J Technol Transf (2018) 43:1291–1317
https://doi.org/10.1007/s10961-017-9561-4

Identifying core topics in technology and innovation management studies: a topic model approach

Hakyeon Lee[1] · **Pilsung Kang**[2]

Published online: 11 February 2017

Abstract The study of technology and innovation management (TIM) has continued to evolve and expand with great speed over the last three decades. This research aims to identify core topics in TIM studies and explore their dynamic changes. The conventional approach, based on discrete assignments by subjective judgment with predetermined categories, cannot effectively capture latent topics from large volumes of scholarly data. Hence, this study adopts the topic model approach, which automatically discovers topics that pervade a large and unstructured collection of documents, to uncover research topics in TIM research. The 50 topics of TIM research are identified through the Latent Dirichlet Allocation model from 11,693 articles published from 1997 to 2016 in 11 TIM journals, and top 10 most popular topics in TIM research are briefly reviewed. We then explore topic trends by examining the changes in topics rankings over different time periods and identifying hot and cold topics of TIM research over the last two decades. For each of the 11 TIM journals, the areas of subspecialty and the effects of editor changes on topic portfolios are also investigated. The findings of this study are expected to provide implications for researchers, journal editors, and policy makers in the field of TIM.

Keywords Technology and innovation management · Research topic · Topic model · Latent Dirichlet Allocation (LDA)

JEL Classification O30 · O322

Electronic supplementary material The online version of this article (doi:10.1007/s10961-017-9561-4) contains supplementary material, which is available to authorized users.

✉ Pilsung Kang
pilsung_kang@korea.ac.kr

1 Department of Industrial and Systems Engineering, Seoul National University of Science and Technology, 232 Gongneung-ro, Nowon-gu, Seoul 139-743, Republic of Korea

2 School of Industrial Management Engineering, Korea University, 145 Anam-ro, Seongbuk-gu 136-713, Republic of Korea

1 Introduction

Over the last three decades, technology and innovation management (TIM) has grown with great speed, and become a self-sustained academic discipline (Pilkington and Teichert 2006). TIM literature has expanded significantly with over a dozen academic journals devoted to this field. The considerable growth of TIM research has boosted bibliometric studies on TIM itself because it is common practice for scholars to turn their attention toward the literature itself once an academic discipline has reached a certain degree of maturity (Ramos-Rodríguez and Ruíz-Navarro 2004).

Studies of TIM literature have been conducted in various directions. Several attempts have been made to identify intellectual pillars of TIM including influential works, journals, and authors by counting the number of publications and citations in a single journal (Biemans et al. 2007; Durisin et al. 2010; Pilkington and Teichert 2006) or a group of journals (Ball and Rigby 2006; Beyhan and Cetindamar 2011). The intellectual relationships in TIM have often been explored through co-citation and/or citation analysis, and visualized as networks of various academic entities such as articles (Beyhan and Cetindamar 2011; Pilkington and Teichert 2006), authors (Beyhan and Cetindamar 2011; McMillan 2008), and journals (Lee 2015; McMillan 2008). There have also been a series of attempts to rank TIM journals based on citation analysis (Cheng et al. 1999; Linton and Thongpapanl 2004; Thongpapanl 2012).

Another important research stream in bibliometric studies of TIM literature is to identify core topics. Finding core topics and tracing their evolution in a scientific discipline has been of great interest to government, industry, and academia (Small et al. 2014). Which research topics are rising or falling in popularity is a useful piece of information for governmental funding boards for grants allocation to promising areas, companies formulating R&D projects for promising technologies, and researchers hoping to work on promising topics. In response, there have been several efforts to analyze research topics in TIM. Some studies investigated the core topics of a single TIM journal, such as *IEEE Transactions on Engineering Management* (Allen and Sosa 2004), *Technovation* (Merino et al. 2006), and *Research Policy* (Teichert and Pilkington 2006). Recent studies have broadened the coverage to multiple journals of TIM to deal with the field in its entirety. Cetindamar et al. (2009) identified and compared common topics of TIM research in developed and developing countries based on 325 articles published in the top 10 TIM journals. Out of the pre-determined 22 topics, 17 common topics were revealed to have been studied in at least 10 articles each, with the top 5 topics accounting for nearly half of all TIM studies. Choi et al. (2012) explored the national characteristics and differences in terms of TIM research topics by analyzing 5239 papers across the 10 TIM journals. They measured and compared the relative research advantage profile of 15 countries based on their own defined 13 research domains of TIM.

These studies employ the conventional approach of discrete assignments in which each article is classified into a single category of predetermined topics based on subjective judgment. However, such discrete assignments cannot effectively capture latent topics from large scholarly data. Firstly, manual allocation based on reading abstracts or author keywords always entails the risk of classification error. It also costs too much time, particularly when the number of articles to be assigned is very high. Secondly, the pre-determined categories are by no means exhaustive; relatively new and emerging topics are likely to be ignored. Convergence topics are also difficult to handle. Thirdly, an article usually contains two or more topics. Multiple themes may be introduced together in an article. Research methods and application domains should also be regarded as topics in themselves. However, assigning an article to a single topic cannot incorporate such aspects of academic research.

The clustering approach using co-occurrence networks, which is the most widely used bibliometric approach to identifying research topics in other academic fields, may accommodate the first two limitations. The relationships between articles or keywords are measured by co-citation or co-word measures, and similar entities are clustered as a topic (Yan et al. 2012). However, when the data is too big, it is extremely difficult to define topics of clusters from thousands of nodes of a co-occurrence network. Furthermore, a node is usually clustered into a single topic, thus multiple topics contained in a single article cannot be captured by the conventional clustering approach.

The topic model approach, which has rapidly gained popularity in recent years, is a promising solution that may overcome these limitations. Topic models are algorithms to automatically discover the core topics that pervade a large and unstructured collection of documents (Blei 2012). These algorithms do not require any prior labeling of the documents; the topics emerge from the analysis of the original texts. Fractional assignments are made by assuming that documents exhibit multiple topics, modeled as probability distributions over terms. Given these advantages, recent years have seen an increased impetus to use the topic model approach to identify topics in a variety of academic domains such as biology (Wang et al. 2011), information retrieval (Yan et al. 2012), statistics (De Battisti et al. 2015), and hydropower (Jiang et al. 2016).

There has also been a recent study that identified research topics in TIM using topic modeling. Antons et al. (2016) employed the topic model approach to map the topic landscape of Journal of Product Innovation Management which is one of the major journals in the field of TIM, but it has the following limitations. Firstly, only the articles published in the single journal were used for deriving topics, thus the results cannot represent the whole field of TIM. Secondly, the study adopted the discrete assignment approach in which articles with higher than 10% loadings are assigned to each topic, which are not distinctive from the conventional approach. It did not exploit the advantage of topic modeling coming from the fractional assignment explained above.

This study identifies TIM research topics using topic modeling based on the articles published in 11 major journals in the field of TIM, and analyzes the topics based on the fractional assignment approach. We derive 50 topics of TIM research by applying the Latent Dirichlet Allocation (LDA) model, which is the most popular topic model, to 11,693 articles published during the last two decades (from January 1997 to June 2016) across the 11 TIM journals. We then explore topic trends by examining the changes in topics rankings over different time periods and identifying hot and cold topics of TIM research over the last two decades. For each of the 11 TIM journals, the areas of subspecialty and the effects of editor changes on topic portfolios are also investigated.

The rest of this paper is organized as follows: Sect. 2 reviews the topic model approach and its applications to identifying research topics. The methods and procedure for identifying topics are explained in Sect. 3. Section 4 lists the 50 topics of TIM research and provides brief reviews on top 10 most popular topics. Section 5 deals with TIM research trends, and Sect. 6 reports the core topics at the journal level. Finally, the conclusions are provided in Sect. 7, along with limitations and directions for future research.

2 Topic model approach to identifying research topics

Topic models are a set of statistical algorithms that uncover the main themes that permeate in a large collection of unstructured documents known as corpus (Blei 2012; Zhang 2012). They are considered "generative" models as they assume that there exists a certain probabilistic document generation process. Topic models also assume that (1) each document is a mixture of topics, and (2) each topic has its own probability distribution over words. Hence, the purpose of topic modeling algorithms is to estimate the parameters of this probabilistic document generation process, i.e., topic distribution per document and word distribution per topic, by observing the words used in actual documents in the corpus (Blei 2012).

Topic models have many advantages from both algorithmic and practical perspectives. First, they have principled mathematical foundations that help us understand the mechanism of document generation. Second, they do not require any prior labeling of documents so that documents can be analyzed without the help of human experts (Blei et al. 2003). Finally, they can organize and summarize documents automatically (Griffiths and Steyvers 2004). Due to these benefits, topic models have recently garnered significant attention and been successfully applied to a wide range of text mining tasks (Yan 2014).

LDA, proposed by Blei (2012), is the most widely adopted topic modeling algorithm and has some practical advantages over other topic modeling algorithms. It can estimate the mixture of existing topics in new documents without the need to update the current model. Moreover, since it has a fixed number of parameters irrespective of the corpus size, it can handle high volumes of documents (Blei et al. 2003). Hence, LDA can avoid the problem of overfitting, enabling it to be easily applied to text mining tasks in which massive volumes of text documents are constantly generated.

The basic idea of LDA is that documents are expressed as random mixtures over latent topics, each of which is characterized by a distribution over words. Figure 1 shows the document generation process of LDA with an illustrative example. In Fig. 1, $w_{d,i}$ is the ith word in the dth document and $z_{d,i}$ is the topic assignment to the word $w_{d,i}$. θ_d is the topic proportions for the dth document, whereas ϕ_k is the word distribution for the kth topic. α and β are the Dirichlet hyper-parameters for θ_d and ϕ_k, respectively. N, D, and K denote the number of words in a document, number of documents in a corpus, and number of topics across the corpus, respectively. Nodes are random variables and edges indicate dependence. In addition, only the shaded node is observable while the others are hidden or latent. Finally, plates indicate replicated variables. In LDA, the following process generates each word in a document. First, per-topic word distributions, i.e., the frequency of each word's usage within a topic, are estimated for the whole corpus. Then, for each document, a per-document topic proportion, which indicates the prevalence of each topic in the target document, is determined. Finally, each word is chosen by taking into account the assigned topic and its distribution over words.

