Keywords are generally extracted from the titles, abstracts, customized keywords, and contents of publications or patents, the number of which is vast so that we have to deal with much larger networks than usual (Yi & Choi, 2012). An effective method is to select some representative or "high quality" keywords to create a network (Haunschild et al., 2019; Lee, Su, & Chan, 2010). The second challenge is that compiling a list of representative keywords is not straightforward. Because keywords appeared in various forms, linguistic structures, and situational contexts, their boundaries are hard to recognize and segment (Li & Yan, 2019). Thus, it is desirable to combine an appropriate keyword extraction algorithm with some form of manual labeling practice.

2.3. Coupling measurement of knowledge networks

To construct network linkages, the coupling between scientific and technological knowledge networks should be determined. *Network coupling,* a tendency of nodes in one network to be interdependent with neighboring nodes in another network, is associated with the dynamic interaction of interdependent networks with common nodes or links (Hu et al., 2013). Currently, there are two main ways to calculate network coupling (Parshani, Rozenblat, Ietri, Ducruet, & Havlin, 2010; Wang, Jia, Wang, & Liu, 2019; Yan & Ding, 2012). The first is to explore the similarities of interdependent networks in terms of different coupling methods or network structures, such as inter-similarity (Parshani et al., 2010), multiple support-dependency relations (Shao, Buldyrev, Havlin, & Stanley, 2011), assortativity (Valdez, Macri, Stanley. & Braunstein, 2013), clustering (Huang et al., 2013), degree distribution (Emmerich, Bunde, & Havlin, 2014), and spatially embedded networks (Bashan, Berezin, Buldyrev, & Havlin, 2013). These are primarily used to quantify the coupling of interdependent networks in terms of node coupling and structural coupling.

The second way is to draw the common nodes and edges of interdependent networks to construct an independent coupled network and to inspect and analyze the structure of this coupled network and its evolution (Wang et al., 2019). For example, by coupling references and words from titles and abstracts, Boyack and Klavans (2010) built a bibliographic coupling-based, hybrid citation-text network to represent the research front and found that it outperformed other networks, such as co-citation networks, bibliographic coupling networks, and direct citation networks. Perianes-Rodriguez, Waltman, and Van Eck (2016) constructed journal coupling networks based on the citation networks of different journals to identify the journals that were most strongly related to a specific focal journal. To investigate research interactions and scholarly communications, Yan and Ding (2012) examined the similarities among six types of scholarly networks and tried to build a hybrid scholarly network based on the coupling relationships of these six networks. More recently, Boyack and Klavans (2020) investigated the performance of scholarly networks based on three types of relatedness (textual, citation, and hybrid) on the task of topic clustering and found that hybrid network models outperformed those based on single relatedness measures. Similar results were observed in the case of author co-citation networks. More sub-fields and details can be identified as more relationships are integrated into an author co-citation network (Bu, Ni, & Huang, 2017). It has also been demonstrated that a coupling-based hybrid network is superior to a single-type network for identifying topics (Glänzel & Thijs, 2017), representing research fronts (Kleminski, Kazienko, & Kajdanowicz, 2020), and mapping the backbone of science or technology (Ahlgren, Chen, Colliander, & van Eck, 2020).

3. Research framework and methodology

3.1. Research framework

As shown in Fig.1, our research framework consists of three phases. The first step is to detect sentence boundaries, tokenize each detected sentence, recognize entity mentions, filter stopwords, perform part-of-speech (POS) tagging and stemming, and then, in the preprocessing phase, evaluate the meaningfulness of the extracted keywords. In the second phase, the frequency of the keywords is counted, and the words are divided into four groups: Top-N high-weight terms, shared terms that appear concurrently in both S&T, terms neighboring shared terms, and other terms. Multiple term combination strategies are adopted to proxy for the knowledge base of S&T and to construct additional appropriate knowledge networks for S&T, based on the concurrence of terms. The final phase consists of calculating the coupling of the constructed knowledge networks between S&T and then building network linkages. We describe each phase in more detail in the subsections below.

3.2. Keyword extraction

While author keywords are provided for scientific publications, this is not the case for patents. Previous studies have shown that scientific publications and patents use different glossaries in different rhetorical ways (Xu et al., 2019), and freely chosen author keywords increase this heterogeneity. To alleviate this problem, we discarded the authors' keywords and instead extracted keywords from the titles and abstracts of the scientific publications and patents. An unsupervised and domain-independent algorithm, rapid

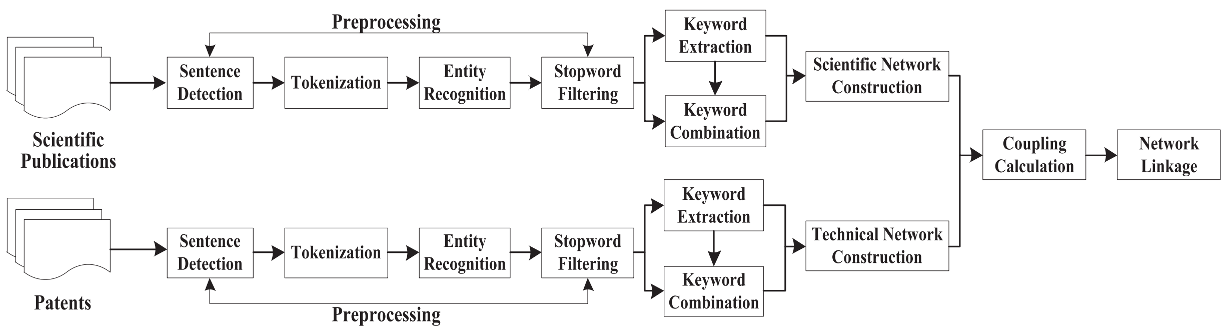

Fig. 1. Research framework of network linkages between S&T.

automatic keyword extraction (RAKE), was employed in the extraction process (Rose, Engel, Cramer, & Cowley, 2010). RAKE first obtains the candidate keywords by punctuation and stopword segmentation. These candidates are then scored based on the ratio of their degree in the co-occurrence network to term frequency (deg(w)/freq(w)). Furthermore, this algorithm supports the extraction of keywords consisting of multiple words. Keywords that occur more frequently and dominantly in longer candidates are given higher scores (Li & Yan, 2019; Rose et al., 2010).

However, RAKE tends to rank long phrases higher than shorter ones, which is inconsistent with the empirical practice that author keywords primarily consist of two to three words (Li & Yan, 2019). Without limitations, phrases that contains more than three words are likely to occupy the top positions of the RAKE output results. Following the practice of Li and Yan (2019), we filtered bigrams and trigrams that matched the part-of-speech templates below. They were selected as valid keywords to perform subsequent analyses.

- Noun + Noun
- Adjective + Noun
- Adjective + Adjective + Noun
- Adjective + Noun+ Noun
- Noun+ Adjective + Noun
- Noun+ Noun+ Noun
- Noun+ Preposition + Noun

On average, in our dataset, scientific publications had 5.06 author keywords. Therefore, we chose the five keywords with the highest rankings as the final, valid keywords to represent a document. The same extraction process was applied to the titles and abstracts separately. Of the five keywords, two keywords that represented the core idea of the publication were extracted from the title, and the other three were taken from the abstract. If fewer than two keywords were extracted from the title, we took more keywords from the abstract to complete the set of five.

3.3. Knowledge network construction of S&T

Building appropriate scientific and technological knowledge networks is a critical step in the analysis of science-technology linkages from a network-coupling perspective. However, it is difficult to determine the rationality and effectiveness of an exploration of science-technology linkages when only one type of scientific and technological knowledge network is examined. In contrast, building multiple knowledge networks not only offers reference for appropriate network construction but also verifies the robustness of the approach by making it possible to compare the coupling results based on different network models.

Based on the distribution and frequency of the keywords in the S&T knowledge systems, we divided them into four groups: Top-N high-weight terms, shared terms that appear concurrently in both S&T, terms neighboring shared terms, and other terms. Scientific and technological knowledge networks were constructed according to the following term combination strategies: all terms, Top-N high-frequency terms, Top-N + shared terms, Top-N + shared + neighbor terms. Network coupling measured with the all terms strategy provides a holistic view of the relationship between the S&T knowledge networks, whereas high-weight terms represent important knowledge in the scientific or technological network with less degree of overlap. On the other hand, shared terms reflect common focuses in S&T. They proxy for other types of investigation of focuses in S&T, together with the neighbor knowledge. Fig.2 presents an example of network linkages between S&T.

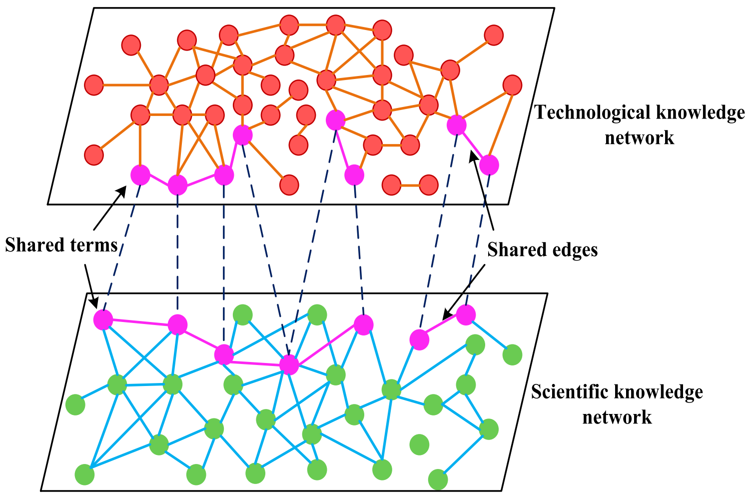

Fig. 2. An example of network linkages between S&T.

3.4. Coupling measurement of knowledge networks between S&T

Network coupling is embodied in node coupling and structural coupling between or among networks (Barrat, Barthélemy, & Vespignani, 2004; Liu, Li, & Jia, 2016). The node coupling of knowledge networks reveals the knowledge linkages between them formed by the correspondence of network nodes, while structural coupling consists of structural linkages arising from the consistency of network edges. Therefore, we propose a new approach to measuring the coupling of knowledge networks between S&T, namely, gauging their node-coupling strength and structural-coupling strength.

