Information Processing and Management 57 (2020) 102258

Contents lists available at ScienceDirect

Information Processing and Management

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

Semantic Matching Efficiency of Supply and Demand Texts on Online Technology Trading Platforms: Taking the Electronic Information of Three Platforms as an Example

Xijun He[a], Xue Meng[a,*], Yuying Wu[a], Chee Seng Chan[b], Ting Pang[b]

[a] School of economics and management, Beijing University of Technology, No. 100 Ping Le Yuan, Chaoyang District, Beijing, 100124, China
[b] Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia

ARTICLE INFO

Keywords:
technology supply and demand texts
semantic matching
matching efficiency
online technology trading platforms (OTTPs)
Supply-Demand Matching Efficiency (SDME)

ABSTRACT

We calculated the matching values of technology supply and demand texts based on texts semantic similarity with Word2Vec and Cosine similarity algorithms, and then proposed a new index named Supply-Demand Matching Efficiency (SDME) to measure the matching efficiency of online technology trading platforms (OTTPs). Through the empirical research on the three types of OTTPs, the findings are as follows: First, the SDME of Zhejiang Market (Government-Owned, Government-Operated, GOGO), Technology E Market (Government-Owned, Contractor-Operated, GOCO), and Keyi Market (Market-Owned, Market-Operated, MOMO) are 64.69%, 54.38% and 28.99% respectively, indicating that the government plays an important role in attracting effective technology suppliers and demanders to participate in online trade and standardizing information expression, thereby improving the SDME. Second, by comparing the SDME and the newly announced signing rate of each OTTP, we found that the OTTP with high SDME also has high signing rate, and the changing trend of the two is consistent. Third, we used the TextRank and Latent Dirichlet Allocation (LDA) to study the topic distribution of technology supply and demand, and calculated the topic differences of each OTTP, which are 70%, 75%, 84% respectively. The Technology E Market and Zhejiang Market have low topic differences and high SDME, while Keyi Market has high topic differences and low SDME, which indicated that the topic differences have a negative effect on SDME. Intuitively, measuring the semantic matching efficiency of supply and demand texts on OTTPs can help the suppliers and demanders to retrieve information accurately, and assist the OTTPs to carry out trade promotion and evaluate trade performance.

1. Introduction

Recently, the Chinese government had successively approved the establishment of more than 450 technology trading institutions and platforms, with the vast majority providing online technology trading services. The OTTPs (i) ease the problems of information asymmetry in technology trade, (ii) break the limitations of distance and time, and (iii) provide support for micro-, small-, and medium-sized enterprises to release supply and demand information, assist in finding technology partners quickly, and actively participate in technology transfer. However, the low OTTP signing rates is the bottleneck of the development of online technology

* Corresponding author.
E-mail addresses: hexijun@bjut.edu.cn (X. He), mengxue@emails.bjut.edu.cn (X. Meng), wuyuying@bjut.edu.cn (Y. Wu), cs.chan@um.edu.my (C.S. Chan), pt@xxmu.edu.cn (T. Pang).

https://doi.org/10.1016/j.ipm.2020.102258
Received 23 December 2019; Received in revised form 27 March 2020; Accepted 27 March 2020
0306-4573/

market. As of January, 2020, Keyi Market had released 434,733 technology achievements but only signed 9740 contracts (2.24%), and Zhejiang Market had released 246,717 technology achievements and signed 47,411 contracts (19.22%).

According to Liu et al. (Liu, Bi, & Ye, 2016), there are many reasons for the low OTTP signing rates, such as the uncertainty of trading environment (the trust) of trading parties (Hua & Jiang, 2018), the price of technology commodity (He, & Li, 2014) and the technology convertibility (Hu, Tao, & Yuan, 2018). However, Yang et al. (Yang, Xia, & Ma, 2017) stated that it is the matching between *the tacit and explicit knowledge* and *the demand problems contained in technology texts* is the premise of effective docking and trade between the suppliers and demanders. Generally, the supply and demand information of OTTPs are mostly unstructured texts in natural language, and they differ in focus, i.e., the supply texts are focused on the technical components, materials, processes, and efficacy; while the demand texts are focused on the product defects and technical problems. Nonetheless, the supply and demand texts are usually short in length, often only consisting of a title and introduction, and with no significant topic features. Therefore, the semantic matching of technology supply and demand texts of OTTPs is facing a serious challenge.