Training of LDA is equivalent to inferring the latent variables, i.e., the per-topic word distributions ϕ_k, per-document topic proportions θ_d, and per-word topic assignment $z_{d,i}$. A well trained LDA model is one where the chosen words for a document are very similar to the words appearing in that document. Therefore, based on the observed words in each document, the LDA parameters are optimized to maximize the likelihood of those word appearances estimated by the model. Appendix A provides more details on the LDA training procedure and parameter estimation.

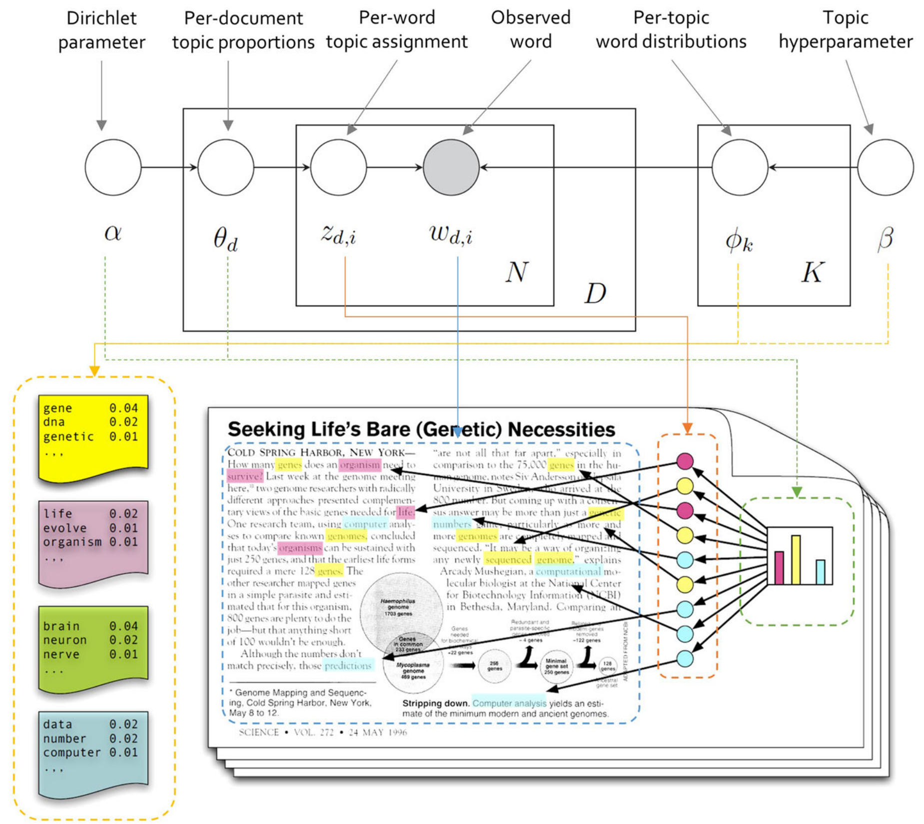

Fig. 1 Document generation process and two outputs of LDA. The illustrative example is taken from Blei et al. (2003) with permission

Once the LDA inference is completed for K topics, a D by K per-document topic proportion matrix $\boldsymbol{\theta}$ and a V by K per-topic word distribution matrix $\boldsymbol{\phi}$ are obtained as shown in Fig. 2. Note that the sum of each row is 1 for $\boldsymbol{\theta}$, whereas the sum of each column is 1 for $\boldsymbol{\phi}$. In addition, note that if $\boldsymbol{\theta}$ averages to a one-dimensional row vector, it can be understood as the corpus-level topic distributions. With these inferred distributions, subsequent text analysis tasks can be conducted; $\boldsymbol{\theta}$ can be used to reduce the dimensionality for document categorization or classification, whereas $\boldsymbol{\phi}$ can be used to exploit the relationship

(a)

	Topic 1	Topic 2	Topic 3	...	Topic K	Sum
Doc 1	0.20	0.50	0.10	...	0.10	1
Doc 2	0.50	0.02	0.01	...	0.40	1
Doc 3	0.05	0.12	0.48	...	0.15	1
...	...	...	...	...	...	1
Doc N	0.14	0.25	0.33	...	0.14	1

(b)

	Topic 1	Topic 2	Topic 3	...	Topic K
word 1	0.01	0.05	0.05	...	0.10
word 2	0.02	0.02	0.01	...	0.03
word 3	0.05	0.12	0.08	...	0.02
...	...	...	...	...	...
word V	0.04	0.01	0.03	...	0.07
Sum	1	1	1	1	1

Fig. 2 LDA output examples. **a** Per-document topic proportions (θ_d). **b** Per-topic word distributions (ϕ_k)

between latent topics, to trace the change in topic distributions over time, and to detect topics rising and falling in popularity (Griffiths and Steyvers 2004; Yan 2014).

Recently, there have been several attempts to take advantage of LDA to identify research topics in a specific academic field. De Battisti et al. (2015) identified the 30 most prominent topics in statistics based on 3060 papers published in three statistical journals, and investigated the evolutionary path of research topics in statistics. Yan et al. (2012) detected 10 topics in the field of information retrieval across 52,762 articles over three periods, and explored the topic dynamics based on topic similarities. Yan (2014) analyzed the research dynamics of the information science field by defining six types of topic continuity and three types of topic popularity based on 27,796 articles. In the same research field, Ding (2011) discovered five communities from a total of 12,146 articles by considering both topological and topical relationships. He et al. (2013) employed topic models to analyze an interdisciplinary collaboration in co-authorship network in computer science field based on the titles of 230,000 articles. Jiang et al. (2016) identified 29 topics of hydropower research using 5093 articles and clustered them in terms of term-level similarity and document-level similarity. Wang et al. (2011) developed the Bio-LDA algorithm to identify latent topics in biology. Using 336,899 articles published in PubMed, they extracted 50 topics of biology and explored bio-term relationships to identify new opportunities such as discovering potential drugs for diseases. Although topic modeling has not enjoyed many applications so far because it is a relatively new approach, the above studies demonstrated the usefulness and the validity of topic modeling in identifying research topics in an academic field. This study also adopts the LDA model for identifying latent topics in the field of TIM.

3 Methods

This section explains the methods and procedure for identifying topics in TIM research through LDA. The procedure comprises three steps: (1) target journal selection and data collection, (2) text preprocessing and LDA inference, and (3) topic labeling, as shown in Fig. 3.

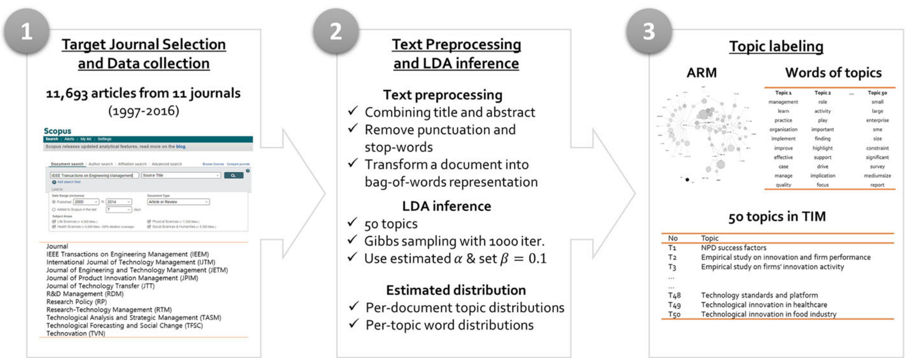

Fig. 3 Procedure for identifying topics

3.1 Target journal selection and data collection

The first step is to select relevant and leading TIM journals for article collection. Selecting leading TIM journals out of dozens of journals is complicated by the multidisciplinary nature of TIM, but recent studies dealing with TIM literature provide clues about journal selection. Building on the study of Cheng et al. (1999), Linton and Thongpapanl (2004) selected the following 10 TIM journals as base journals for ranking analysis: *IEEE Transactions on Engineering Management, International Journal of Technology Management, Journal of Engineering and Technology Management, Journal of Product Innovation Management, R&D Management, Research Policy, Research-Technology Management, Technological Analysis and Strategic Management, Technological Forecasting and Social Change,* and *Technovation.* Although there are some exceptions which include additional journals (Ball and Rigby 2006; Thongpapanl 2012), most subsequent studies also regarded the same set of ten journals as leading TIM journals (Beyhan and Cetindamar 2011; Cetindamar et al. 2009; Choi et al. 2012; Lee 2015). This study also relies on these ten journals, and considers an additional journal, *Journal of Technology Transfer*. Several recent studies acknowledged that this journal deserves to be considered as one of the leading TIM journals due to its long history, relatively high impact, and cohesion with other TIM journals (Thongpapanl 2012; Lee 2015). Table 1 lists the selected 11 journals and the respective number of articles included in this study. For each journal, the following process is conducted to collect the raw data for the targeted articles. First, we search documents that are original articles and review papers from the SCOPUS database using the source title for each journal from January 1997 to June 2016 as shown in Fig. 3. Then, for the selected articles, four meta-variables, namely source title, publication year, article title, and abstract, are selectively downloaded. In total, we collected 11,693 articles published in the 11 journals from 1997 to 2016.