In our approach, the node-coupling strength of knowledge networks between S&T is calculated based on similarities in the degree distribution of the coupling nodes in the two networks, rather than on the coupling nodes' quantitative characteristics. This measurement takes into account structural information about the coupling nodes in the S&T knowledge networks. While the structural-coupling strength of the two networks is calculated based on similarities in the weight distribution of the coupling edges, the coupling strength of the networks is calculated with the following steps.

Step 1: Node sets N_1 and N_2 are extracted from the scientific knowledge network G_1 and the technological knowledge network G_2, respectively, and extended to a larger node set N, where $N = N_1 \cup N_2$.
Step 2: The degree distribution f_1 of N is calculated according to the weighted degree of centrality of each node in G_1; it is recorded as 0 when nodes do not exist in G_1. The degree distribution f_2 is calculated in a similar fashion for G_2.
Step 3: The similarities between degree distribution f_1 and f_2 are calculated as the node-coupling strength of G_1 and G_2 at time t, namely,$f_{G_1G_2}(t)$.
Step 4: Edge sets E_1 and E_2 are extracted from the scientific knowledge network G_1 and the technological knowledge network G_2, respectively, and extended to a larger edge set E, where $E = E_1 \cup E_2$.
Step 5: The edge weight distribution g_1 of E is calculated according to the weight of each edge in G_1, where it is recorded as 0 when edges do not exist in G_1. The edge weight distribution g_2 in G_2 is calculated in a similar way.
Step 6: The similarities between the edge weight distribution g_1 and g_2 are calculated as the edge-coupling strength of G_1 and G_2 at time t, namely,$g_{G_1G_2}(t)$.
Step 7: The final coupling strength between the scientific knowledge network G_1 and the technological knowledge network G_2 at time t is calculated as follows:

$$\xi_{G_1G_2}(t) = \alpha f_{G_1G_2}(t) + \beta g_{G_1G_2}(t),$$

where$\xi_{G_1G_2}(t)$indicates the coupling strength of G_1 and G_2 at time t, and α and β are the weight coefficients of the node-coupling strength and the structural-coupling strength, which are determined by their effects on the network coupling.

4. Empirical analysis

4.1. Data sources

The proposed approach was tested by analyzing scientific publications and patents granted in the field of energy conservation. This field, which addresses urgent problems such as global climate change, maintaining sustainable economic development, and adequately coping with the contradictions of energy consumption, requires a deep integration of S&T for successful innovation to take place (Croucher, 2011; Peterman, Kourula, & Levitt, 2014). Scientific publications and patents granted in this field were gathered from the Web of Science database (WoS) and the United States Patent and Trademarks Office (USPTO) database, respectively. Data collection was performed on September 5, 2020. After eliminating duplicate publications (with identical titles) and duplicate patent documents (with identical patent numbers), 100,736 publications and 66,343 patents granted between 1979 and 2019 were obtained. However, only a small number of scientific energy conservation publications and patents were published throughout the 1990s, while there was remarkable growth during the 2000s, as shown in Fig. 3. Considering that the knowledge networks of S&T were exceedingly sparse, we carried out our research based on the energy conservation publications and patents from 2000 to 2019, which resulted in 82,031 articles and 64,203 patents. Note that there was a sudden drop in patent numbers in 2001. The two database-specific queries described in the next two sections were applied to gather the data.

4.1.1. Selecting the scientific publications

To select the scientific publications from the WoS database, we utilized the following search query:

TI=((“power-efficient”) OR (“energy efficiency”) OR (“low-power”) OR (“low-energy”) OR (“energy-saving”) OR (“fuel-efficient”) OR (“energy-conscious”) OR (“high-efficiency”) OR (“low power-consumption”) OR (“energy efficient”) OR (“power allocation”) OR (“energy consumption”) OR (“smart grid”) OR (“smart metering”) OR (“enhancement of efficiency”) OR (“energy management”) OR (“energy performance”) OR (“rational use”) OR (“electricity consumption”) OR (“system efficiency”) OR (“energy efficient prosperity”) OR (“eco-power”) OR (“energy-efficient window”) OR (“energy analysis”) OR (“efficiency factor”) OR (“energy-saving awareness”) OR (“energy efficiency measures”) OR (“energy utilization”) OR (“energy conservation”) OR (“intelligent control”) OR (“energy demand”) OR (“sustainable energy”) OR (“energy policy”) OR (“energy efficient systems”) OR (“energy usage”) OR (“maximum efficiency”) OR (“energy efficiency ratio”) OR (“energy efficiency strategies”) OR (“energy simulation”) OR (“energy use”) OR (“energy storage”) OR (“energy forecasting”) OR (“consumption of energy”) OR (“conscious use”)) AND DT=“Article” AND PY=(1979–2019).

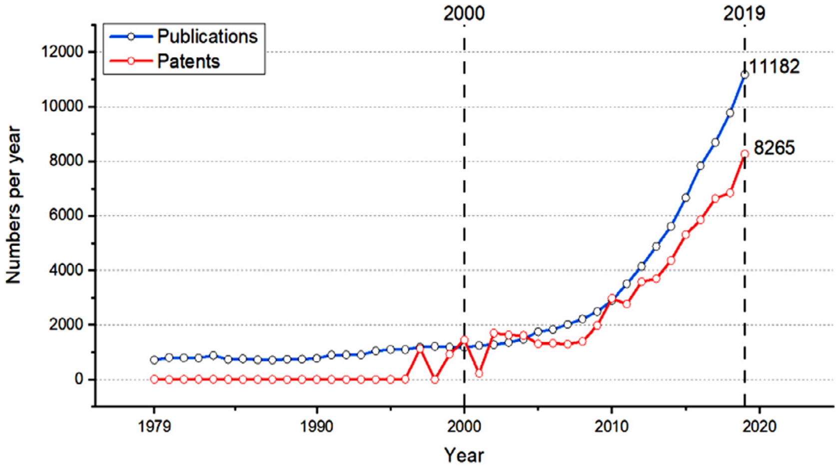

Fig. 3. The number of energy conservation publications and patents granted per year from 1979 to 2019.

4.1.2. Selecting the patent publications

The United States Patent and Trademark Office (USPTO) provides a special, detailed search guide for identifying and capturing energy conservation patents, which is more accurate than searching through terms based on other databases. Thus, we conducted a retrieval operation based on the following IPC lists.

Energy conservation	USPC	IPC
Energy storage or distribution	307; 700/295+	H01M; H02J
Thermal	702/130+	F24H 7/00
Insulating	52/404.1+	E04B 1/74, E04B 2/00
Static structures	52/404.1	E04B 1/62; E04D; E06B 7/098; E06B
Transportation	244/172.7	B64G 1/44
Land vehicle	105; 180; 280	B60L 8/00; B60K 16/00
Alternative-power vehicle (e.g. hydrogen)	180/2.1–2.2, 54.1	B60L 8/00
Electric vehicle	180/65.1; 180/65.21; 701/22+	B60L 9/00
Fuel-cell-powered vehicles	180/65.21, 180/65.31	B60L 11/18
Drag reduction	105/1.1; 296/180.1+; 296/181.5	B62D 35/00
Human-powered vehicle	180/205; 280/200+	B62M
Hybrid-powered vehicle	180/65.21–65.29	B60K 6/20, 6/00
Rail vehicle	105/49	B61C 1/00
Roadway (e.g. recycled surface, all-weather bikeways)	404	E01C 1/00
Wave-powered boat motors	440/9	B63H 19/02
Wind-powered boat motors	440/8	B63H 13/00
Wind-powered ships	114/102.1+	B63H 9/00

The above official classifications of energy conservation technology can be retrieved from https://www.uspto.gov/web/patents/classification/ international/est_concordance.htm.

4.2. Constructing knowledge networks

4.2.1. Knowledge extraction

To evaluate the keyword extraction performance for scientific publications and patents, we compared the extraction results based on several different algorithms, including TF-IDF, TextRank, and RAKE. We adopted the coding framework developed by Li and Yan (2019), which divides the meaningfulness of keywords into three levels:

- Level 1 (scored as 1): The extracted keyword is meaningful and reflects the major topic of the document.
- Level 2 (scored as 0.5): The extracted keyword is meaningful but does not relate to the major topic of the document.
- Level 3 (scored as 0): The extracted keyword is irrelevant and, therefore, meaningless for our purposes.

Two coders individually rated 1000 extracted keywords from 200 randomly selected scientific publications and patents. Five documents were selected from each year represented in the dataset. Inter-coder reliability was tested with Cohen's weight kappa (Ben-David, 2008); the resulting Cohen's kappa (κ) was 0.783, showing a substantial strength of agreement.

Fig. 4 presents the mean scores for keyword meaningfulness in the sample publications and patents from 2000 to 2019. The average score of each algorithm is represented by a horizontal line followed by its numerical value. It can be seen that RAKE had the best extraction performance (the average scores were 0.757 and 0.772 for scientific publications and patents, respectively), while both RAKE and TextRank outperform TF-IDF significantly. Furthermore, patents yielded better results from all three algorithms than

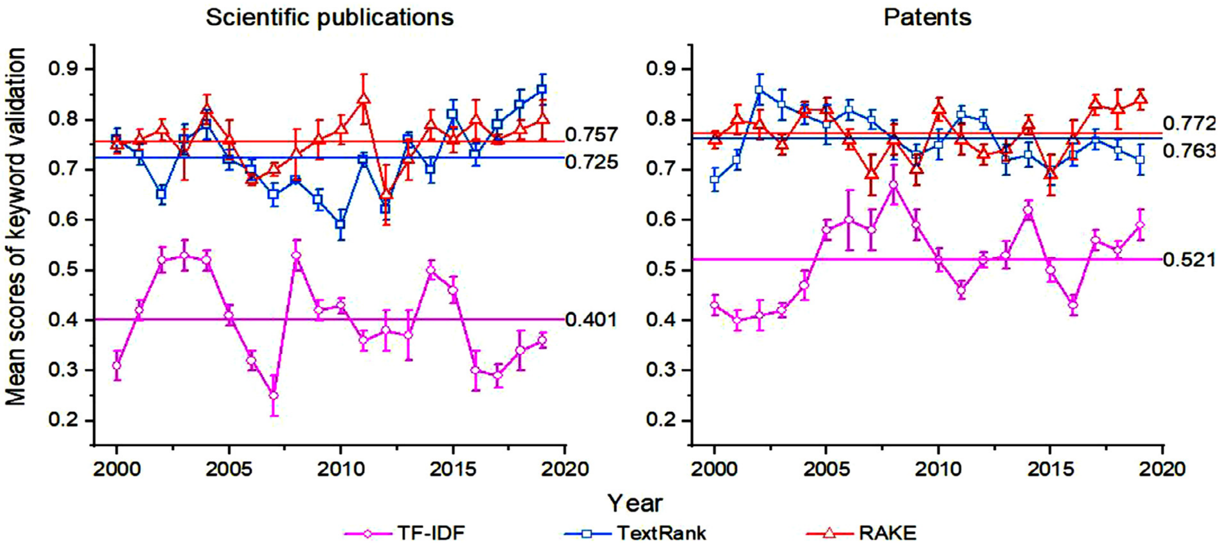

Fig. 4. TF-IDF, TextRank, and RAKE performance results for keyword extraction from scientific publications and patents.