Literately, there are a few researches on semantic matching between technology supply and demand texts. For instance, based on the features of technology transfer network platform and the demand orientation, Yang et al. (Yang, Xia, & Ma, 2017) proposed a new idea to measure the matching efficiency of OTTPs with supply and demand texts by using space vector model, TF-IDF function and similarity calculation. This inspired us to adopt the idea and proposed a new index named Supply-Demand Matching Efficiency (SDME) to measure the efficiency of OTTPs. Besides that, we also excavated the semantic features of the texts, thus providing ideas for tacit knowledge mining and knowledge matching of technology supply and demand texts.

2. Research objective and contribution

Our research objective is to extract the semantic features of technology supply and demand texts of OTTPs, and then use the new index – SDME to measure the supply and demand matching efficiency for OTTPs. It is based on the semantic similarity of technology supply and demand texts using Word2Vec and cosine similarity algorithms. According to the development modes of Chinese OTTPs, we select three representative OTTPs for empirical research where we calculate the matching values of technology supply and demand texts, further measure the SDME of the OTTPs and compare the SDME differences of the three OTTPs. Moreover, we use TextRank and Latent Dirichlet Allocation (LDA) to study the topic distribution of technology supply and demand of OTTPs. Finally, we put forward some suggestions to improve the SDME of OTTPs and promote supply and demand docking and trading.

As a summary, our main contributions are as follows:

(1) We use the semantic similarity between technology supply and demand texts to calculate the matching values, which can more accurately mine the tacit knowledge contained in technology texts. On this basis, we propose a new index named SDME, which provides ideas for estimating the trade possibility of technology supply and demand on OTTPs before technology trade occurs.
(2) We measure the SDME of the three representative OTTPs, and find that the government plays an important role in attracting effective technology suppliers and demanders to participate in online trade and standardizing information expression. By comparing the SDME with the newly announced signing rate of each OTTP, we find that the OTTP with high SDME also has high signing rate, and the changing trend of the two is consistent.
(3) We compare the SDME and the topic differences of technology supply and demand of OTTPs, and find that Technology E Market and Zhejiang Market have low topic differences and high SDME, while Keyi Market has high topic differences and low SDME, which shows that the topic differences have a negative effect on SDME.

The remainder of this paper is organized as follows: Section 3 reviews the previous research, Section 4 introduces the measurement process, Section 5 reports the results of the empirical research, Section 6 concludes and discusses our work.

3. Related work

In this section, we first provide a brief introduction on OTTPs, and follow by works that are related to us.

3.1. Characteristics of OTTP

The OTTP is an intermediary connecting technology suppliers and demanders. It provides specialized technology information and service resources for technology suppliers and demanders in the process of technology trade (Lichtenthaler, & Ernst, 2007). Also, it plays the role of knowledge public facilities (Cooke-Davies, 2002), and promotes the identification of technology trade opportunities (Li, Lan, & Liu, 2015). Compared with other intermediaries, the OTTPs can break the limitations of distance and time, accelerate the marketization of technology achievements, and reduce the cost of technology transfer (Disdier, & Head, 2008). In general, the OTTPs mainly provide the release and retrieval of technology supply and demand information, the communication of technology suppliers and demanders online, the technology trade consultation and other services. The technology information of the OTTPs generally exists in the form of texts. The matching and docking of technology supply and demand mainly depend on the active retrieval of technology suppliers and demanders and the human judgment of staff and experts (He, Ma, Wu, & Jiang, 2019).