3.2 Text preprocessing and LDA inference

Once the articles are collected, some pre-processing is required before conducting the LDA inference. Since an article's title contains the most representative words while its abstract

Table 1 List of TIM journals

Journal	Number of articles
IEEE Transactions on Engineering Management (IEEM)	837
International Journal of Technology Management (IJTM)	1433
Journal of Engineering and Technology Management (JETM)	351
Journal of Product Innovation Management (JPIM)	724
Journal of Technology Transfer (JTT)	597
R&D Management (RDM)	583
Research Policy (RP)	1999
Research-Technology Management (RTM)	732
Technology Analysis and Strategic Management (TASM)	813
Technological Forecasting and Social Change (TFSC)	2416
Technovation (TVN)	1208
Total	11,693

precisely summarizes the background, objective, methodology, and findings of the study, we merged the title and abstract of each article to treat them as a single document in the LDA. For each document, i.e., the merged title and abstract of each article, we removed all punctuation and numbers, and transformed all characters into the lower-case. Then, we eliminated all stop-words, whose main role is to make a sentence grammatically correct, e.g., articles (a, an, the) and prepositions (of, by, from, etc.), as they are not semantically meaningful. It is also acceptable to apply user-defined stop-words lists for analytical purposes. Since some general words appear in most articles, e.g., "study", "paper", "find", "effect", and "discuss", we constructed a list of additional stop-words and removed them from the corpus.

After stop-words removal, lemmatization was conducted to find the lemma of each word to reduce dimensionality (i.e., the total number of words in a corpus) without loss of generality. In a natural language, a word can be used in different forms to comply with grammar. For example, the word *innovate* may be used as either *innovate*, *innovates*, or *innovated* as a verb, or it can be used as either *innovation* or *innovations* as a noun, or *innovative* as an adjective. Stemming and lemmatization are two representative methods to find the root of a word. The main difference between them is that stemming reduces a word to its base form, whereas lemmatization reduces a word to its lemma. Since stemming finds any base form of a word, it does not have to be an existing word. For example, all the above six words related to *innovate* will be stemmed to *innovat*, which is the longest common string among them. Although stemming can more significantly reduce the dimensionality than lemmatization, it has a risk of losing part-of-speech information and having a non-word base form that may be difficult to interpret. Contrary to stemming, lemmatization finds the lemma that preserves both meaning and part-of-speech information. In the above example, *innovate*, *innovates*, and *innovated* will be lemmatized into *innovate*, *innovation* and *innovations* will be lemmatized into *innovation*, and *innovative* will be lemmatized into *innovative*. Since the interpretation of word distribution for the topics is one of the major research objectives in our study, lemmatization is more appropriate than stemming.

The last preprocessing step is to transform the documents into the bag-of-words representation. In the bag-of-words model, each document is represented in a vector of an unordered collection of words. In other words, if there are a total of V words in the corpus, then each document becomes a V-dimensional vector in which the value of an element is the frequency of the corresponding word in the document.

We used the R package named *"topicmodels"* to fit the LDA model (Hornik and Grün 2011). Table 2 shows the experimental settings used in our study. Two candidate inference algorithms are employed: variational expectation method (VEM) and Gibbs sampling (Gibbs). In order to determine the number of topics K, we vary it from 10 to 100 in stepped increases of 10, while four different iterations are tested for Gibbs sampling. We used the default values of the package for Dirichlet parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$; the former is estimated based on the corpus and the latter is set to 0.1 to encourage distinctive word distributions among the topics. When deciding the K and Gibbs iterations, a quantitative performance measure, such as perplexity, can be used from a modeling perspective. For the purpose of qualitative analysis, it is commonly determined by domain experts to have more interesting and interpretable topics (Andrzejewski et al. 2007). After thorough investigation into per-topic word distributions for different K values, we finally decided to use 50 topics identified after 1000 iterations of the Gibbs sampling.

Table 2 LDA experimental settings

Component	Candidates
Inference algorithm	Variational expectation method, collapsed Gibbs sampling
The number of topics K	10, 20, 30, 40, 50, 60, 70, 80, 90, 100
Gibbs sampling iteration	1000, 3000, 5000, 10,000
Dirichlet parameter α	Estimated from the corpus
Dirichlet parameter β	0.1 (default value of the LDA function in the topicmodels package

3.3 Topic labeling

One of the most challenging tasks in topic modeling is to name unlabeled topics. Topics are usually labeled based on their most frequent words, but designating a topic based on frequent words and their distributions is by no means straightforward. A unit constituting a topic is a word, not a term; a word itself may not be indicative of the topic. For more effective labeling of topics, we employed association rule mining (ARM), a data mining technique for discovering interesting relationships among variables in a large database, to identify the relations between words included in the same topic. Pairs of words with lift values greater than 1 were identified with their support and confidence values. In addition, abstracts of the 10 highest loading articles for each topic were extracted to capture what was actually studied under each topic.

Topic labeling was conducted by a group of five experts with more than ten years of academic experiences in the field of TIM. The three pieces of information, the lists of the most frequent words, the ARM results, and the abstracts of the 10 highest loading articles, were provided to the five experts prior to the workshop, and each expert was asked to label each topic individually in advance. The expert workshop was then held to collect individual labeling results. The expert group proposed identical or very similar labels for 39 topics, but 11 topics were labeled differently or not even labeled. An in-depth discussion was made for labeling the 11 topics, and the expert group reached a consensus on the labels of the nine topics. The remaining two topics whose most probable words are either not relevant to each other or too general to designate a specific label were named *miscellaneous.* These 50 topics serve as the backbone of the subsequent analysis in this study.

4 Topic landscape of TIM research

4.1 The 50 TIM topics

The extracted and labeled 50 topics of TIM research are shown in Table 3 with their five frequent and relevant words. Topics are numbered in a descending order of proportions in the whole collection of articles (i.e., fractional assignment), and the number of articles in which each topic has the highest proportion (i.e., discrete assignment) is also provided in Table 3. The difference between discrete and fractional assignments is clearly observed in Table 3. There exist significant differences in the topic rankings between the two types of assignments. For example, *[T2] Empirical study on innovation and firm performance* is ranked second in terms of proportions while its ranking in terms of the number of assigned

Table 3 Topics of TIM research

Topic	Frequent words (lemmas)	Share (Rank)			
		Proportion (%)		Number of articles	
[T1] NPD success factors	Product; development; NPD; customer; success	2.4046	(1)	455	(1)
[T2] Empirical study on innovation and firm performance	Performance; relationship; positive; hypothesis; empirical	2.3581	(2)	289	(17)
[T3] Empirical study on firms' innovation activity	Firm; innovative; evidence; empirical; sample	2.2648	(3)	311	(12)
[T4] Innovation diffusion modeling	Model; diffusion; forecast; curve; logistic	2.2216	(4)	306	(14)
[T5] Open innovation	Innovation; open; radical; intermediary; openness	2.2025	(5)	268	(18)
[T6] R&D globalization	Country; global; China; international; foreign	2.1871	(6)	409	(3)
[T7] Public R&D policy	Policy; public; government; support; program	2.1817	(7)	382	(4)
[T8] Technology planning and roadmapping	Approach; roadmapping; intelligence; planning; technology	2.1562	(8)	295	(16)
[T9] Technological forecasting and foresight	Forecast; future; scenario; foresight; Delphi	2.0837	(9)	350	(8)
[T10] Knowledge flow and transfer	Knowledge; external; source; internal; flow	2.0561	(10)	304	(15)
[T11] Technological transitions	Transition; sustainable; infrastructure; regime; society	2.0547	(11)	363	(7)
[T12] Energy policy and sustainability	Energy; climate; emission; policy; power	2.0509	(12)	342	(9)
[T13] Consumer behavior and satisfaction	Consumer; community; perceive; behavior; satisfaction	2.0506	(13)	307	(13)
[T14] Technology transfer and commercialization	University; transfer; commercialization; spinoff; entrepreneurship	2.0496	(14)	421	(2)
[T15] Technology advancement and fusion	Technology; development; emerge; advance; fusion	2.0486	(15)	151	(37)
[T16] Organizational learning	Organizational; process; learn; development; experience	2.0482	(16)	190	(32)
[T17] Technology development and economic growth	Growth; investment; economic; productivity; return	2.0466	(17)	243	(25)
[T18] Organizational culture and organizational innovation	Organization; organizational; manger; culture; leadership	2.0448	(18)	196	(29)
[T19] Technological collaboration and R&D alliances	Network; collaboration; alliance; partner; cooperation	2.0415	(19)	379	(5)
[T20] Social network analysis	Network; social; interaction; tie; externality	2.0396	(20)	133	(42)
[T21] Multiple criteria decision making	Measure; evaluation; assessment; criterion; selection	2.0332	(21)	246	(22)
[T22] NPD team management	Project; team; development; member; management	2.0279	(22)	327	(10)
[T23] Corporate management	Company; business; corporate; challenge; opportunity	2.0213	(23)	264	(19)