Table 1
Top 10 keywords with the highest frequency extracted from the scientific publications and patents.

Rank	Scientific publications	Frequency	Patents	Frequency
1	energy efficiency	3157	fuel cell	3932
2	energy consumption	2401	fuel cell system	2310
3	wireless sensor network	1682	battery pack	2257
4	high efficiency	1593	secondary battery	1800
5	low power	1244	fuel cell stack	1727
6	thermal energy storage	1059	positive electrode	1412
7	low energy	1030	negative electrode	1383
8	smart grid	972	battery cell	1272
9	energy storage	934	lithium secondary battery	1071
10	energy storage system	868	membrane electrode assembly	1042

scientific publications. These results suggest that most of the keywords extracted by RAKE are meaningful terms that can be used to construct knowledge networks for S&T.

Based on the selected keyword extraction algorithm (i.e., RAKE), we obtained 229,776 distinct keywords from the scientific publications, with an average of 4.76 keywords per document, and 105,391 distinct keywords from the patents, with an average of 4.45 keywords per document. Since some documents in earlier years had a short or even no abstract, the figures were slightly lower than our extraction limit (5). Altogether, there were 326,952 distinct keywords in the S&T documents, of which 8215 keywords were shared by both scientific publications and patents. This confirms the existence of knowledge heterogeneity between science and technology, at least from the perspective of lexical expression.

Table 1 shows the ten most common keywords found in scientific publications and patents. It can be seen that they reflect the different focuses of science and technology in the field of energy conservation. Researchers who have investigated problems in this field from a macroscopic perspective have tended to focus on problem-oriented terms (e.g., *energy efficiency, consumption*, and *storage*) to deliver their findings. On the other hand, the words used in the titles and abstracts of the patents were more object-oriented, which mostly concerned specific technologies or products such as fuel cells and batteries. On average, the frequency of the occurrence of technological keywords was significantly higher than that of scientific keywords (2.71 vs. 1.70, Mann-Whitney U test, $p < 0.001$). This suggests an intriguing phenomenon that keywords in patents are more likely to be reused than those in scientific publications.

4.2.2. Descriptive statistics of the S&T knowledge networks

To construct a temporal knowledge network, the scientific publications and patents were grouped by year, resulting in 20 knowledge networks for S&T, respectively. Table 2 presents the descriptive statistics of full knowledge networks of S&T (using the all-terms strategy). In general, the size of the scientific knowledge networks was larger than the size of the technological ones, since the number of scientific publications was consistently higher than that of the patents after 2003. The former had a lower density but a slightly higher average clustering coefficient than the latter. It is evident that S&T networks consist of many disconnected components; that is because some keywords were used in only a small set of documents. We extracted the largest component from each network and found that the largest component contained most of the nodes and edges in the full networks and that the ratio increased with

Table 2
Descriptive statistics of S&T knowledge network properties (all terms strategy).

Network properties	Scientific knowledge networks (N = 20)			Technological knowledge networks (N = 20)		
	Mean	Min	Max	Mean	Min	Max
Full network						
Nodes	15,246.95	4760	39,054	8809.75	936	18,721
Edges	36,630.65	10,055	101,247	22,998.45	1848	53,844
Density	4.59×10^{-4}	1.33×10^{-4}	8.95×10^{-4}	8.56×10^{-4}	3.07×10^{-4}	4.22×10^{-3}
Avg. clustering coefficient	0.95	0.93	0.96	0.91	0.88	0.96
Components	1119.70	532	1979	471.15	152	697
Largest component						
Nodes	9258.00	696	28,912	6571.35	49	15,641
Edges	24,872.25	1695	81,499	18,929	118	48,333
Diameter	15.6	12	21	12.75	6	18
Avg. clustering coefficient	0.92	0.91	0.93	0.89	0.86	0.90
Avg. shortest path length	5.61	4.71	7.27	4.53	3.02	5.55

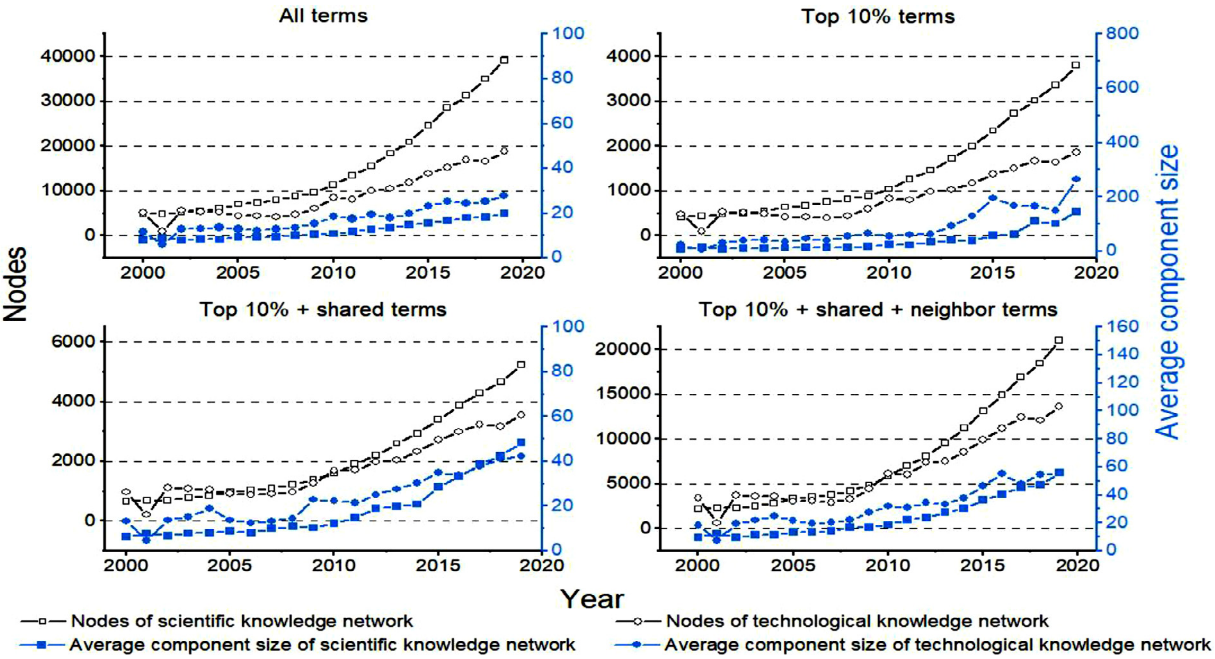

Fig. 5. Annual sizes and average component sizes of the S&T knowledge networks.

time. Furthermore, the largest components had higher average clustering coefficients (Sci: 0.92 vs. 0.0004; Tech: 0.89 vs. 0.001) and smaller average shortest path lengths (Sci: 5.61 vs. 6.30; Tech: 4.53 vs. 5.72) than random Erdős-Rényi networks with the same size and density. This suggests that the small world effect exists in S&T knowledge networks, whereby most knowledge can connect with any other knowledge through a small number of steps (Watts & Strogatz, 1998).

Fig. 5 presents the annual sizes and average component sizes of the S&T knowledge networks (other temporal statistics of the S&T networks are provided in Appendix A). It is seen in Fig. 5 that the gaps between the sizes of the S&T knowledge networks became greater over time for the four network construction strategies. As the network sizes increased, the two networks showed an upward trend in the number of components but an opposite trend in density (see Table A.1). Moreover, it is interesting that the number of components in the technological knowledge network was not only smaller than the number in the scientific knowledge network, even when the technological network was larger than the scientific network, but it had a consistently larger average size except in the year 2001. The increase in the number of components indicates that a group of new keywords appeared in the same document, forming a fully connected network that was disconnected from the existing communities. Therefore, in the scientific knowledge network, a higher component number but a smaller component size suggests that scientific researchers apply more new keywords in their work than patent seekers, either by combining existing keywords or by proposing novel ones. This is consistent with our findings in Section 4.2.1, that patents tend to reuse existing keywords.

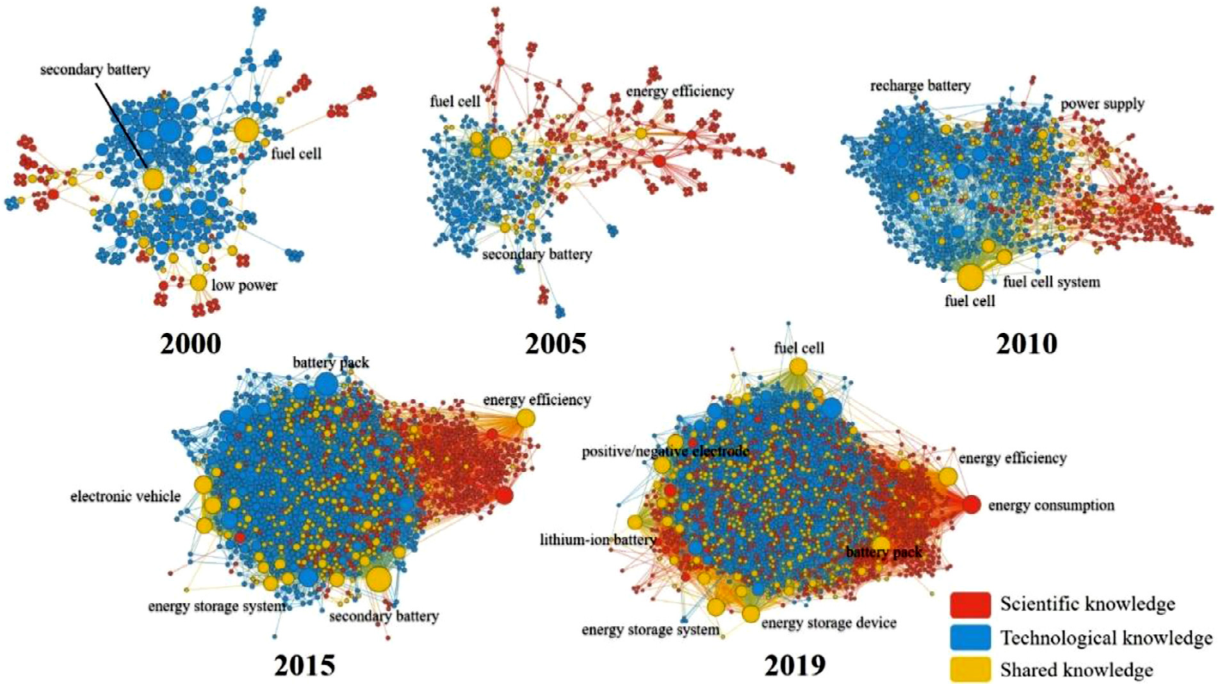

Fig. 6. Temporal snapshots of the S&T knowledge networks (nodes with an above-average degree).