3.2. Operational status of OTTPs

OTTPs in developed countries are relatively early in the building process (Li, Huang, & Zeng, 2018). For example,

(1) America: The National Technology Transfer Center (NTTC) was established in 1989 to develop a network platform that can provide comprehensive technology trading information and professional consulting services. Established in 1999, Yet2.com, the pioneer of virtual technology trading platforms, mainly carries out supply and demand search, intellectual property portfolio listing, and patent trading. InnoCentive was founded in 2001 as a pioneer of open innovation and crowdsourcing (Li, Huang, & Zeng, 2018).
(2) Britain: In 1991, the British government transferred the British Technology Group (BTG) to a joint consortium composed of British venture capital company, British bank, and BTG, etc. to realize its privatization, forming a model authorized by the British government and operated by the market. BTG was committed to select technology projects based on market demand and pushed them to the market, which involved finding, screening, and obtaining technology, evaluating technological achievements, protecting patents, and assisting in technology trading and commercial development.
(3) Germany: The Steinbeis Foundation for Economic Development (STW) established Steinbeis Transfer Centers (STC) in 1998, expanding its business from pure technology transfer to include technology consulting, research, and development (Li, Huang, & Zeng, 2018).
(4) European Union: In 1995, the European Commission established the Innovation Relay Center (IRC) (Wang, 2014) to serve small- and medium-sized enterprises, actively identify potential technology demands, implement network services, provide online technology information queries, and focus on cross-regional cooperation. In 2008, IRC and the Europe Information Center (EIC) merged to form the Enterprise Europe Network (EEN), which includes 17 industries such as agriculture, biotechnology, and environmental (Zhou, Zhang, & Tang, 2016). Similar technology trading platforms include Technomart in Japan and the Korea Technology Transfer Center (KTTC).

China: OTTPs in China were built relatively late. There are three main operational modes: (i) Government-Owned, Government-Operated (GOGO), (ii) Government-Owned, Contractor-Operated (GOCO), (iii) Market-Owned, Market-Operated (MOMO). Established in 2002, Zhejiang Market is the first public welfare technology trading platform owned by the government in China (Xiang, 2013), which is GOGO. Keyi Market was founded by Xiamen Keyi Technology Co., Ltd., in 2007, launched the first online technology trade service system, Keyi Bao, in 2013, and is the representative of MOMO. The Technology E Market was built in 2004 by China Technology Exchange Co., Ltd. and Chinese government is the representative of GOCO, providing a variety of integrated services such as research and development (R&D), consulting, trading, etc.

3.3. Texts matching measurement based on semantic features

The traditional method based on keyword matching represents each text as a group of keywords, without considering the semantic information, which leads to semantic confusion caused by polysemy, content mismatch caused by synonyms and so on (Wu, Zhu, Li, et al., 2017). Research on semantic similarity matching based on large-scale corpus training has developed rapidly (Martinez-Gil, & Chaves-Gonzalez, 2019). The matching method based on the semantic co-occurrence method (Chen, Zhou, & Zhang, 2018) is widely used in word sense disambiguation (Duque, Stevenson, Martinez-Romo, et al., 2018), document classification (Benedetti, Beneventano, Bergamaschi, et al., 2019), clustering analysis (Soares, Campello, Nourashrafeddin, et al., 2019), novelty detection (Kumar, & Bhatia, 2020) and so on. Compared with the previous linear semantic similarity matching method based on SVD (singular value decomposition, SVD) (Hawashin, Alzubi, Kanan, et al., 2019), its performance is significantly improved. However, the co-occurrence analysis only considers the words co-occurrence at the text level and equates co-occurrence and correlation. When co-occurrence times are the same, the correlation strength cannot be determined, and the possible semantic correlation between the non-co-occurrence keywords is ignored (Rule, Cointet, & Bearman, 2015). The semantic matching methods based on Wikipedia represents each text as a concept vector in the Wikipedia semantic space so as to obtain the semantic information contained in the text. According to the semantic representation of each text, the similarity between texts is calculated, which has better performance in the text classification (Wu, Zhu, Li, et al., 2017). However, the immense Wikipedia reference space results in that it needs to conduct a large number of full-text keyword matching operations to generate a document concept vector, thereby reducing the Wikipedia matching efficiency. The cosine similarity integrating contexts semantic information (e.g., vocabulary, knowledge) (Gu, Xu, & Zhou, 2018), semantic similarity calculation method based on word semantic comparison and corpus training (Shajalal, & Aono, 2019), and hierarchical semantic similarity calculation (Jia, Yang, Wu, et al., 2020) are widely used in text similarity, concept similarity and other fields. Measurement based on information content makes up for the lack of semantic representation ability of co-occurrence analysis, calculating the concept similarity according to the concept distribution in the labeled texts corpus (Racharak, Suntisrivaraporn, & Two, 2018), but it strongly depends on the availability of the corpus (Hussain, Wasti, Huang, et al., 2020) and manual labeling (Jiang, Bai, Zhang, et al., 2017; Xu, Dong, Liu, et al., 2017). The premise of domain ontology is that domain experts predefine ontology, which has high cost and weak scalability (Sousa, Silva, & Pesquita, 2020).