Table 3 continued

Topic	Frequent words (lemmas)	Share (Rank)			
		Proportion (%)		Number of articles	
[T24] Empirical study on innovation success factors	Factor; influence; success; variable; survey	2.0184	(24)	116	(45)
[T25] Academic research management	Research; scientific; basic; laboratory; academic	2.0142	(25)	258	(21)
[T26] Bibliometric analysis of scientific publications	Field; article; journal; author; bibliometric	1.9890	(26)	194	(30)
[T27] Patent analysis and intellectual property management	Patent; property; intellectual; invention; citation	1.9793	(27)	379	(5)
[T28] Product design	Product; design; concept; process; function	1.9792	(28)	194	(30)
[T29] Miscellaneous 1	Issue; question; address; challenge; attention	1.9789	(29)	103	(46)
[T30] Types and patterns of innovation	Innovation; type; difference, characteristic, pattern	1.9766	(30)	76	(49)
[T31] Strategic portfolio management	Strategic; portfolio; management; competence; core	1.9562	(31)	160	(36)
[T32] Technological change and evolution	Technological; change; evolution; path; trajectory	1.9516	(32)	140	(41)
[T33] Innovation cycle	Time; stage; cycle; generation; period	1.9457	(33)	129	(44)
[T34] Disruptive innovation	Market; competition; disruptive; entry; incumbent	1.9398	(34)	164	(35)
[T35] Innovation in high-tech industries	Industry; pharmaceutical; high-tech; semiconductor; electronics	1.9221	(35)	201	(28)
[T36] Regional innovation system and industrial clusters	Cluster; regional; local; park; location	1.9208	(36)	325	(11)
[T37] Systems of innovation approach	System; interaction; approach; complex; NIS	1.9170	(37)	149	(38)
[T38] R&D human resource development	Work; experience; employee; skill; professional	1.9156	(38)	245	(23)
[T39] Total quality management and continuous improvement	Quality; implementation; improvement; continuous; total	1.8885	(39)	202	(27)
[T40] Entrepreneurship and venture capital	Capital; venture; financial; entrepreneurial; startup	1.8885	(40)	264	(19)
[T41] Innovation capability and absorptive capacity	Role; capability; dynamic; capacity; absorptive	1.8827	(41)	81	(47)
[T42] Miscellaneous 2	Activity; mechanism; unit; variety; formal	1.8567	(42)	81	(47)
[T43] Supply chain management	Cost; supplier; supply; chain; manufacturer	1.8350	(43)	245	(23)
[T44] Technology adoption	Level; adoption; benefit; condition; intention	1.8341	(44)	48	(50)
[T45] ICT service management	Service; communication; information; mobile; Internet	1.8301	(45)	232	(26)
[T46] Innovation in SMEs	Manufacturing; small; enterprise; large; SME	1.8253	(46)	146	(40)
[T47] Software development and management	Problem; software; solution; develop; application	1.7993	(47)	130	(43)

Table 3 continued

Topic	Frequent words (lemmas)	Share (Rank)			
		Proportion (%)		Number of articles	
[T48] Technology standards and platform	Control; standard; component; platform; architecture	1.7884	(48)	175	(34)
[T49] Technological innovation in healthcare	People; age; health; medical; care	1.7778	(49)	177	(33)
[T50] Technological innovation in food industry	Production; good; material; food; agricultural	1.7153	(50)	148	(39)

NPD new product development, *NIS* national innovation system, *ICT* information and communication technology, *SME* small and medium enterprise

articles is placed at the middle (17th). This indicates that the relationship between innovation and firm performance has been empirically investigated in a number of articles, but the topic is not picked out in many cases because those articles also encompass other topics with higher proportions. The proportions of such general topics are likely to be underestimated by discrete assignment, which hinders the capture of the real distributions of topics. This study employs the proportions of topics as a measure of topic share for identifying hot and cold topics as well as core topics.

The types of topics can be categorized into three types: themes, methods, and fields. While most of the topics convey research "themes," some of them encompass the "methods" employed or "fields" being studied. For example, the following topics cover widely used techniques and approaches employed in TIM research: *[T20] Social network analysis*, *[T21] Multiple criteria decision making,* and *[T26] Bibliometric analysis of scientific publications*. Some topics represent specific fields and industries of interest in TIM research such as *[T12] Energy policy and sustainability*, *[T35] Innovation in high-tech industries*, and *[T45] ICT service management*, *[T49] Technological innovation in healthcare*, and *[T50] Technological innovation in food industry*. Such diversity in types of topics demonstrates the advantage of fractional assignments of LDA. For example, Kirkels and Duysters (2010) analyzed the collaboration network of SMEs in high-tech industries using social network analysis, which indicates that the article covers at least 4 out of the 50 topics identified: *[T19] Technological collaboration and R&D alliances* as a theme, *[T46] Innovation in SMEs* and *[T35] Innovation in high-tech industries* as fields, and *[T20] Social network analysis* as a method. The conventional approach of discrete assignments forces the article to be assigned to only one of the four topics, but LDA enables all the four topics to be exhibited by the article with their respective proportions.

4.2 Review of top 10 topics

Reviewing the core topics that have been most widely studied in TIM research can provide a quick overview of 20 years of TIM research. This section provides brief reviews of the top 10 topics with relevant article samples included in our data set.

[T1] NPD success factors is found to be the most popular topic. Since NPD is recongized as the most critical driver of competitive success of firms, much of the TIM literature has been devoted to identifying factors influencing NPD performance (Kalluri and Kodali 2014). The determinants of NPD performance have been examined and

identified from various perspectives such as product factors (e.g. product innovativeness), strategic factors (e.g. strategic fit), process factors (e.g. customer involvement), market factors (e.g. market potential), and organizational factors (e.g. cross-functional integration) (Evanschitzky et al. 2012).

The second most popular topic is *[T2] Empirical study on innovation and firm performance*. A number of studies have sought empirical evidence of the impact of innovation on firm's performance including productivity, profitability, and growth (Mansury and Love 2008). Earlier studies were centered on the impact of innovation inputs mainly measured by the amount of R&D, while the focus of recent works has shifted to innovation output including patents, product and/or process innovations, and new product sales (Hashi and Stojčić 2013). Factors influencing the relationship between innovation and firm performance have also been identified ranging from firm-specific factors such as size and organizational characteristics, to external factors including market competition and industry attractiveness (Koellinger 2008).

Much attention has also been paid to *[T3] Empirical study on firms' innovation activity*. Measuring firm's innovation activities is a critical input to innovation policy making and a prerequisite for exploring the relationship between innovation and firm performance (Evangelista et al. 1997). The development of harmonized surveys such as the Community Innovation Survey have boosted empirical studies of various aspects of firm's innovation activity including frequency of different types of innovations, sources of innovation, and barriers to innovation (e.g. Karlsson and Tavassoli 2016). Many attempts have also been made to identify different patterns and characteristics of innovation activities across industries (e.g. García-Piqueres et al. 2016) and countries (e.g. Faber and Hesen 2004).

[T4] Innovation diffusion modeling has been one of the most important topics in the fields of TIM as well as marketing (Kim et al. 2016). S-shaped diffusion models such as the Logistic and Bass models have been extensively employed to model the diffusion process of technological innovations. Many efforts have been made to modify and extend the basic diffusion models by incorporating additional variables influencing adoption of new technology to enhance the model validity and predictive power (e.g. Michalakelis et al. 2010). In the meanwhile, many applications have also been made to explore and compare different diffusion patterns across various technological innovations (e.g. Teng et al. 2002).

Since the publication of Chesbrough (2006), recent years have seen a huge increase in the number of articles dealing with *[T5] Open innovation*. Various approaches have been taken in open innovation research such as developing conceptual frameworks for types, process, and contexts of open innovation (e.g. Huizingh 2011), empirically investigating firms' open innovation activity and its effect on firm performance (e.g. Hung and Chou 2013), and deriving managerial implications from case studies (e.g. Carayannis and Meissner 2016). The topics of open innovation research include various aspects of open innovation such as its processes, tools, culture, and business models as well as partnering and alliances (Friesike et al. 2015).

[T6] R&D globalization has long been studied due to the growth of foreign R&D investment of multinational enterprises (MNEs) (Iwata et al. 2006). Much discussion has been made to the motivations of R&D off-shoring such as exploiting the capabilities and advantages at home and seeking new knowledge (Ambos and Ambos 2011). Some studies investigated whether R&D off-shoring has positive effects on host country firms from the perspective of technology spillovers (e.g. Qu et al. 2013), while others explored the benefits of R&D off-shoring in terms of reverse technology transfer and productivity growth at home (e.g. Castellani and Pieri 2013).

[T7] Public R&D policy is considered one of the most important themes in the field of public policy as well as TIM. Since the core instrument of public R&D policy is government R&D funding, many attempts have been made to measure the performance of government-funded R&D and explore the effect of subsidy on both public and private R&D (e.g. Xu et al. 2014). The findings revealed the role of government funding for university research is generally accepted, but there is controversy for private R&D (Feldman and Kelley 2006). Many efforts have also been made to derive policy implications for promoting collaboration and technology transfer between university and industry by identifying its facilitating factors and barriers (e.g. Plewa et al. 2013).

Much progress has been made in *[T8] Technology planning and roadmapping*. Technology roadmapping has been positioned as the most widely used technique for strategic technology planning. Earlier studies focused on developing basic frameworks of technology roadmaps such as formats and elements, and establishing the roadmapping process and procedure (e.g. Phaal et al. 2004). In an effort to increase the effectiveness of technology roadmapping process, recent years have witnessed many attempts to combine other techniques with technology roadmaps such as Delphi, analytic hierarchy process, and quality function deployment (e.g. Geum et al. 2011).

[T9] Technological forecasting and foresight has long been a major concern for TIM researchers, practitioners, and policy makers as it provides critical inputs to *[T7] public policy* and *[T8] Technology planning and roadmapping.* A variety of normative and/or exploratory methods have been developed to forecast technology futures, and can be classified into nine categories: expert opinion, trend analysis, monitoring and intelligence, statistical methods, modeling and simulation, scenarios, valuing/decision/economic methods, descriptive and matrices methods, and creativity (Technology Futures Analysis Methods Working Group 2004). An important research stream in recent years is to apply data mining techniques to science and technology database such as publications, patents, and projects (Porter 2007).

[T10] Knowledge flow and transfer has attracted much attention as the flow of knowledge plays a crucial role in the creation and diffusion of innovation (Mu et al. 2008). Many attempts have been made to measure knowledge flows at different levels such as inter-country (e.g. Wu and Mathews 2012), inter-organization (e.g. Azagra-Caro and Consoli 2016), and intra-organization (e.g. Lai et al. 2016). The popular approach is to use patent citations as a proxy for knowledge flows and analyzed the relationships from the perspective of social networks (e.g. Sorenson et al. 2006). Much of empirical research has also been devoted to identifying determinants of knowledge flows such as spatial and social proximity (e.g. Ensign et al. 2014).