By overlapping the S&T knowledge networks in five different years, Fig. 6 provides insight into their coupling and evolution patterns from a visualization perspective. The nodes represent terms that appear in scientific (red), technological (blue) and both (yellow) knowledge networks. The edges represent their co-occurrence relationships. Their sizes are determined by their frequency of occurrence. It is evident that the network size and coupling strength both increase with time. In 2015 and 2019, terms from the two networks cannot be easily separated by the force-directed layout algorithm, and nodes of different colors are strongly intertwined. *Fuel cell* is a prominent shared term appearing in all the networks, while terms like *energy efficiency, energy storage system, electronic vehicle*, and *lithium-ion battery* play important roles in bridging other terms between S&T.

Nonetheless, a pair-wise network comparison is needed to quantitatively reveal the coupling effect between the S&T networks. We applied the proposed coupling measurement to the empirical dataset, and we report the results in the next section. Comparisons of the different network construction strategies and other network coupling measurements are also presented.

4.3. *Measuring the coupling of S&T knowledge networks*

Fig. 7 displays the pair-wise coupling results of the S&T knowledge networks from 2000 to 2019. The color of each cell represents the coupling strength. A striking characteristic is that pairs with the highest coupling strength are located in the top-right corner of the heat maps, whereas the lowest coupling pairs appear at the bottom, despite the occurrence of an abnormal technological knowledge network in 2001.

Another remarkable feature is that the knowledge networks of S&T exhibit extremely low coupling strength before 2007 under the Top-10% terms strategy. By conducting a literature review and consulting domain experts, we found that scientific studies in the field of energy conservation covered a wide variety of research topics, ranging from energy consumption modeling and estimation (Al-Shehri, 2000), energy conservation policy (Viklund, 2004), and clean production (Momirlan & Veziroglu, 2005) to the design of various high-efficiency devices (Liu, Bai, Li, & Wei, 2017; Zhang, Qin, Buchholz, & Passerini, 2018). There was also a large number of scientists from diverse disciplines involved in the field, including physics, engineering, environmental science, economics, and so forth. Due to their practical nature, the topics of patents were concentrated on novel systems, devices, and materials that promote energy efficiency. For example, protective coatings for electrodes and a power supply system for electric vehicles. These topics had not become the major concerns in the academic until later years, therefore, the scientific knowledge networks in Top-10% terms strategy before 2007 loosely coupled with the technological knowledge networks.

It should be considered that different network construction strategies yield different results. Since the Top-10% strategy highlights the core concepts in S&T, it produces the lowest average coupling strength. Under this strategy, the highest value for coupling strength is 0.67, based on a comparison of the 2019 scientific knowledge network with the 2017 technological knowledge network. During these two years, scientific studies and patents shared a common interest in electronic vehicles. While scientists paid attention to energy management strategies, energy storage systems, and autopilot algorithms, patent inventors mainly focused on the technical components of vehicles, especially batteries and motors. The strategy of Top-10% + shared terms shows the highest average coupling

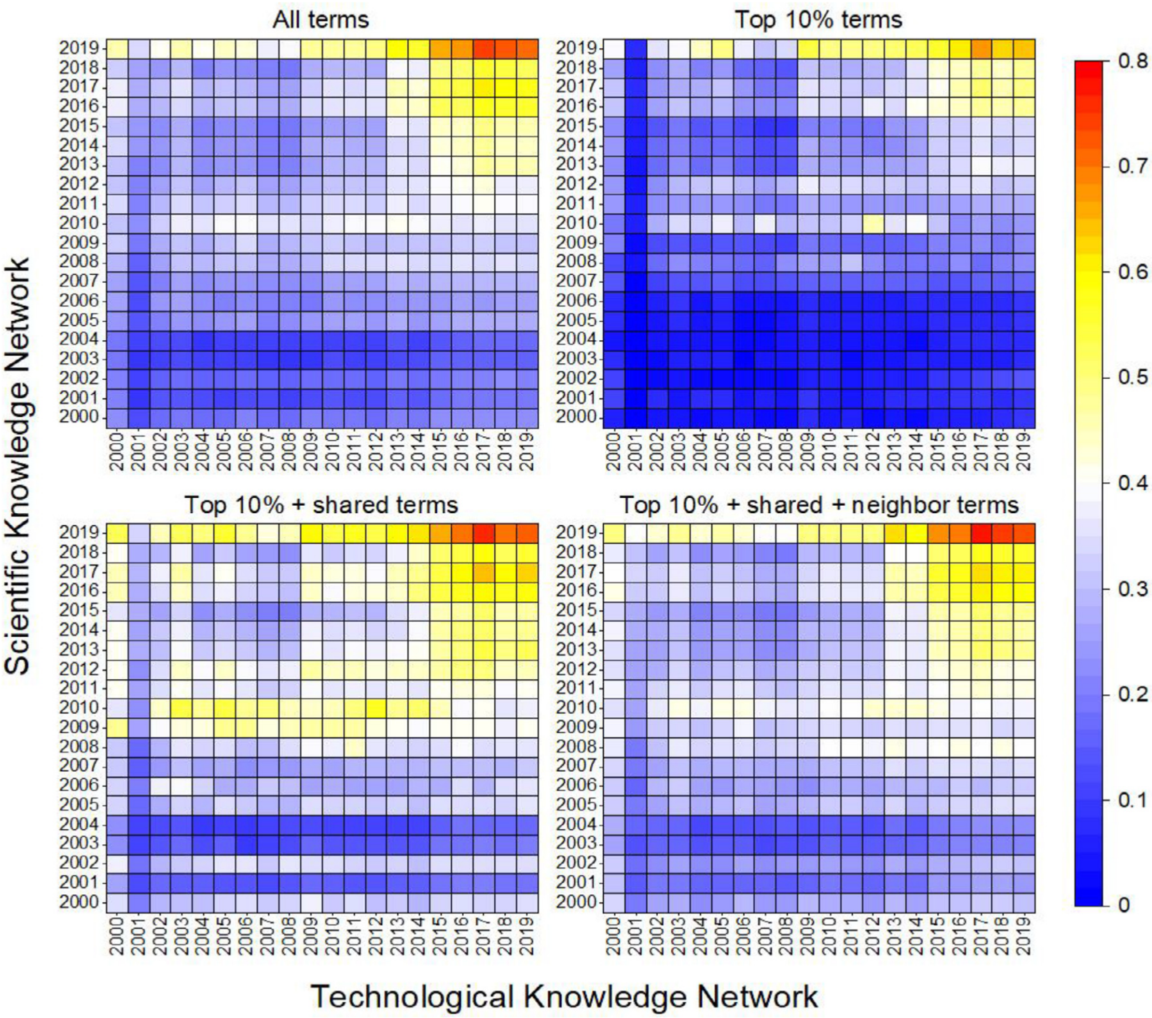

Fig. 7. The coupling strength of the S&T knowledge networks under the four network construction strategies.

strength. It is worth noting that the strategy discovered a horizontal area in the middle of the heat map, which exhibits much higher coupling strength than the cells nearby. This horizontal area corresponds to the scientific knowledge networks in 2009 and 2010 and a wide range of technological knowledge networks. Such a pattern is less prominent, or barely noticeable, under the other three strategies. While battery technology was not one of the most prominent scientific research topics (Top-10% terms) in 2009 and 2010, it laid the foundation for the development of energy management and storage strategies. Scientific papers mentioned the underlying battery technology of their research in the abstracts and this was captured by the Top-10% + shared terms network construction strategy. Since battery technology remained a central topic of patents in most years, they demonstrated a relatively strong coupling strength with scientific knowledge networks containing battery knowledge in this strategy.

To investigate more in-depth how the S&T knowledge networks were associated, we further compared the coupling strength between them from three perspectives: (1) comparing S&T knowledge networks in the same year, (2) over the preceding three years, (3) and over the subsequent three years. Fig.8 shows these comparisons under the four network construction strategies.