Compared to the above methods, the word vector model can use the contexts semantic information to express the texts in the form of a word vector through unsupervised learning (Tien, Le, Tomohiro, et al. 2019), without human labeling of the corpus, and it has strong scalability (Jiang, Li, & Huang, 2017; Khatua, Khatua, & Cambria, 2019). Therefore, we use the Word2Vec model proposed by Mikolov et al. (Mikolov et al., 2013) and integrate multiple corpora for model training, so as to express technology supply and demand in low-dimensional real number vectors and use cosine similarity to calculate the SDME.

3.4. Matching efficiency measurement of technology supply and demand

There is no unified definition of the matching efficiency of technology supply and demand. The transfer, application, and promotion rates of technology achievements are common indices (Danquah, 2018; Link, & Hasselt, 2019; Xu, Yang, & Luan, 2019; Lin, & Mao, 2019). However, these indices based on quantitative calculation cannot truly reflect the actual economic value or social benefits of technology achievements. The most representative indices of transformation efficiency considering value factors was proposed by the European Experts Committee of Knowledge Transfer Measurement (Xiong, 2017), including the agreements number of R&D cooperation, patent license revenue, the number of companies deriving from universities, and so on (Vinig & Lips, 2015; Ye, Yang, Han, et al., 2015; Sun & Liu, 2016; Sun & Grimes, 2017).

Most of these methods are quantitative analysis after technology trade from the perspective of input-output, which is difficult to measure the potential of transformation and utilization of technology achievements. Therefore, measuring the matching efficiency of OTTPs before the trade can help the suppliers and demanders to retrieve information accurately, and assist the OTTPs to carry out trade promotion and evaluate trade performance.

4. SDME measurement of technology supply and demand texts based on word vector

The main steps of the measurement process are as follows: (i) Supply and demand texts collection and word set extraction; (ii) Word vector model training; (iii) Semantic matching values and SDME calculation based on word vector; (iv) Topic differences analysis of supply and demand texts.

4.1. Word set extraction from supply and demand texts

We collected supply and demand texts from OTTPs by python and preprocessed them to obtain supply texts set $S = \{S_1, S_2, \ldots S_i, \ldots S_m\}$ and demand texts set $D = \{D_1, D_2, \ldots D_j, \ldots D_n\}$, where m and n are, respectively, the numbers of supply and demand texts. With the word segmentation tool, each supply and demand texts was transformed to supply and demand word sets $S_i = \{S_{i\_1}, S_{i\_2}, \ldots, S_{i\_p}\}(i = 1, 2, \ldots, m)$ and $D_j = \{D_{j\_1}, D_{j\_2}, \ldots, D_{j\_q}\}(j = 1, 2, \ldots, n)$, where p and q are, respectively, the numbers of supply and demand words in the supply word set S_i and demand word set D_j. More of the same words in two-word sets indicate higher grammatical similarity, but the meanings of the words and their semantic relationships with other words are ignored. Therefore, after removing repeated words from the supply and demand word sets, the supply word set S_i became $\bar{S}_i = \{\bar{S}_{i\_1}, \bar{S}_{i\_2}, \ldots, \bar{S}_{i\_p'}\}$, and the demand word set D_j became $\bar{D}_j = \{\bar{D}_{j\_1}, \bar{D}_{j\_2}, \ldots, \bar{D}_{j\_q'}\}$, where p' and q' are, respectively, the numbers of supply and demand words in the supply word set $\bar{S}_i$ and demand word set $\bar{D}_j$, and r is the number of repeated words, satisfying $0 \le r \le min(p, q)$.