5 Topic trends of TIM research

Examining the dynamic changes in popular topics over time can also provide fruitful implications for TIM researchers. The 20-year period under study is divided into four periods: Period 1 (1997–2001), Period 2 (2002–2005), Period 3 (2006–2010), and Period 4 (2011–2016). The length of the period is not even: 5 years for Period 1 and 3, 4 years for Period 2, and 5.5 years for Period 4. The reason for this is to make the linkage to the findings of previous studies. Linton and Thongpapanl (2004) reported the rankings of TIM journals based on a citation analysis of the articles published between 1997 and 2001 which corresponds to Period 1 of our analysis. The same analysis was conducted by

Thongpapanl (2012) to update the rankings of the journals for the period 2006–2010 corresponding to Period 3 of our analysis. Aligning the period of analysis along with the two previous studies enables us to recognize what topics actually led the TIM journals to the rankings in the corresponding periods.

Table 4 provides the topic rankings in terms of their proportions for the four different time periods. We can observe what topics were popular in each period and how topic rankings have changed from Table 4. Some topics show continuously increasing or decreasing trends. For example, the ranking of *[T1] NPD success factors* have been continuously decreasing since Period 2 although it is still included in top 10 topics in Period 4. On the contrary, *[T2] Empirical study on innovation and firm performance* started from 42th in Period 1, but is ranked first in Period 4.

The above analysis enables us to capture the overall changes in core topics of TIM research, but is less effective in examining the rise and fall of individual topics. The discrete period analysis does not allow the examination of annual fluctuations within the whole period under study. Investigating whether each topic is rising (hot topic) or falling (cold topic) in popularity over the whole time period may provide richer implications. Hot and cold topics can be identified based on the results of the LDA as those topics that have significantly increased or decreased topic proportions over time, respectively. Increased topic proportion indicates that researchers pay increasingly greater attention to that topic which can be perceived as an emerging research area, whereas decreased topic proportion would indicate the opposite. We employed a linear regression to identify hot and cold topics (Griffiths and Steyvers 2004). The year index was used as the input variable and topic distributions in the corresponding years were used as the response variable. Topics whose regression slopes are positive (negative) at a statistical significance level of 0.05 are determined as hot (cold) topics.

As a result, 14 hot topics and 16 cold topics have been derived as shown in Table 5. Figure 4 depicts the proportion trends of the top five (a) hot topics and (b) cold topics. As expected from Table 4, *[T2] Empirical study on innovation and firm performance* is found to be the hottest topic. Its proportion increased almost continuously from 2000 to 2009, with a sharp increase between 2009 and 2010. There is much overlap between the top 10 topics and the hot topics identified. Except for *[T1] NPD success factors* and *[T7] Public R&D policy (T7)*, eight of the top 10 topics are identified as hot topics. In addition, all of the top 10 topics of Period 1 are also included in hot topics except *[T1] NPD success factors*. Indeed, *[T1] NPD success factors* is even classified as a cold topic. Although it is found to be the most widely studied topic throughout the overall period, its popularity have continuously declined. The list of cold topics also overlaps significantly with the top 10 list of Period 1. Eight of the top 10 topics of Period 1 are revealed as cold topics except *[T7] Public R&D policy* and *[T6] R&D globalization*.

The remarkable topics in the hot topic list are *[T27] Patent analysis and intellectual property management* and *[T20] Social network analysis*. Both of the two topics are not highly ranked in the overall topic rankings. Nevertheless, recent years have seen an increase in the use of patent analysis and social network analysis as tools for TIM research. Since patents are a unique source of information about technological innovations, patent indicators have widely been used for measuring performance of innovation activities or technological innovation itself (e.g. Hagedoorn and Cloodt 2003), also for measuring technological relationships between technology domains or between innovative firms (e.g. Lee et al. 2016). Social network analysis, originally developed in social sciences, has often been employed for visualizing and grasping a technology structure as a technology domain can be described as a networked system of interconnected technological entities (Lee et al.

Table 4 Changes in topic rankings for the four periods

Topic	Period 1 (1997–2001)	Period 2 (2002–2005)	Period 3 (2006–2010)	Period 4 (2011–2016)
[T1] NPD success factors	1	1	2	8
[T2] Empirical study on innovation and firm performance	42	23	4	1
[T3] Empirical study on firms' innovation activity	24	3	1	9
[T4] Innovation diffusion modeling	26	8	27	3
[T5] Open innovation	48	36	3	4
[T6] R&D globalization	5	7	6	15
[T7] Public R&D policy	4	10	15	13
[T8] Technology planning and roadmapping	27	20	11	7
[T9] Technological forecasting and foresight	23	42	25	10
[T10] Knowledge flow and transfer	39	28	5	16
[T11] Technological transitions	47	47	18	5
[T12] Energy policy and sustainability	44	50	48	2
[T13] Consumer behavior and satisfaction	50	45	21	6
[T14] Technology transfer and commercialization	36	25	9	14
[T15] Technology advancement and fusion	6	9	33	27
[T16] Organizational learning	8	5	12	39
[T17] Technology development and economic growth	16	15	23	18
[T18] Organizational culture and organizational innovation	2	18	14	41
[T19] Technological collaboration and R&D alliances	34	11	10	17
[T20] Social network analysis	38	40	22	11
[T21] Multiple criteria decision making	15	14	28	22
[T22] NPD team management	11	12	8	38
[T23] Corporate management	10	4	7	46
[T24] Empirical study on innovation success factors	21	21	19	20
[T25] Academic research management	20	22	13	24
[T26] Bibliometric analysis of scientific publications	31	37	17	19
[T27] Patent analysis and intellectual property management	49	48	16	12
[T28] Product design	7	35	26	36
[T29] Miscellaneous 1	9	24	31	33
[T30] Types and patterns of innovation	28	27	29	21
[T31] Strategic portfolio management	17	13	24	40
[T32] Technological change and evolution	14	16	42	28

Table 4 continued

Topic	Period 1 (1997–2001)	Period 2 (2002–2005)	Period 3 (2006–2010)	Period 4 (2011–2016)
[T33] Innovation cycle	19	29	38	26
[T34] Disruptive innovation	33	19	37	25
[T35] Innovation in high-tech industries	13	2	30	50
[T36] Regional innovation system and industrial clusters	32	6	20	45
[T37] Systems of innovation approach	22	34	35	30
[T38] R&D human resource development	18	26	40	32
[T39] Total quality management and continuous improvement	3	31	41	48
[T40] Entrepreneurship and venture capital	46	17	32	34
[T41] Innovation capability and absorptive capacity	29	38	34	37
[T42] Miscellaneous 2	37	44	36	31
[T43] Supply chain management	25	33	47	42
[T44] Technology adoption	41	43	44	29
[T45] ICT service management	45	32	43	35
[T46] Innovation in SMEs	12	30	45	49
[T47] Software development and management	30	39	39	47
[T48] Technology standards and platform	40	41	46	43
[T49] Technological innovation in healthcare	43	49	49	23
[T50] Technological innovation in agriculture	35	46	50	44

2009). It has also been considered as a powerful tool for studying technological collaboration networks (e.g. De Prato and Nepelski 2014).

The coldest topic is *[T39] Total quality management and continuous improvement* followed by *[T18] Organizational culture and organizational innovation*. Those topics enjoyed much attention in the early 2000s and in the late 1990s, but their proportions have steadily decreased in TIM research. Including these two topics, many of the cold topics are related to the management of corporate functions such as human resource management, organizational management, and strategic management. This implies that there has been a shift of interest in TIM research from general topics of business management to those that are more specific to technology, innovation, and R&D as TIM has become a self-sustained discipline. This does not mean that the importance of business management topics has declined in the overall management literature, but only indicates a continuing fall in the share of such topics in the field of TIM.

Table 5 Hot and cold topics in TIM research

No.	Topic	Slope (×1000)
(a) Hot topics		
1	[T2] Empirical study on innovation and firm performance	0.6400
2	[T12] Energy policy and sustainability	0.6106
3	[T5] Open innovation	0.5752
4	[T11] Technological transitions	0.5415
5	[T13] Consumer behavior and satisfaction	0.4760
6	[T27] Patent analysis and intellectual property management	0.4322
7	[T4] Innovation diffusion modeling	0.3289
8	[T20] Social network analysis	0.2962
9	[T9] Technological forecasting and foresight	0.2336
10	[T8] Technology planning and roadmapping	0.2280
11	[T14] Technology transfer and commercialization	0.2216
12	[T3] Empirical study on firms' innovation activity	0.1900
13	[T10] Knowledge flow and transfer	0.1850
14	[T44] Technology adoption	0.0740
(b) Cold topics		
1	[T39] Total quality management and continuous improvement	−0.5000
2	[T18] Organizational culture and organizational innovation	−0.4654
3	[T35] Innovation in high-tech industries	−0.4113
4	[T46] Innovation in SMEs	−0.3689
5	[T16] Organizational learning	−0.3366
6	[T23] Corporate management	−0.3350
7	[T1] NPD success factors	−0.3185
8	[T15] Technology advancement and fusion	−0.2886
9	[T22] NPD team management	−0.2691
10	[T28] Product design	−0.2667
11	[T29] Miscellaneous 1	−0.2540
12	[T31] Strategic portfolio management	−0.2269
13	[T32] Technological change and evolution	−0.1865
14	[T38] R&D human resource development	−0.1531
15	[T43] Supply chain management	−0.1415
16	[T21] Multiple criteria decision making	−0.1320

Hot topics and cold topics have positive and negative regression slopes, respectively, at a statistical significance level of 0.05

6 Topics of TIM journals

6.1 Areas of subspecialty

TIM is interdisciplinary in nature (Lee 2015). The 11 TIM journals have different origins such as engineering, management, and economics. Each journal also has its own interests and scope, and core topics may differ across journals; thus, it is also worthwhile to