It can be seen that the black lines (the coupling strength between S&T networks in the same year) in the main graphs increased over time, with some fluctuations. This pattern is found for every construction strategy, suggesting that the coupling strength of S&T networks, in terms of knowledge linkages and structural linkages, gradually increases over time. This finding is consistent with the growing need for S&T integration to facilitate technological innovation in the energy conservation field (Croucher, 2011). In all of the cases, there is a point where the blue line rises above the other lines and continues for a period of time (see the highlighted areas in Fig. 8), which suggests that for the succeeding-three-year comparisons during that span of time in each graph, the scientific knowledge network had a stronger coupling strength with the technological knowledge network than for the same-year or preceding-three-year comparisons. This implies that a science-leading relationship may have existed between S&T during the highlighted time ranges. That is, knowledge and structural linkages of the scientific knowledge network may be found later in the succeeding technological knowledge networks. Under all of the construction strategies, the gap between blue and red lines first

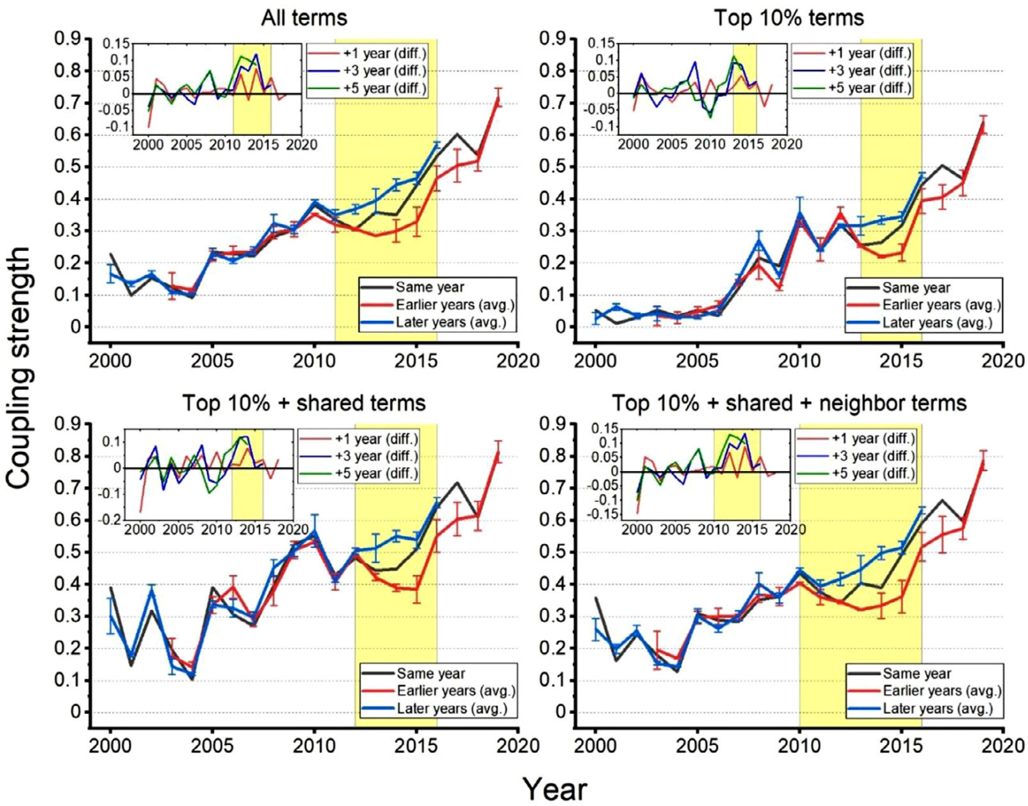

Fig. 8. Coupling strength between S&T knowledge networks in the same year, earlier years, and later years. Black lines represent same-year comparisons and are essentially the values that lie in the antidiagonals of the four heat maps in Fig. 7. Blue lines denote the average coupling strength between the scientific knowledge network in a given year and the technological knowledge networks over the preceding three years, which consists of the average of the three values to the left of the antidiagonal. Similarly, red lines represent the average of the three values to the right of the antidiagonal. The time ranges when blue lines rise above red lines are highlighted in yellow.

enlarges and then shrinks. However, the duration of the science network's lead varies across the strategies, with the longest under the Top-10% + shared + neighbor terms strategy (2010–2016) and the shortest under the Top-10% terms strategy (2013–2016). The inset plots detail the coupling strength between scientific networks and the technological networks in one (red), three (blue), and five (green) years later, subtracting the coupling strength between S&T networks in the same year. Before the highlighted time range, all of the lines fluctuate around zero, and no continuous science-leading pattern is seen. The blue and green lines are consistently above zero and reach a peak in the highlighted area, suggesting that during this time the scientific knowledge network resembles the technological knowledge network three and five years in its future. According to a report from the United Nations, scientific advancements in third-generation solar cells (e.g., perovskite solar cell) have greatly improved the photoelectric efficiencies between 2012 and 2015, while reducing manufacturing costs (United Nations Conference on Trade & Development, 2019). This subsequently enables a larger scale of renewable energy deployment and stimulates related patent output. Other advancements in material and biochemistry science, e.g., printable organic solar cells and artificial photosynthesis, also promote renewable energy's share in the total worldwide energy supply through the transformation of scientific achievements. Therefore, the proposed approach may serve as an effective and quantitative method to detect the lead-lag relationship between S&T.

4.4. Comparing the results of different coupling measurements

Since the proposed approach to gauging the network coupling effect takes into consideration both knowledge content (node/keyword correspondences) and structural information (the weight of nodes and edges), it is beneficial to compare our approach to others that consider only one aspect of the two. For a content-based approach, we chose the Jaccard Index, as it is based only on the ratio of common nodes to all of the nodes in the S&T networks. For structure-based measurement, network portrait divergence (NPD) was employed (Bagrow & Bollt, 2019), which is a mathematically principled comparison approach that is applicable to all networks (Tantardini, Ieva, Tajoli, & Piccardi, 2019). It first constructs a portrait of each network, represented by a matrix M, where entry M_{lk}, $l = 0, 1, 2, \ldots, d$ (d is the diameter of the network), $k = 0, 1, 2, \ldots, N\text{-}1$ (N is the number of nodes in the network), is the number of nodes having k nodes at the shortest path distance l. The network portrait serves as an effective summary of the global topological characteristics of the network. Network portrait divergence is calculated by the Jensen-Shannon divergence over

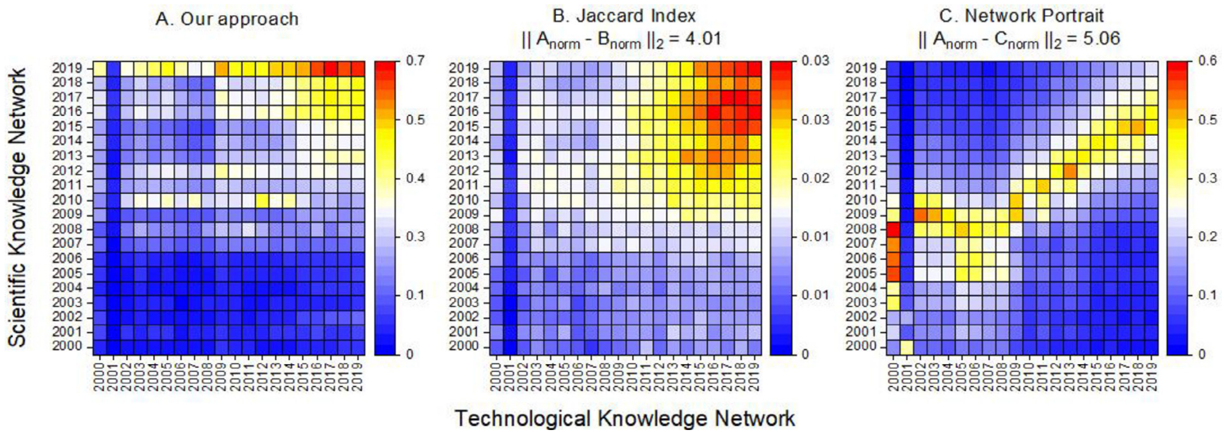

Fig. 9. Comparison of the coupling measurements of the Top-10% knowledge networks.

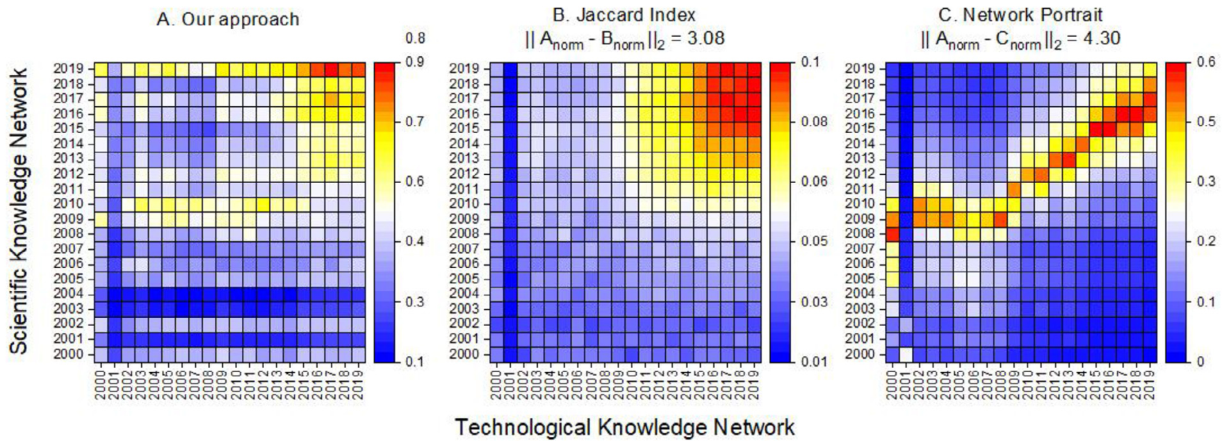

Fig. 10. Comparison of the coupling measurements of the Top-10% + shared knowledge networks.

the probability $p(k, l)$ (Bagrow & Bollt, 2019). Because this measurement relies only on a network's structural properties, comparing networks that have the same structure but different nodes result in an NPD of 0. We subtracted the NPD from one to convert it into a similarity measurement rather than a distance measurement.

Figs. 9 and 10 show the results for the three approaches in relation to the two types of network construction methods just defined. All terms and Top-10% + shared + neighbor terms strategies have not been included, due to the computational inefficiency of calculating NPD for large networks. Since we focused on the distribution patterns of the resulting matrices, we normalized each matrix through min-max normalization and calculated the L2 norm (Euclidean distance) between the proposed approach and the two other methods. The result shows that the content-based approach (i.e., the Jaccard Index) is more similar to our approach than the structure-based approach (i.e., network portrait divergence). Cells with a high coupling strength measured by Jaccard Index are also located in the upper-right area of the heat map, suggesting that S&T knowledge networks have experienced the most coupling in the later years. Unlike the results seen with the proposed approach, these highly coupling pairs are symmetrical about the antidiagonal. This symmetrical pattern becomes more prominent as the network size increases. For the structure-based approach (i.e., NPD), the distribution of coupling strength shows a distinct pattern. Most of the highly coupling pairs lie on the antidiagonal of the heat map except those between 2000 and 2005. This indicates that S&T knowledge networks in the same or close years had higher structural coupling strength. Combining the results of the three approaches reveals that the S&T knowledge networks are coupled the most in the later years, regarding both content and structure.