4.2. Word vector model training

Taking data including the supply and demand information collected from OTTPs, Wikipedia data, and domain patent texts (title and introduction) in the Incopat patent database as the corpus, and using Word2Vec to train the word vector model, the words in the corpus were mapped into high-dimensional space to obtain the spatial word vector model.

4.3. SDME calculation

We used the word set similarity method based on word vectors to calculate the semantic matching values between supply and demand texts (Cui, Cai, & Feng, 2017; He, Ma, & Wu, 2018). First, we calculated the word similarity and constructed a semantic similarity matrix of the supply word set $\bar{S}_i$ and demand word set $\bar{D}_j$ in the form of a Cartesian product, and the semantic similarity matrix M_1 was as follows,

$$M_1 = \begin{pmatrix} Sim(\bar{S}_{i\_1}, \bar{D}_{j\_1}) & Sim(\bar{S}_{i\_1}, \bar{D}_{j\_2}) \cdots & Sim(\bar{S}_{i\_1}, \bar{D}_{j\_q'}) \\ Sim(\bar{S}_{i\_2}, \bar{D}_{j\_1}) & Sim(\bar{S}_{i\_2}, \bar{D}_{j\_2}) \cdots & Sim(\bar{S}_{i\_2}, \bar{D}_{j\_q'}) \\ \vdots & \vdots & \vdots \\ Sim(\bar{S}_{i\_p'}, \bar{D}_{j\_1}) & Sim(\bar{S}_{i\_p'}, \bar{D}_{j\_2}) \cdots & Sim(\bar{S}_{i\_p'}, \bar{D}_{j\_q'}) \end{pmatrix},$$

where $Sim(\bar{S}_{i\_1}, \bar{D}_{j\_1})$ in the matrix M_1 represents the semantic similarity between the word $\bar{S}_{i\_1}$ in word set $\bar{S}_i$ and the word $\bar{D}_{j\_1}$ in word set $\bar{D}_j$. The word vectors of $\bar{S}_{i\_1}$ and $\bar{D}_{j\_1}$ were, respectively, represented as a_i and b_i, which could be obtained by the word vector model in section 4.2, and h was the dimension of the word vectors. Then we used the cosine similarity algorithm to calculate the word similarity,

$$Sim(\bar{S}_{i\_1}, \bar{D}_{j\_1}) = \frac{\sum_{i=1}^{h}(a_i \times b_i)}{\sqrt{\sum_{i=1}^{h}(a_i)^2} \times \sqrt{\sum_{i=1}^{h}(b_i)^2}} \tag{1}$$

Second, we calculated the word set similarity. We found the maximum element $Sim(\bar{S}_{i\_k}, \bar{D}_{j\_v})$ in matrix M_1, added it to the set R, and deleted the elements in row k and column v where $Sim(\bar{S}_{i\_k}, \bar{D}_{j\_v})$ was located. We repeated this process until the number of elements T in set R was equal to $min(p', q')$, then all the elements in the matrix M_1 were deleted, and we finally obtained the set $R = \{Sim_1, Sim_2, ..., Sim_T\}$. The matching values of the supply word set $\bar{S}_i$ and the demand word set $\bar{D}_j$ was the weighted average of all elements in the set R. According to the equality of the elements in the set, the weighted average here was calculated as an arithmetic average. We standardized the results and calculated the final similarity,

$$Sim(S_i, D_j) = Sim(\bar{S}_i, \bar{D}_j) = \frac{(p+q) \times (r + \sum_{t=1}^{T} Sim_t)}{2pq} \tag{2}$$