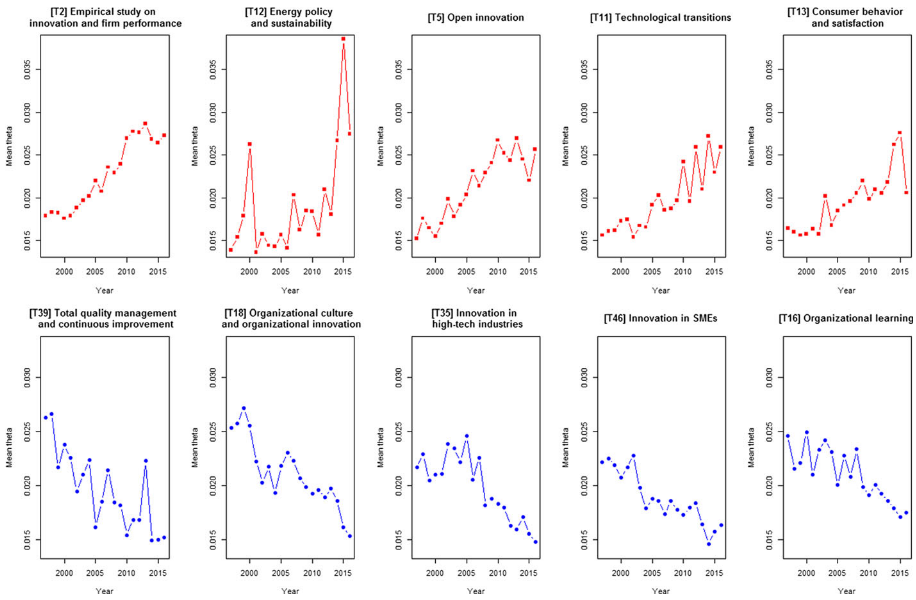

Fig. 4 Proportion trends of top five hot and cold topics in TIM research

compare topic distributions of the 11 TIM journals. Table 6 presents the top 10 topics for each of the 11 TIM journals. Remarkable differences in core topics are found between journals. The first ranked topics of the journals all differ except *[T22] NPD team management* for both IEEM and RDM. *[T12] Energy policy and sustainability* is ranked first in TFSC, but it is not included as top 10 topics in any of the other journals. The most frequently appearing topics are *[T3] Empirical study on firms' innovation activity* and *[T5] Open innovation* both of which are included as top 10 topics for seven journals. On the other hand, nine of the 50 topics never appear on any of the top 10 lists.

It is shown that the unique aims and scope of each journal is well represented by its top 10 topics. For example, JPIM places greater emphasis on product development and innovation. JTT is devoted to several topics related with technology transfer and commercialization. Most of the core topics of RP are centered on empirical findings that may provide policy and managerial implications for R&D. TFSC mainly deals with several topics on technological forecasting and planning. TASM encompasses various facets of technology-based organizations and corporate functions related with technology. In summary, while multidisciplinary efforts to link engineering, science, and management for dealing with various topics of technology, innovation, and R&D underlie all 11 TIM journals, each journal is observed to have somewhat different origins and interests.

Furthermore, some of the journals are relatively closer to each other in terms of overall topic distributions. Figure 5 shows the heatmap of journal similarity for all of the 50 topics with a dendrogram obtained from hierarchical clustering. The 11 TIM journals can be classified into three clusters. Cluster 1 including TASM, TFSC, RP, and JTT is more focused on innovation planning and policy. The five journals, RDM, JETM, JPIM, IEEM, and RTM are grouped into Cluster 2 dominated by engineering management. Cluster 3, comprising TVN and IJTM, is mainly concerned with the interfaces between technology and business functions.

Table 6 Top 10 topics of each of the 11 TIM journals

Rank	IEEM	IJTM	JETM	JPIM
1	[T22] NPD team management	[T6] R&D globalization	[T2] Empirical study on innovation and firm performance	[T1] NPD success factors
2	[T4] Innovation diffusion modeling	[T16] Organizational learning	[T22] NPD team management	[T2] Empirical study on innovation and firm performance
3	[T21] Multiple criteria decision making	[T39] Total quality management and continuous improvement	[T1] NPD success factors	[T13] Consumer behavior and satisfaction
4	[T43] Supply chain management	[T10] Knowledge flow and transfer	[T18] Organizational culture and organizational innovation	[T22] NPD team management
5	[T2] Empirical study on innovation and firm performance	[T23] Corporate management	[T3] Empirical study on firms' innovation activity	[T34] Disruptive innovation
6	[T13] Consumer behavior and satisfaction	[T18] Organizational culture and organizational innovation	[T10] Knowledge flow and transfer	[T18] Organizational culture and organizational innovation
7	[T28] Product design	[T35] Innovation in high-tech industries	[T31] Strategic portfolio management	[T28] Product design
8	[T18] Organizational culture and organizational innovation	[T31] Strategic portfolio management	[T20] Social network analysis	[T5] Open innovation
9	[T1] NPD success factors	[T5] Open innovation	[T24] Empirical study on innovation success factors	[T3] Empirical study on firms' innovation activity
10	[T39] Total quality management and continuous improvement	[T3] Empirical study on firms' innovation activity	[T41] Innovation capability and absorptive capacity	[T20] Social network analysis

Rank	JTT	RDM	RP	RTM
1	[T14] Technology transfer and commercialization	[T22] NPD team management	[T3] Empirical study on firms' innovation activity	[T23] Corporate management
2	[T25] Academic research management	[T5] Open innovation	[T27] Patent analysis and intellectual property management	[T18] Organizational culture and organizational innovation
3	[T7] Public R&D policy	[T2] Empirical study on innovation and firm performance	[T7] Public R&D policy	[T1] NPD success factors
4	[T27] Patent analysis and intellectual property management	[T1] NPD success factors	[T25] Academic research management	[T25] Academic research management

Table 6 continued

Rank	JTT	RDM	RP	RTM
5	[T17] Technology development and economic growth	[T18] Organizational culture and organizational innovation	[T17] Technology development and economic growth	[T38] R&D human resource development
6	[T36] Regional innovation system and industrial clusters	[T10] Knowledge flow and transfer	[T36] Regional innovation system and industrial clusters	[T22] NPD team management
7	[T4] Entrepreneurship and venture capital	[T3] Empirical study on firms' innovation activity	[T14] Technology transfer and commercialization	[T17] Technology development and economic growth
8	[T6] R&D globalization	[T23] Corporate management	[T5] Open innovation	[T5] Open innovation
9	[T19] Technological collaboration and R&D alliances	[T16] Organizational learning	[T6] R&D globalization	[T31] Strategic portfolio management
10	[T3] Empirical study on firms' innovation activity	[T25] Academic research management	[T10] Knowledge flow and transfer	[T28] Product design

Rank	TFSC	TASM	TVN
1	[T12] Energy policy and sustainability	[T11] Technological transitions	[T3] Empirical study on firms' innovation activity
2	[T4] Innovation diffusion modeling	[T5] Open innovation	[T46] Innovation in SMEs
3	[T9] Technological forecasting and foresight	[T31] Strategic portfolio management	[T5] Open innovation
4	[T8] Technology planning and roadmapping	[T30] Types and patterns of innovation	[T2] Empirical study on innovation and firm performance
5	[T11] Technological transitions	[T15] Technology advancement and fusion	[T35] Innovation in high-tech industries
6	[T49] Technological innovation in healthcare	[T16] Organizational learning	[T16] Organizational learning
7	[T33] Innovation cycle	[T19] Technological collaboration and R&D alliances	[T6] R&D globalization
8	[T13] Consumer behavior and satisfaction	[T2] Empirical study on innovation and firm performance	[T15] Technology advancement and fusion
9	[T6] R&D globalization	[T6] R&D globalization	[T7] Public R&D policy
10	[T15] Technology advancement and fusion	[T20] Social network analysis	[T24] Empirical study on innovation success factors

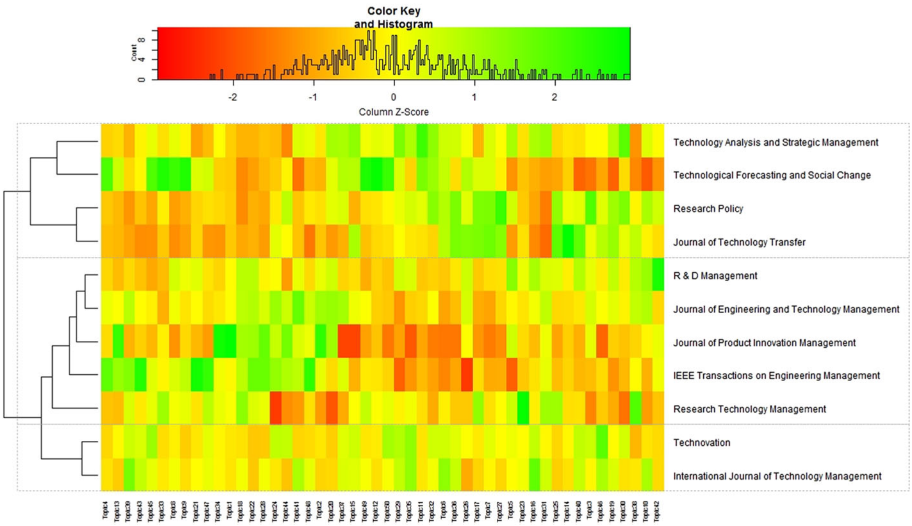

Fig. 5 Heatmap visualization of journal similarity

6.2 Effects of editor changes

One of the main roles and responsibilities of an editor-in-chief (EIC) of an academic journal is to set up the direction of the journal. An EIC establishes and defines journal's editorial policies and scope including topics of interests, and make editorial decisions in line with the policies. A change of EICs usually causes the changes of core topics covered in the journal, thus examining whether the changes of EIC in the TIM journals influence on journals' topic landscapes is also worthy of investigation.