4.5. Comparing the results with randomly reshuffled networks

To validate the patterns found by the proposed coupling measurement method, we constructed random networks as null models by shuffling the edges of the original S&T knowledge networks (20 years × 4 strategies × 2 types = 160 networks). After that, the proposed approach was applied to the null models and it generated the coupling strengths shown in Fig.11. The Euclidean distances between the actual and randomized networks are presented in Table 3. The coupling strengths of the randomized networks are much

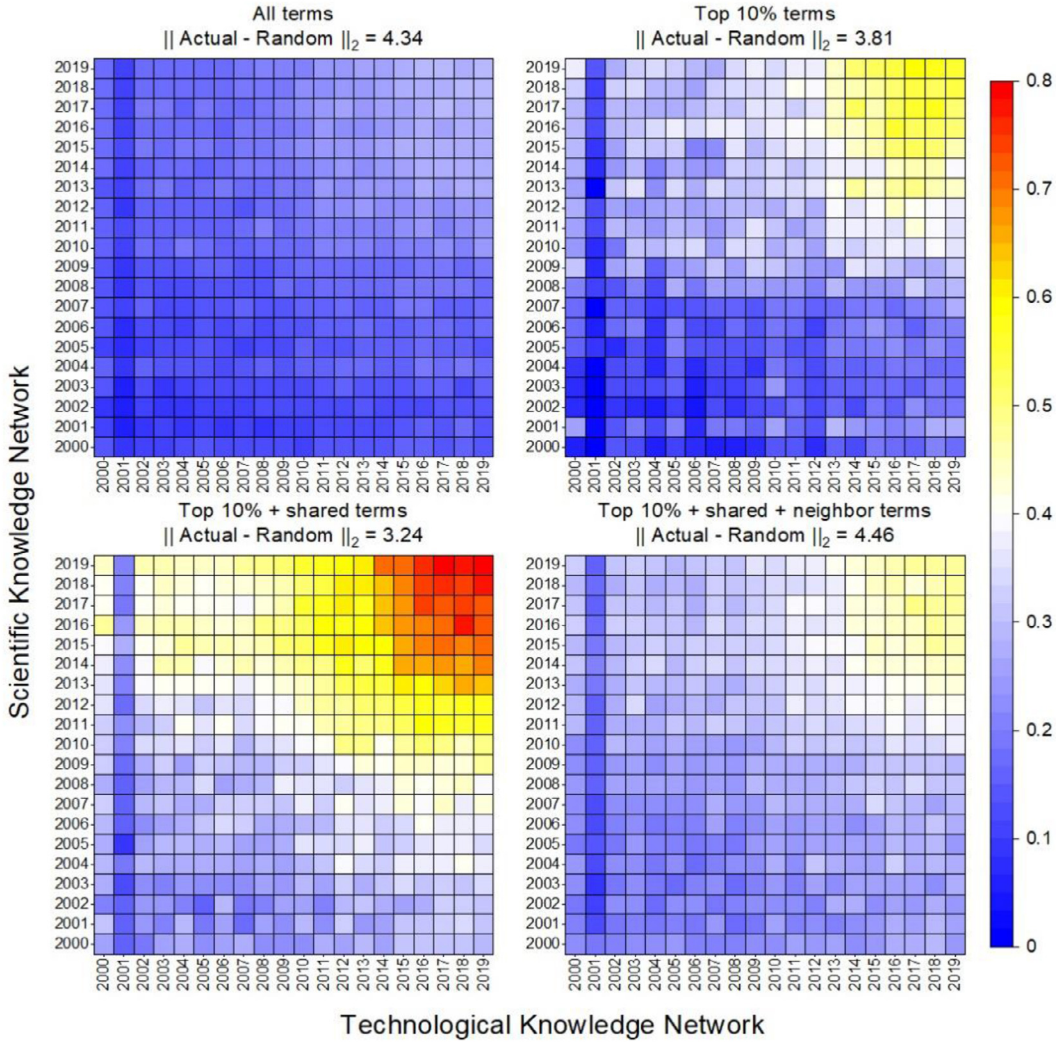

Fig. 11. The coupling strengths of randomly reshuffled S&T knowledge networks based on four network construction strategies.

Table 3
L2 norms (Euclidean distances) between the matrices of coupling strength calculated for the actual and randomly reshuffled S&T knowledge networks. The corresponding pairs of actual and random networks are highlighted in bold.

	Actual networks				Randomly reshuffled networks			
	All	T10	T10S	T10SN	All	T10	T10S	T10SN
Actual networks All terms	0	1.70	2.10	0.68	**4.34**	3.66	3.48	4.51
Top-10% terms	1.70	0	2.82	2.18	4.61	**3.81**	3.75	4.82
Top-10% + shared terms	2.10	2.82	0	1.90	3.73	3.18	**3.24**	3.73
Top-10% + shared + neighbor terms	0.68	2.18	1.90	0	4.35	3.75	3.56	**4.46**

Note: All, T10, T10S, and T10SN denote all terms, Top-10% terms, Top-10% + shared terms, and Top-10% terms + shared + neighbor terms, respectively.

lower than those of their actual counterparts. Moreover, the patterns found in the actual S&T knowledge networks are absent from the results of the random networks. For instance, the low coupling area at the bottom of the matrix using the Top-10% terms strategy and the horizontal high coupling area in the matrix with Top-10% + shared terms strategy visible in Fig. 7 are missing in Fig.11.

From a quantitative perspective, the distances between the matrices of the actual and random networks are relatively large (Table 3). They are larger than any distances calculated within the matrices of actual S&T knowledge networks, but similar to the distances between the proposed approach and the two other methods. This suggests that the patterns found in the S&T energy conservation knowledge networks are not random.

5. Discussion

There is an interplay and multifaceted relationship between S&T, which requires more sophisticated indicators and methods to capture the linkages. This study proposes a novel knowledge network coupling approach to measuring science-technology linkage by combining their knowledge linkages and structural linkages. To the best of our knowledge, it is one of the first studies to explore science-technology linkage from a network coupling perspective. Unlike prior research, which has focused on linear linkages between S&T, such as citation relations, author-inventor matching, and topic linkages, we contribute to the literature on S&T interactions by investigating network linkages between scientific knowledge systems and technological knowledge systems. An experimental study in the field of energy conservation confirms that our approach is indeed successful in revealing interactions between S&T.

5.1. Main findings

To construct the appropriate scientific and technological knowledge networks to explore their linkages, we used four term-combination strategies. Although the interaction patterns among the four model networks seemed to be in rough agreement, namely, in showing a gradually increasing coupling trend, the coupling performance for the four network models differed. The Top-10% strategy produced the lowest average coupling strength because it omitted a lot of the shared terms between S&T (Chen & Xiao, 2016). The number of these shared terms was far below the number of unique terms in the science or technology samples. Thus, the all-term strategy also omitted these shared terms and failed to clearly recognize the linkages between S&T, despite its holistic view of the scientific and technological knowledge systems. Of them, the Top-N + shared + neighbor terms construction strategy seems that it might be the most appropriate strategy for analyzing the interactions between S&T because it retains important knowledge, shared knowledge, and the associations between shared knowledge and knowledge unique to each network.

To verify the effectiveness of our network linkage approach, we compared it with approaches that consider only knowledge content or structural information. Content-based approaches, such as the Jaccard Index, gauge similarities between two knowledge networks based on the distribution of terms shared between them. However, they fail to recognize structural linkages. In contrast, structure-based approaches such as network portrait divergence (NPD) rely only on structural properties to measure similarities between networks. Especially, the similarity of two networks is 1 based on this approach when they hold the same structure but different nodes. Moreover, the distribution of coupling results based on such a one-sided approach is usually highly concentrated (Bagrow & Bollt, 2019). Therefore, it may be beneficial to measure science-technology linkage by taking into account both their knowledge linkages and structural linkages.

A striking characteristic of science-technology network linkage in the energy conservation field is that the coupling strength of the knowledge networks of S&T is gradually increasing over time. To investigate whether a lead-lag relation exists between them, we further compared the coupling strength between the S&T networks for a given year, as well as for earlier and later years. We found that the scientific knowledge network had a stronger coupling strength with the later technological knowledge networks than those in the same year or earlier years during a specific period of time. Specifically, scientific knowledge networks during that period were more strongly coupling with technological knowledge networks three or five years later than earlier ones. This implies that scientific developments precede technological advances. Knowledge linkages and structural linkages of the scientific knowledge network may be later found in the subsequent technological knowledge networks. Furthermore, the coupling patterns discovered in the actual networks differ from those in the randomized networks, which suggests that the proposed approach is robust and that the patterns revealed are not random.

5.2. Theoretical and practical implications

This study contributes to innovation research and practice in several ways. Theoretically, taking knowledge networks as proxies for scientific and technological knowledge systems, the novel approach to network coupling proposed in the paper can reveal interactions between S&T, which is conducive to quantifying knowledge linkages and structural linkages between S&T at the micro-content level. Unlike prior research, which has mainly investigated linear interactions between S&T, our approach captures complex S&T linkages from a different vantage point. The introduction of network coupling enriches the current methodology for measuring S&T linkages.

Also, our research offers a much more comprehensive understanding of interactions between S&T networks constructed with different strategies. It is difficult to see the rationale for the exploration of science-technology linkages and its effectiveness when only one type of scientific or technological knowledge network is involved. We performed an analysis using multiple knowledge networks with different designs to compare coupling patterns, which provides reference points that enable researchers to investigate dynamic coevolution and knowledge flows between networks.

Practically, the methodologies and findings of our study can serve as useful resources for policymakers. First, we provide a practical approach to measuring non-linear coupling relationships between S&T, which makes it possible to detect lead-lag relationships in a timely manner. The identified relationships can support corresponding policy adjustments to promote S&T transitions of innovation, which eventually enhances socioeconomic prosperity (Jones, 2011). For instance, special laws could be enacted to provide a faster patent-licensing process for scientific breakthroughs, as well as incentives for product manufacturing. Second, our findings suggest that the coupling strength of the S&T systems in the field of energy conservation is increasing. A science-leading period after 2010 was also identified. This implies the increasingly vital role of scientific advances in driving technological progress, which should encourage governments to emphasize science, technology, engineering, and mathematics (STEM) education and allocate appropriate resources (e.g., funding, labor, and policy changes). China is an example of success in this area. The Ministry of Science and Technology and the National Natural Science Foundation of China (NSFC) have established specialized funding projects for electric vehicles and related studies, fostering numerous scientific advances (Chen, Li, Yu, Chen, & Li, 2019; Xiao et al., 2020). Economic policies have also been introduced to compensate and reward firms that embrace scientific innovations. Consequently, China has become a leading country in renewable energy technology.