We used the similarity of word sets to express the similarity of texts, and then the similarity matrix M_2 of technology supply and demand texts was obtained:

$$M_2 = \begin{pmatrix} Sim(S_1, D_1) & Sim(S_1, D_2)\cdots & Sim(S_1, D_n) \\ Sim(S_2, D_1) & Sim(S_2, D_2)\cdots & Sim(S_2, D_n) \\ \vdots & \vdots & \vdots \\ Sim(S_m, D_1) & Sim(S_m, D_2)\cdots & Sim(S_m, D_n) \end{pmatrix},$$

where $Sim(S_i, D_j)$ was the matching values between S_i and D_j. According to the above screening process, we obtained the set $G = \{Sim_1, Sim_2, ..., Sim_g\}$, which referred to the matching values and matching pairs of technology supply and demand texts after screening. We took the matching pairs with matching values greater than 0.50 as the successful matching pairs and counted the number of these as c. The formula for the *SDME* of technology supply and demand is

$$SDME = \frac{c}{min(m, n)} \times 100\% \tag{3}$$

4.4. Supply similarity and demand similarity

We used the word set similarity method based on word vectors to calculate the supply similarity and demand similarity matrices, analyzed the concentration of supply and demand, and further studied the relationship between SDME and concentration. For example, the demand similarity matrix M_3 is

$$M_3 = \begin{pmatrix} Sim(D_1, D_1) & Sim(D_1, D_2)\cdots & Sim(D_1, D_n) \\ Sim(D_2, D_1) & Sim(D_2, D_2)\cdots & Sim(D_2, D_n) \\ \vdots & \vdots & \vdots \\ Sim(D_n, D_1) & Sim(D_n, D_2)\cdots & Sim(D_n, D_n) \end{pmatrix}.$$

The final demand similarity of technology demand texts D_i referred to the average similarity between D_i and all demand texts in the technology demand set, which can be calculated by the arithmetic average of all elements in row i of matrix M_3 according to formula (4):

$$SimD_i = \frac{\sum_{j=1}^{n} Sim(D_i, D_j)}{n} \tag{4}$$

We similarly obtained the supply similarity matrix M_4 and the final supply similarity of each supply texts.

4.5. Supply and demand topic differences

When both the supply similarity and the demand similarity are very high, but the matching rate of supply and demand is very low, we need to analyze the difference of supply and demand topics. In this section, the TextRank algorithm and LDA are combined to study the distribution and difference of supply and demand topics. First, according to the basic idea of TextRank, a word graph was constructed, where the adjacency relation of the constituent words in the technology supply (or demand) texts has been preprocessed. Second, according to formula (5), the TextRank value of each node in the word graph was calculated iteratively until it converges (Lin, Miao, & Zhang, 2019):

$$TR(V_i) = (1-d) + d \times \sum_{V_j \in In(V_i)} \frac{w_{ji}}{\sum_{V_k \in Out(V_j)} w_{jk}} TR(V_j) \tag{5}$$

where $d \in [0, 1]$ is the damping coefficient, representing the probability that a particular node in the word graph points to any other node, usually taking the value 0.85; w_{ji} is the connection weight from word V_j to word V_i; $Out(V_j)$ is the word set pointed to by word V_j; and $In(V_i)$ is the word set that points to word V_i. The initial default weight of each word node is 1. The weight contribution of a word node is transferred to the adjacent nodes in the form of an equal weight (Ning & Liu, 2016).

Third, reversed the TextRank values of the nodes, and took the first x words as the keyword set, which was regarded as the domain dictionary for texts preprocessing of the LDA in the next step.

Fourthly, the LDA was used to study the topic distribution of technology supply and demand. The basic principle of topic extraction with LDA is that each texts is regarded as the probability distribution of a potential topic, each regarded as a probability distribution of the words contained in the texts set; hence, the high-probability topic was screened for the analysis of supply and demand difference (Feng & Zhang, 2017).