The TIM journals experienced the EIC changes once or twice during the period of analysis except IJTM and JTT. M.A. Dorgham, the founding editor of IJTM, is still in charge of EIC in the journal, and Albert N. Link has been the EIC of JTT since 1996. Also, RP has been edited by the editorial team composed of several editors, not by an EIC. We thus investigated the topic differences between before and after EIC changes in the other eight journals. Table 7 summarizes the history of EIC changes of each journal. IEEM and JETM underwent the EIC changes twice during the period of analysis while there was only one change in the other journals.

Statistical hypothesis tests (Student's t test) were conducted to verify whether topic proportions of a journal change significantly when its EIC changes. We divided the entire 20-year period into two or three periods according to the number of EIC changes. Owing to the time difference between the manuscript submission and its publication, the beginning year of a new EIC was excluded from the analysis. For example, the EICs of IEEM were changed twice in 2003 and 2011; thus, the comparing periods are: 1997–2002 (former former EIC), 2004–2010 (former EIC), and 2012–2016 (current EIC). Then, for each topic, its mean proportions in the papers published in two consecutive periods were tested. Table 8 shows the number of increasing and decreasing topics after the change of EIC at a significance level of 0.05.

Table 7 EIC Changes of TIM journals

Journal	Current EIC	Former EIC	Former Former EIC
IEEM	Rajiv Sabherwal (2011-present)	George F. Farris (2003–2010)	Dundar F. Kocaoglu (1986–2002)
JETM	Jeremy K. Hall (2011-present)	Michael K. Badawy (1981–2010)	
JPIM	Gloria Barczak (2013-present)	C. Anthony Di Benedetto (2004–2012)	Abbie Griffin (1998–2003)
RDM	Ellen Enkel (2012-present)	Jeff Butler (1993–2011)	
RTM	James A. Euchner (2010-present)	Michael F. Wolff (1982–2009)	
TFSC	Fred Young Phillips (2011-present)	Harold A. Linstone (1969–2010)	
TASM	James Fleck (2014-present)	Harry Rothman (1989–2013)	
TVN	Jonathan Linton (2006-present)	G. Hayward (1984–2005)	

Table 8 The number of significantly changed topics after EIC change

Journal	Current versus Former			Former versus Former Former		
	Increasing	Decreasing	Total	Increasing	Decreasing	Total
IEEM	7	5	12	6	5	11
JETM	9	10	19	5	6	11
JPIM	2	3	5			
RDM	5	11	16			
RTM	5	8	13			
TFSC	8	15	23			
TASM	8	5	13			
TVN	13	12	25			

It is shown that a change of EIC in a journal has an impact on its topics covered to some extent. On average, 14.8 topics (29.6%) exhibited significant changes in their proportions between before and after EIC changes. Exactly half of the 50 topics changed in the case of TVN while only five topics were found to have differences for JPIM. It can be interpreted that an alternation of EIC of a journal yields considerable changes in its topic landscape, but the change is not so radical. The topics of interest may change by the editorial direction of a new EIC, but the basic identity does not readily change. However, it should be noted that the differences between periods may not be solely caused by the EIC change. As shown in Table 4, topic popularity have changed over time in TIM research, thus it cannot be denied that overall research trends before and after EIC changes have influences on topic popularity to some extent.

7 Conclusions

This study identified 50 topics of TIM research and explored their changes and relationships by applying the topic model approach to more than ten thousand articles published during the last twenty years. It was shown that the LDA model based on fractional assignments could automatically capture multiple topics underlying each article from huge volumes of scholarly data without predefined categories, which cannot be done by the conventional discrete assignments approach with predetermined categories. We also briefly reviewed the top 10 topics of TIM research, and examined the differences in interest across the journals through individual journal analysis. Dynamic changes in popular topics of TIM were also examined by uncovering the hot and cold topics, which revealed that the shift of TIM research focus from topics of general management to those of technology and innovation management. It was also revealed that the changes of EIC caused considerable changes in topic portfolios in the TIM journals.

Over the last three decades, TIM has continued to evolve and expand with great speed to be considered a self-sustained discipline. For continual progress in TIM research, the state of the art of TIM research must be grasped. The findings of this study are expected to provide fruitful implications for researchers, journal editors, and policy makers in the field of TIM. Researchers can judge whether their current research topics are hot or cold, and select appropriate journals to submit their works based on the list of core topics lists for each journal. Future research directions may be guided by the core and hot topics, and other related areas of research in the TIM network. Newcomers to the field, in both academia and practice, may easily gain an overview of TIM research. Journal editors can confirm that past publications have coincided with their own editorial policy, and may be guided in setting new editorial vision and direction. The list of hot topics can be useful references for grant allocation for government funding boards and designing national R&D programs for promoting TIM research.

Nevertheless, this study also has some limitations that may serve as avenues for further research. Firstly, we identified the topics for the entire 20-year period and used them as a common set in exploring dynamic changes in topic distributions for the four periods. However, a topic evolves over time by broadening its coverage or narrowing its scope. We cannot detect the evolution of topics with such a cumulative approach. Identifying topics for each period and linking the topics of different periods based on their similarity can enable us to trace the evolution of topics over time. Secondly, we used the basic LDA model to discover topics and their popularity trends over time. Since topic independence is assumed and the order of documents does not matter in the basic LDA, it is often too strict to be applied to certain domains. Recently, certain LDA extensions have been developed to relieve these assumptions. By employing them, we can explore the research dynamics of the TIM field more thoroughly with the resultant topic correlations, effect of publication time, and scholarly impact of articles. Thirdly, we ranked the 50 topics by their proportions as if they were at the same semantic level, but it seems that not all topics are placed at the same level in a topic hierarchy of TIM research. Directly comparing and ranking the topics at different levels may underrepresent or overrepresent the importance of topics. For example, *[T36] Regional innovation system and industrial clusters* can be considered as a sub-topic of *[T37] Systems of innovation approach*. If the two topics are merged into a single topic, the topic is likely to be more highly ranked. This problem can be addressed by employing the hierarchical LDA model in which a topic hierarchy is estimated based on a Bayesian approach (Blei et al. 2010). Once a topic hierarchy is constructed, the topics placed at the same level can be straightforwardly compared. Constructing the topic hierarchy of TIM research could also be a fruitful area of future research. Fourthly, this study only delineated what has been done in TIM research, but did not depict what the future of TIM research will look like. Applying some forecasting techniques to topic proportion changes may provide some clues for promising topics of future research.

Acknowledgements This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by both the Ministry of Science, ICT, and Future Planning (NRF-2014R1A1A1004648, NRF-2015R1A2A2A04007359) and the Ministry of Education (NRF-2016R1D1A1A09917423, NRF-2016R1D1A1B03930729).

References

Allen, T. J., & Sosa, M. L. (2004). 50 years of engineering management through the lens of the IEEE transactions. *IEEE Transactions on Engineering Management, 51*(4), 391–395.

Ambos, B., & Ambos, T. C. (2011). Meeting the challenge of offshoring R&D: An examination of firm- and location-specific factors. *R&D Management, 41*(2), 107–119.

Andrzejewski, D., Mulhern, A., Liblit, B., & Zhu, X. (2007). Statistical debugging using latent topic models. In *European conference on machine learning* (pp. 6–17). Springer.

Antons, D., Kleer, R., & Salge, T. O. (2016). Mapping the topic landscape of JPIM, 1984–2013. In search of hidden structures and development trajectories. *Journal of Product Innovation Management, 33*(6), 726–749.

Azagra-Caro, J. M., & Consoli, D. (2016). Knowledge flows, the influence of national R&D structure and the moderating role of public–private cooperation. *The Journal of Technology Transfer, 41*(1), 152–172.

Ball, D. F., & Rigby, J. (2006). Disseminating research in management of technology: Journals and authors. *R&D Management, 36*(2), 205–215.

Beyhan, B., & Cetindamar, D. (2011). No escape from the dominant theories: The analysis of intellectual pillars of technology management in developing countries. *Technological Forecasting and Social Change, 78*(1), 103–115.

Biemans, W., Griffin, A., & Moenaert, R. (2007). Twenty years of the journal of product innovation management: History, participants, and knowledge stock and flows. *Journal of Product Innovation Management, 24*(3), 193–213.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77–84.

Blei, D. M., Griffiths, T. L., & Jordan, M. I. (2010). The nested chinese restaurant process and bayesian nonparametric inference of topic hierarchies. *Journal of the ACM (JACM), 57*(2), 7.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3,* 993–1022.

Carayannis, E. G., & Meissner, D. (2016). Glocal targeted open innovation: Challenges, opportunities and implications for theory, policy and practice. *The Journal of Technology Transfer*. doi:10.1007/s10961-016-9497-0.

Castellani, D., & Pieri, F. (2013). R&D offshoring and the productivity growth of European regions. *Research Policy, 42*(9), 1581–1594.

Cetindamar, D., Wasti, S. N., Ansal, H., & Beyhan, B. (2009). Does technology management research diverge or converge in developing and developed countries? *Technovation, 29*(1), 45–58.

Cheng, C. H., Kumar, A., Motwani, J. G., Reisman, A., & Madan, M. S. (1999). A citation analysis of the technology innovation management journals. *IEEE Transactions on Engineering Management, 46*(1), 4–13.

Chesbrough, H. W. (2006). *Open innovation: The new imperative for creating and profiting from technology*. MA: Harvard Business Press.

Choi, D. G., Lee, Y., Jung, M., & Lee, H. (2012). National characteristics and competitiveness in MOT research: A comparative analysis of ten specialty journals, 2000–2009. *Technovation, 32*(1), 9–18.