5.3. Limitations and future research

Admittedly, this study has several limitations, which suggest possibilities for future research. First, it is crucial to extract and select representative keywords when constructing knowledge networks to represent S&T. We plan to formulate some extraction rules for collecting these keywords from full-text information. To cope with differences in word expression between scientific publications and patents, we will include more effective algorithms for word alignment. Second, there are multi-level and multi-type relationships between keywords (i.e., co-word, citations, cocitation, and semantic), which may be equivalent or complementary. We will examine the imbalances, equivalences, and similarities of knowledge networks that are constructed based on different keyword links and further analyze the coupling characteristics of various S&T knowledge networks. It may be productive to adopt a multi-link fusion approach for unveiling associations between the two systems. Third, introducing suitable evaluation indexes or expert knowledge, we will compare the performance of network-based, citation-based, and topic-based approaches for measuring science-technology linkage, which can help us understand and optimize their utilization. Lastly, it is necessary to explore the coevolutionary characteristics of S&T knowledge network coupling using community identification and subnetwork decomposition methods. In sum, investigating the relationship between the coupling strength of S&T and innovation performance can be beneficial to uncovering the mutual effects of S&T.

6. Conclusion

The results from our study suggest that it is feasible to develop a novel network coupling approach to measure science-technology linkage. Not only do we describe a methodology beyond traditional linkage measurement, but we also provide detailed results for the significant case study involving energy conservation. It is shown that the coupling strength of S&T systems of energy conservation increases over time, with a specific period of time when a science-leading relationship existed. To summarize, this study makes three major contributions:

First, by turning the linear model into a network model, our study captures the interplay of the multifaceted relationship between S&T. Network linkages contain both correspondences between knowledge content and the correspondences between knowledge structures. Although knowledge linkages have been discovered to play a major bridging role in S&T, structural linkages between scientific knowledge networks and technological knowledge networks also supply abundant interaction information about the relationship between S&T. Thus, the integration of these two forms of linkage offers a more accurate measurement of science-technology linkage.

Second, calculating the coupling between S&T knowledge networks is key to revealing their interactions from a network perspective. Measuring the linkage between S&T based only on the number of coupling nodes and coupling edges shared between them may be inadequate due to a lack of consideration for the consistency of structural characteristics. Therefore, we computed the node-coupling strength of S&T knowledge networks based on the similarity of coupling nodes' degree distribution and the structural-coupling strength based on the similarity of coupling edges' weight distribution. These calculations can capture rich structural information about the two knowledge systems.

Third, appropriate S&T knowledge networks are examined to explore science-technology linkage. Using a variety of knowledge networks with different designs not only provided reference points for network construction but also verified the robustness of our proposed approach by making it possible to compare coupling performance based on different network models.

CRediT authorship contribution statement

Zhichao Ba: Conceptualization, Data curation, Formal analysis, Writing - review & editing. **Zhentao Liang:** Conceptualization, Data curation, Formal analysis, Writing - review & editing.

Acknowledgments

This work was jointly supported by the National Natural Science Foundation of China (Nos. 72004094, 71921002, and 71790612). In addition, we thank the three anonymous reviewers for valuable comments and suggestions which helped us to re-shape and improve this paper.

Appendix A

Table A.1.

Table A.1
Temporal network properties of S&T knowledge networks.

Year	Scientific knowledge network Nodes	Edges	Density	Comp.	Avg. C.C.	Technical knowledge network Nodes	Edges	Density	Comp.	Avg. C.C.
2000	4760	10,055	8.88E-04	599	0.96	5093	11,543	8.90E-04	441	0.93
2001	4827	10,423	8.95E-04	532	0.95	936	1848	4.22E-03	152	0.96
2002	5156	10,858	8.17E-04	657	0.95	5612	12,978	8.24E-04	435	0.92
2003	5495	11,832	7.84E-04	644	0.96	5362	12,415	8.64E-04	412	0.92
2004	5988	12,830	7.16E-04	722	0.96	5236	12,306	8.98E-04	381	0.92
2005	7026	15,416	6.25E-04	756	0.96	4392	10,159	1.05E-03	342	0.92
2006	7288	15,987	6.02E-04	781	0.96	4389	10,105	1.05E-03	360	0.92
2007	8023	17,521	5.44E-04	860	0.95	4196	9692	1.10E-03	327	0.92
2008	8742	19,375	5.07E-04	866	0.96	4636	10,875	1.01E-03	340	0.92
2009	9661	21,780	4.67E-04	902	0.95	6106	14,805	7.94E-04	395	0.91
2010	11,307	25,552	4.00E-04	1047	0.95	8455	21,570	6.04E-04	456	0.90
2011	13,446	30,851	3.41E-04	1129	0.95	8079	20,285	6.22E-04	460	0.91
2012	15,565	36,463	3.01E-04	1211	0.94	10,065	25,969	5.13E-04	521	0.90
2013	18,366	43,695	2.59E-04	1343	0.94	10,461	26,641	4.87E-04	583	0.90
2014	20,909	50,387	2.31E-04	1419	0.94	11,867	31,466	4.47E-04	598	0.90
2015	24,564	59,956	1.99E-04	1577	0.94	13,873	37,984	3.95E-04	597	0.89
2016	28,523	70,812	1.74E-04	1709	0.94	15,221	41,933	3.62E-04	600	0.88
2017	31,270	78,572	1.61E-04	1742	0.94	16,928	47,301	3.30E-04	697	0.89
2018	34,969	89,001	1.46E-04	1919	0.94	16,567	46,250	3.37E-04	657	0.88
2019	39,054	101,247	1.33E-04	1979	0.93	18,721	53,844	3.07E-04	669	0.88

Note. Comp denotes the number of components and CC denotes the clustering coefficient.

References

Ahlgren, P., Chen, Y., Colliander, C., & van Eck, N. J. (2020). Enhancing direct citations: A comparison of relatedness measures for community detection in a large set of PubMed publications. *Quantitative Science Studies, 1*(2), 714–729.

Ahmadpoor, M., & Jones, B. F. (2017). The dual frontier: Patented inventions and prior scientific advance. *Science (New York, N.Y.), 357*(6351), 583–587.

Albert, T. (2016). *Measuring technology maturity: Operationalizing information from patents, scientific publications, and the web*. Wiesbaden, Germany: Springer Gabler.

Al-Shehri, A. (2000). A simple forecasting model for industrial electric energy consumption. *International Journal of Energy Research, 24*(8), 719–726.

Bagrow, J. P., & Bollt, E. M. (2019). An information-theoretic, all-scales approach to comparing networks. *Applied Network Science, 4*(1), 1–15.

Barrat, A., Barthélemy, Marc, & Vespignani, A. (2004). Weighted evolving networks: Coupling topology and weights dynamics. *Physical Review Letters, 92*(22), Article 228701.

Bashan, A., Berezin, Y., Buldyrev, S. V., & Havlin, S. (2013). The extreme vulnerability of interdependent spatially embedded networks. *Nature Physics, 9*(10), 667–672.

Bassecouolard, E., & Zitt, M. (2004). Patents and publications: The lexical connection. In H. F. Moed, W. Glänzel, & U. Schoch (Eds.), *Handbook of quantitative science and technology research: The use of publication and patent statistics in studies of S&T systems* (pp. 665–694).

Behrouzi, S., Shafaeipour Sarmoor, Z., Hajsadeghi, K., & Kavousi, K. (2020). Predicting scientific research trends based on link prediction in keyword networks. *Journal of Informetrics, 14*(4), Article 101079.

Ben-David, A. (2008). Comparison of classification accuracy using Cohen's weighted kappa. *Expert Systems with Applications, 34*(2), 825–832.

Boyack, K. W., & Klavans, R. (2008). Measuring science-technology interaction using rare inventor-author names. *Journal of Informetrics, 2*(3), 173–182.

Boyack, K. W., & Klavans, R. (2020). A comparison of large-scale science models based on textual, direct citation and hybrid relatedness. *Quantitative Science Studies*, 1–19.

Boyack, K. W., Klavans, R., Small, H., & Ungar, L. (2014). Characterizing the emergence of two nanotechnology topics using a contemporaneous global micro-model of science. *Journal of Engineering and Technology Management, 32*(2), 147–159.

Boyack, K. W., & Klavans, R. (2010). Cocitation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology, 61*(12), 2389–2404.

Breschi, S., & Catalini, C. (2010). Tracing the links between science and technology: An exploratory analysis of scientists' and inventors' networks. *Research Policy, 39*(1), 14–26.

Bu, Y., Ni, S., & Huang, W. B. (2017). Combining multiple scholarly relationships with author cocitation analysis: A preliminary exploration on improving knowledge domain mappings. *Journal of Informetrics, 11*(3), 810–822.

Cassiman, B., Glenisson, P., & Looy, B. V. (2007). Measuring industry-science links through inventor-author relations: A profiling methodology. *Scientometrics, 70*(2), 379–391.

Chen, G., & Xiao, L. (2016). Selecting publication keywords for domain analysis in bibliometrics: A comparison of three methods. *Journal of Informetrics, 10*(1), 212–223.

Chen, R., Li, Q., Yu, X., Chen, L., & Li, H. (2019). Approaching practically accessible solid-state batteries: Stability issues related to solid electrolytes and interfaces. *Chemical Reviews, 120*(14), 6820–6877.

Criscuolo, P., & Verspagen, B. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Research Policy, 37*(10), 1892–1908.

Croucher, M. (2011). Potential problems and limitations of energy conservation and energy efficiency. *Energy Policy, 39*(10), 5795–5799.

Ding, Y., Chowdhury, G. G., & Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis. *Information Processing & Management, 37*(6), 817–842.

Du, J., Li, P. X., Guo, Q. Y., & Tang, X. L. (2019). Measuring the knowledge translation and convergence in pharmaceutical innovation by funding-science-technology-innovation linkages analysis. *Journal of Informetrics, 13*(1), 132–148.