Finally, we calculated the topic differences of supply and demand texts. Specifically, we counted the total number of different topic words in each corresponding supply topic and demand topic, and then used the number of topics to standardize, as shown in formula (6):

$$d = \frac{\sum_{i=1}^{n} s_i}{w} \times 100\% \tag{6}$$

where d is the topic differences of supply and demand texts of an OTTP, S_i is the number of the topic words which is different in the supply topic i and demand topic i, n is the number of the topics of supply and demand of an OTTP, and w is the number of all topic words of supply and demand of an OTTP.

5. Empirical research on SDME of three OTTPs

5.1. Data collection

As typical representatives for empirical analysis, we selected three OTTPs. Zhejiang Market is Government-Owned, Government-Operated (GOGO), Keyi Market is Market-Owned, Market-Operated (MOMO), and Technology E Market Government-Owned, Contractor-Operated (GOCO). And we obtained the technology supply and demand information in the field of electronic information with python, including the title and introduction of supply and demand projects. After removing repeated records, we had 475 supplies and 405 demands for Zhejiang Market, 500 supplies and 445 demands for Keyi Market, and 495 supplies and 480 demands for Technology E Market.

5.2. Validity test of supply and demand matching method

Take Technology E Market as an example. First, 40 pairs of technology supply and demand texts were taken as validation samples. After manual labeling, 20 pairs were positive samples and 20 were negative samples, so the ratio of positive and negative samples was 1:1. The matching values of supply and demand texts pairs was calculated using the supply and demand matching method. If the result was greater than 0.5, it was determined as matching, and otherwise, it was not matching. We compared the precision P, recall R, and F_1 values of the method with the results of manual labeling:

$$P = \frac{TP}{TP + FP},\ R = \frac{TP}{TP + FN},\ F_1 = \frac{2 \times R \times P}{R + P} \tag{7}$$

where TP is the number of positive samples predicted as positive samples, FP is the number of negative samples predicted as positive samples, and FN is the number of positive samples predicted as negative samples. The precision of the method was 0.739, the recall was 0.850, and the F_1 value was 0.791, indicating a valid method.

5.3. SDME calculation

We calculated the semantic matching values of supply and demand texts for the three OTTPs and took the matching pairs with matching values higher than 0.50 as successful matching pairs and calculated the SDME of OTTPs based on the idea of two-sided matching. The supply similarity of each OTTP in Table 1 was calculated by the average of the final supply similarity of all supply texts in the OTTPs. The calculation of demand similarity of each OTTP in Table 1 was the same. Table 1 shows the SDME, supply similarity, and demand similarity of each OTTP, while Fig. 1 shows the distribution of matching values, supply average similarity, and demand average similarity (For each OTTP, supply average similarity refers to the average value of the similarity between each supply texts and other supply texts except itself. The calculation principle of demand average similarity is similar).

According to the SDME, the GOGO Zhejiang Market (64.69%) > the GOCO Technology E Market (54.38%)> the MOMO Keyi Market (28.99%), which shows that government support can attract technology suppliers and demanders to participate in OTTP trading activities, and standardize the expression of supply and demand information, thus promoting the SDME.

According to the newly announced signing rates, the signing rates of the three OTTPs are 19.22%[1], 7.49%[2] and 2.24%[3] respectively. By comparing the SDME and the signing rate of each OTTP, we find that the OTTP with high SDME also has high signing rate, and the changing trend of the two is consistent.

According to the matching values of supply and demand, among the successful matching pairs of Zhejiang Market, Technology E

[1] Data source: http://www.51jishu.com
[2] Data source: http://us.ctex.cn/article/xcp/201706/20170600040966.shtml
[3] Data source: https://www.1633.com

Table 1
SDME, supply similarity, and demand similarity of the three OTTPs.

Index/OTTPs	Technology E Market	Keyi Market	Zhejiang Market
SDME (%)	54.38	28.99	64.69
supply similarity	0.5635	0.5317	0.5283
demand similarity	0.4827	0.5390	0.4833

Market, and Keyi Market, the proportion of matching values of $0.6-0.7$ is the greatest for each, at 90%, 81%, and 93%, respectively, which shows that the matching values are low.