De Battisti, F., Ferrara, A., & Salini, S. (2015). A decade of research in statistics: A topic model approach. *Scientometrics, 103*(2), 413–433.

De Prato, G., & Nepelski, D. (2014). Global technological collaboration network: Network analysis of international co-inventions. *The Journal of Technology Transfer, 39*(3), 358–375.

Ding, Y. (2011). Community detection: Topological vs. topical. *Journal of Informetrics, 5*(4), 498–514.

Durisin, B., Calabretta, G., & Parmeggiani, V. (2010). The intellectual structure of product innovation research: A bibliometric study of the journal of product innovation management, 1984–2004. *Journal of Product Innovation Management, 27*(3), 437–451.

Ensign, P. C., Lin, C.-D., Chreim, S., & Persaud, A. (2014). Proximity, knowledge transfer, and innovation in technology-based mergers and acquisitions. *International Journal of Technology Management, 66*(1), 1–31.

Evangelista, R., Perani, G., Rapiti, F., & Archibugi, D. (1997). Nature and impact of innovation in manufacturing industry: Some evidence from the Italian innovation survey. *Research Policy*, *26*(4–5), 521–536.

Evanschitzky, H., Eisend, M., Calantone, R. J., & Jiang, Y. (2012). Success factors of product innovation: an updated meta-analysis. *Journal of Product Innovation Management, 29,* 21–37.

Faber, J., & Hesen, A. B. (2004). Innovation capabilities of European nations: Cross-national analyses of patents and sales of product innovations. *Research Policy, 33*(2), 193–207.

Feldman, M. P., & Kelley, M. R. (2006). The ex ante assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior. *Research Policy, 35*(10), 1509–1521.

Friesike, S., Widenmayer, B., Gassmann, O., & Schildhauer, T. (2015). Opening science: Towards an agenda of open science in academia and industry. *The Journal of Technology Transfer, 40*(4), 581–601.

García-Piqueres, G., Serrano-Bedia, A. M., & López-Fernández, M. C. (2016). Sector innovation capacity determinants: Modelling and empirical evidence from Spain. *R&D Management, 46*(1), 80–95.

Geum, Y., Lee, S., Kang, D., & Park, Y. (2011). Technology roadmapping for technology-based product–service integration: A case study. *Journal of Engineering and Technology Management, 28*(3), 128–146.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences, 101*(suppl 1), 5228–5235.

Hagedoorn, J., & Cloodt, M. (2003). Measuring innovative performance: Is there an advantage in using multiple indicators? *Research Policy, 32*(8), 1365–1379.

Hashi, I., & Stojčić, N. (2013). The impact of innovation activities on firm performance using a multi-stage model: Evidence from the community innovation survey 4. *Research Policy, 42*(2), 353–366.

He, B., Ding, Y., Tang, J., Reguramalingam, V., & Bollen, J. (2013). Mining diversity subgraph in multidisciplinary scientific collaboration networks: A meso perspective. *Journal of Informetrics, 7*(1), 117–128.

Hornik, K., & Grün, B. (2011). Topicmodels: An R package for fitting topic models. *Journal of Statistical Software, 40*(13), 1–30.

Huizingh, E. K. R. E. (2011). Open innovation: State of the art and future perspectives. *Technovation, 31*(1), 2–9.

Hung, K.-P., & Chou, C. (2013). The impact of open innovation on firm performance: The moderating effects of internal R&D and environmental turbulence. *Technovation, 33*(10–11), 368–380.

Iwata, S., Kurokawa, S., & Fujisue, K. (2006). An analysis of global R&D activities of Japanese MNCs in the US from the knowledge-based view. *IEEE Transactions on Engineering Management, 53*(3), 361–379.

Jiang, H., Qiang, M., & Lin, P. (2016). A topic modeling based bibliometric exploration of hydropower research. *Renewable and Sustainable Energy Reviews, 57,* 226–237.

Kalluri, V., & Kodali, R. (2014). Analysis of new product development research: 1998–2009. *Benchmarking: An International Journal, 21*(4), 527–618.

Karlsson, C., & Tavassoli, S. (2016). Innovation strategies of firms: What strategies and why? *The Journal of Technology Transfer, 41*(6), 1483–1506.

Kim, T., Hong, J. S., & Lee, H. (2016). Predicting when the mass market starts to develop: The dual market model with delayed entry. *IMA Journal of Management Mathematics, 27*(3), 381–396.

Kirkels, Y., & Duysters, G. (2010). Brokerage in SME networks. *Research Policy, 39*(3), 375–385.

Koellinger, P. (2008). The relationship between technology, innovation, and firm performance—Empirical evidence from e-business in Europe. *Research Policy, 37*(8), 1317–1328.

Lai, J., Lui, S. S., & Tsang, E. W. (2016). Intrafirm knowledge transfer and employee innovative behavior: The role of total and balanced knowledge flows. *Journal of Product Innovation Management, 33*(1), 90–103.

Lee, H. (2015). Uncovering the multidisciplinary nature of technology management: Journal citation network analysis. *Scientometrics, 102*(1), 51–75.

Lee, H., Kim, C., Cho, H., & Park, Y. (2009). An ANP-based technology network for identification of core technologies: A case of telecommunication technologies. *Expert Systems with Applications, 36*(1), 894–908.

Lee, S., Kim, W., Lee, H., & Jeon, J. (2016). Identifying the structure of knowledge networks in the US mobile ecosystems: Patent citation analysis. *Technology Analysis & Strategic Management, 28*(4), 411–434.

Linton, J. D., & Thongpapanl, N. (2004). Perspective: Ranking the technology innovation management journals. *Journal of Product Innovation Management, 21*(2), 123–139.

Mansury, M. A., & Love, J. H. (2008). Innovation, productivity and growth in US business services: A firm-level analysis. *Technovation, 28*(1–2), 52–62.

McMillan, G. S. (2008). Mapping the invisible colleges of R&D Management. *R&D Management, 38*(1), 69–83.

Merino, M. T. G., do Carmo, M. L. P., & Álvarez, M. V. S. (2006). 25 years of technovation: Characterisation and evolution of the journal. *Technovation, 26*(12), 1303–1316.

Michalakelis, C., Varoutas, D., & Sphicopoulos, T. (2010). Innovation diffusion with generation substitution effects. *Technological Forecasting and Social Change, 77*(4), 541–557.

Mu, J., Peng, G., & Love, E. (2008). Interfirm networks, social capital, and knowledge flow. *Journal of Knowledge Management, 12*(4), 86–100.

Phaal, R., Farrukh, C. J. P., & Probert, D. R. (2004). Technology roadmapping—A planning framework for evolution and revolution. *Technological Forecasting and Social Change, 71*(1–2), 5–26.

Pilkington, A., & Teichert, T. (2006). Management of technology: Themes, concepts and relationships. *Technovation, 26*(3), 288–299.

Plewa, C., Korff, N., Baaken, T., & Macpherson, G. (2013). University–industry linkage evolution: An empirical investigation of relational success factors. *R&D Management, 43*(4), 365–380.

Porter, A. L. (2007). How "tech mining" can enhance R&D management. *Research-Technology Management, 50*(2), 15–20.

Qu, Z., Huang, C., Zhang, M., & Zhao, Y. (2013). R&D offshoring, technology learning and R&D efforts of host country firms in emerging economies. *Research Policy, 42*(2), 502–516.

Ramos-Rodríguez, A. R., & Ruíz-Navarro, J. (2004). Changes in the intellectual structure of strategic management research: A bibliometric study of the Strategic Management Journal, 1980–2000. *Strategic Management Journal, 25*(10), 981–1004.

Small, H., Boyack, K. W., & Klavans, R. (2014). Identifying emerging topics in science and technology. *Research Policy*, *43*(8), 1450–1467.

Sorenson, O., Rivkin, J. W., & Fleming, L. (2006). Complexity, networks and knowledge flow. *Research Policy, 35*(7), 994–1017.

Technology Futures Analysis Methods Working Group. (2004). Technology futures analysis: Toward integration of the field and new methods. *Technological Forecasting and Social Change, 71*(3), 287–303.

Teichert, T., & Pilkington, A. (2006). Themes, concepts and relationships in innovation research. In *Proceedings of the 15th IAMOT conference: East meets west* (pp. 22–26).

Teng, J. T., Grover, V., & Guttler, W. (2002). Information technology innovations: General diffusion patterns and its relationships to innovation characteristics. *IEEE Transactions on Engineering Management, 49*(1), 13–27.

Thongpapanl, N. (2012). The changing landscape of technology and innovation management: An updated ranking of journals in the field. *Technovation, 32*(5), 257–271.

Wang, H., Ding, Y., Tang, J., Dong, X., He, B., Qiu, J., et al. (2011). Finding complex biological relationships in recent PubMed articles using Bio-LDA. *PLoS ONE, 6*(3), e17243.

Wu, C.-Y., & Mathews, J. A. (2012). Knowledge flows in the solar photovoltaic industry: Insights from patenting by Taiwan, Korea, and China. *Research Policy, 41*(3), 524–540.

Xu, K., Huang, K.-F., & Xu, E. (2014). Giving fish or teaching to fish? An empirical study of the effects of government research and development policies. *R&D Management, 44*(5), 484–497.

Yan, E. (2014). Research dynamics: Measuring the continuity and popularity of research topics. *Journal of Informetrics, 8*(1), 98–110.

Yan, E., Ding, Y., Milojević, S., & Sugimoto, C. R. (2012). Topics in dynamic research communities: An exploratory study for the field of information retrieval. *Journal of Informetrics, 6*(1), 140–153.

Zhang, L. L. B. (2012). Aspect and entity extraction from opinion documents. In W. W. Chu (Ed.), *Data mining and knowledge discover for big data: Methodologies, challenge and opportunities* (pp. 1–40). Los Angeles, CA: Springer.