Emmerich, T., Bunde, A., & Havlin, S. (2014). Structural and functional properties of spatially embedded scale-free networks. *Physical Review E, 89*(6), Article 062806.

Forti, E., Sobrero, M., & Franzoni, C. (2007). The effect of patenting on the networks and connections of academic scientists. In *11th International conference of the international society for scientometrics and informetrics* (pp. 272–284).

Gao, J. P., Ding, K., Teng, L., & Pang, J. (2012). Hybrid documents co-citation analysis: Making sense of the interaction between science and technology in technology diffusion. *Scientometrics, 93*(2), 459–471.

Glänzel, W., & Meyer, M. (2003). Patents cited in the scientific literature: An exploratory study of 'reverse' citation relations. *Scientometrics, 58*(2), 415–428.

Glänzel, W., & Thijs, B. (2017). Using hybrid methods and 'core documents' for the representation of clusters and topics: The astronomy dataset. *Scientometrics, 111*(2), 1071–1087.

Guan, J., & 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.

Han, F., & Magee, C. L. (2018). Testing the science/technology relationship by analysis of patent citations of scientific papers after decomposition of both science and technology. *Scientometrics, 116*(2), 767–796.

Haunschild, R., Leydesdorff, L., Bornmann, L., Hellsten, I., & Marx, W. (2019). Does the public discuss other topics on climate change than researchers? A comparison of explorative networks based on author keywords and hashtags. *Journal of Informetrics, 13*(2), 695–707.

Hu, X. J., & Rousseau, R. (2018). A new approach to explore the knowledge transition path in the evolution of science & technology: From the biology of restriction enzymes to their application in biotechnology. *Journal of Informetrics, 12*(3), 842–857.

Hu, Y., Zhou, D., Zhang, R., Han, Z., Rozenblat, C., & Havlin, S. (2013). Percolation of interdependent networks with intersimilarity. *Physical Review E, 88*(5), Article 052805.

Huang, M. H., Yang, H. W., & Chen, D. Z. (2015). Increasing science and technology linkage in fuel cells: A cross citation analysis of papers and patents. *Journal of Informetrics, 9*(2), 237–249.

Huang, X., Shao, S., Wang, H., Buldyrev, S. V., Stanley, H. E., & Havlin, S. (2013). The robustness of interdependent clustered networks. *EPL (Europhysics Letters), 101*(1), 18002.

Jones, B. F. (2011). As science evolves, how can science policy? *Innovation Policy and the Economy, 11*(1), 103–131.

Ke, Q. (2020). An analysis of the evolution of science-technology linkage in biomedicine. *Journal of Informetrics, 14*(4), Article 101074.

Kleminski, R., Kazienko, P., & Kajdanowicz, T. (2020). Analysis of direct citation, co-citation and bibliographic coupling in scientific topic identification. *Journal of Information Science*, Article 0165551520962775.

Kwon, S., Liu, X., Porter, A. L., & Youtie, J. (2019). Research addressing emerging technological ideas has greater scientific impact. *Research Policy, 48*(9), 1–16.

Larsen, K. (2008). Knowledge network hubs and measures of research impact, science structure, and publication output in nanostructured solar cell research. *Scientometrics, 74*(1), 123–142.

Lee, P. C., Su, H. N., & Chan, T. Y. (2010). Assessment of ontology-based knowledge network formation by vector-space model. *Scientometrics, 85*(3), 689–703.

Leydesdoref, L., & Meyer, M. (2007). The scientometrics of a triple helix of university-industry-government relations (introduction to the topical issue). *Scientometrics, 70*(2), 207–222.

Li, D., Azoulay, P., & Sampat, B. N. (2017). The applied value of public investments in biomedical research. *Science (New York, N.Y.), 356*(6333), 78–81.

Li, K., & Yan, E. (2019). Are NIH-funded publications fulfilling the proposed research? An examination of concept-matchedness between NIH research grants and their supported publications. *Journal of Informetrics, 13*(1), 226–237.

Liu, B., Bai, L., Li, T., Wei, C., et al. (2017). High efficiency and high open-circuit voltage quadruple-junction silicon thin-film solar cells for future electronic applications. *Energy & Environmental Science, 10*(5), 1134–1141.

Liu, R. R., Li, M., & Jia, C. X. (2016). Cascading failures in coupled networks: The critical role of node-coupling strength across networks. *Scientific Reports, 6*(1). 10.1038/srep35352.

Ma'ayan, A. (2011). Introduction to network analysis in systems biology. *Science Signaling, 4*(190) tr5-tr5.

Meyer, M. (2006). Are patenting scientists the better scholars? An exploratory comparison of inventor-authors with their non-inventing peers in nano-science and technology. *Research Policy, 35*(10), 1646–1662.

Momirlan, M., & Veziroglu, T. N. (2005). The properties of hydrogen as fuel tomorrow in sustainable energy system for a cleaner planet. *International Journal of Hydrogen Energy, 30*(7), 795–802.

Mukherjee, S., Romero, D. M., Jones, B., & Uzzi, B. (2017). The nearly universal link between the age of past knowledge and tomorrow's breakthroughs in science and technology: The hotspot. *Science Advances, 3*(4), Article e1601315.

Murray, F. (2002). Innovation as co-evolution of scientific and technological networks: Exploring tissue engineering. *Research Policy, 31*, 1389–1403.

Nguyen, A. L., Liu, W., Khor, K. A., Nanetti, A., & Cheong, S. A. (2019). The golden eras of graphene science and technology: Bibliographic evidences from journal and patent publications. *Journal of Informetrics, 14*(4), Article 101067.

Noyons, E. C. M., Raan, A. F. J. V., Grupp, H., & Schmoch, U. (1994). Exploring the science and technology interface: Inventor-author relations in laser medicine research. *Research Policy, 23*(4), 443–457.

Parshani, R., Rozenblat, C., Ietri, D., Ducruet, C., & Havlin, S. (2010). Inter-similarity between coupled networks. *Europhysics Letters, 92*(6), 68002.

Perianes-Rodriguez, A., Waltman, L., & Van Eck, N. J. (2016). Constructing bibliometric networks: A comparison between full and fractional counting. *Journal of Informetrics, 10*(4), 1178–1195.

Peterman, A., Kourula, A., & Levitt, R. (2014). Balancing act: Government roles in an energy conservation network. *Research Policy, 43*(6), 1067–1082.

Popp, D. (2017). From science to technology: The value of knowledge from different energy research institutions. *Research Policy, 46*(9), 1580–1594.

Ranaei, S., Suominen, A., & Dedehayir, O. (2017). A topic model analysis of science and technology linkages: A case study in pharmaceutical industry. *2017 IEEE technology & engineering management conference*.

Roach, M., & Cohen, W. M. (2013). Lens or prism? Patent citations as a measure of knowledge flows from public research. *Management Science, 59*, 504–525.

Rose, S., Engel, D., Cramer, N., & Cowley, W. (2010). Automatic keyword extraction from individual documents. In *Text mining: Applications and theory* (pp. 1–20).

Shao, J., Buldyrev, S. V., Havlin, S., & Stanley, H. E. (2011). Cascade of failures in coupled network systems with multiple support-dependence relations. *Physical Review E, 83*(3), Article 036116.

Shibata, N., Kajikawa, Y., & Sakata, I. (2010). Extracting the commercialization gap between science and technology - Case study of a solar cell. *Technological Forecasting and Social Change, 77*(7), 1147–1155.

Tantardini, M., Ieva, F., Tajoli, L., & Piccardi, C. (2019). Comparing methods for comparing networks. *Scientific Reports, 9*(1), 1–19.

United Nations Conference on Trade and Development. (2019). The role of science, technology and innovation in promoting renewable energy by 2030. Retrieved from UNCTAD website: https://unctad.org/webflyer/role-science-technology-and-innovation-promoting-renewable-energy-2030

Valdez, L. D., Macri, P. A., Stanley, H. E., & Braunstein, L. A. (2013). Triple point in correlated interdependent networks. *Physical Review E, 88*(5), Article 050803.

Verbeek, A., Debackere, K., & Luwel, M. (2002). Linking science to technology: Using bibliographic references in patents to build linkage schemes. *Scientometrics, 54*(3), 399–420.

Verhoeven, D., Bakker, J., & Veugelers, R. (2016). Measuring technological novelty with patent-based indicators. *Research Policy, 45*(3), 707–723.

Viklund, M. (2004). Energy policy options-from the perspective of public attitudes and risk perceptions. *Energy Policy, 32*(10), 1159–1171.

Wang, F., Jia, C., Wang, X., Liu, J., et al. (2019). Exploring all-author tripartite citation networks: A case study of gene editing. *Journal of Informetrics, 13*(3), 856–873.

Wang, G., & Guan, J. (2011). Measuring science-technology interactions using patent citations and author-inventor links: An exploration analysis from Chinese nanotechnology. *Journal of Nanoparticle Research, 13*(12), 6245–6262.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*, 440–442.

Xiao, Y., Wang, Y., Bo, S. H., Kim, J. C., Miara, L. J., & Ceder, G. (2020). Understanding interface stability in solid-state batteries. *Nature Reviews Materials, 5*(2), 105–126.

Xu, H. Y., Winnink, J., Yue, Z., Liu, Z., & Yuan, G. (2020). Topic-linked innovation paths in science and technology. *Journal of Informetrics, 14*(2), Article 101014.

Xu, S., Zhai, D., Wang, F., An, X., Pang, H., & Sun, Y. (2019). A novel method for topic linkages between scientific publications and patents. *Journal of the Association for Information Science and Technology, 70*(9), 1–17.

Yan, E., & Ding, Y. (2012). Scholarly network similarities: How bibliographic coupling networks, citation networks, cocitation networks, topical networks, coauthorship networks, and coword networks relate to each other. *Journal of the American Society for Information Science and Technology, 63*(7), 1313–1326.

Yi, S., & Choi, J. (2012). The organization of scientific knowledge: The structural characteristics of keyword networks. *Scientometrics, 90*(3), 1015–1026.

Zhang, G., Liu, L., & Wei, F. (2019). Key nodes mining in the inventor-author knowledge diffusion network. *Scientometrics, 118*(3), 728–735.

Zhang, H., Qin, B., Buchholz, D., & Passerini, S. (2018). High-efficiency sodium-ion battery based on NASICON electrodes with high power and long lifespan. *ACS Applied Energy Materials, 1*(11), 6425–6432.