According to the distribution of supply and demand average similarity, the supply average similarity of Technology E Market and Zhejiang Market is higher than that of demand average similarity, indicating that the technology supply direction is relatively consistent. Combined with high SDME, the differences between the supply and demand directions of the two OTTPs are low. The supply and demand similarity of Keyi Market differs little, but the supply and demand directions of the OTTPs are quite different. To further analyze the difference between technology supply and demand directions on the three OTTPs, LDA was used to extract the topics of supply and demand texts.

5.4. Topic distribution

On the basis of keyword extraction using TextRank, we used LDA to extract the topic, selected the top 50 topic words, and combined the technology and product categories and subcategories in the new-generation information technology industry (chapter 1) and high-end equipment manufacturing industry (chapter 2) of *The Catalogue of Key Products and Services in Strategic Emerging Industries (No. 1 document in 2017)* to classify the 50 topic words based on manual classification and subject word classification. The results are shown in Tables 2–4.

According to Tables 2–4, Zhejiang Market has the lowest difference in the distribution of supply and demand topics, which are divided into eight categories: artificial intelligence software and equipment, optical communication equipment, network equipment and electronic instrument, satellite mobile communication and navigation terminal, digital video monitoring system, information terminal equipment, network and information security software, and e-commerce. To a certain extent, the low difference of the topics types and topics words can improve their SDME.

Keyi Market has the highest difference in the distribution of supply and demand topics, among which the consistent types include artificial intelligence software and equipment, optical communication equipment, network equipment and electronic instrument. The technology supply of this OTTP also involves satellite mobile communication and navigation terminal, virtual reality technology, and digital video monitoring system, while the technology demand focuses on information terminal equipment, efficient energy-saving technology and equipment in the electronic industry, as well as e-commerce and electronic information. This OTTP has the lowest SDME.

There is a relatively low difference in the distribution of supply and demand topics in Technology E Market, among which five categories are consistent: artificial intelligence software and equipment, optical communication equipment, network equipment and electronic instruments, satellite mobile communication and navigation terminal, and digital video monitoring system. The technology supply of this OTTP also involves virtual reality technology and network and information security software, while the technology demand focuses on cloud computing device and integrated circuit. The SDME of this OTTP is relatively high.

5.5. Topic differences

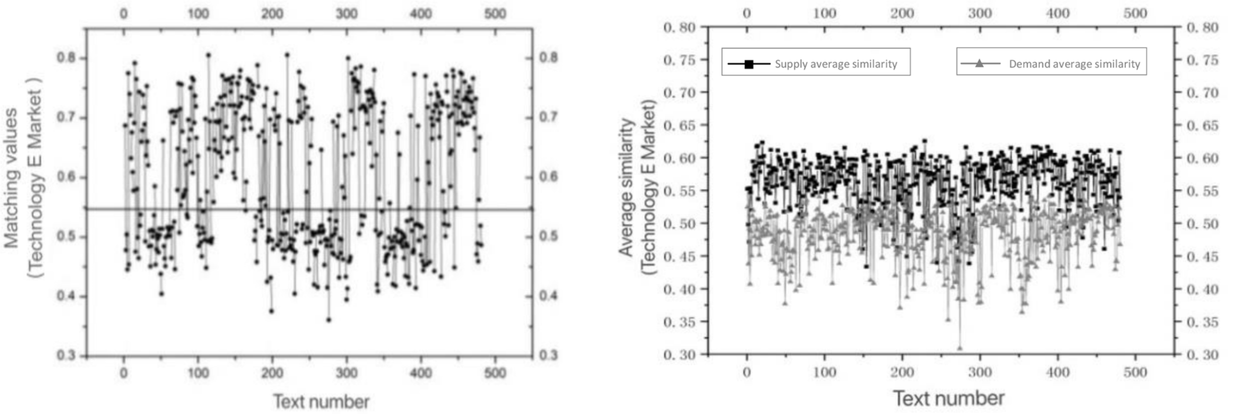
We calculated the topic differences of each OTTP. The results are shown in Table 5.

Considering and comparing the SDME and topic differences, we found that Technology E Market and Zhejiang Market have low topic differences and high SDME, while Keyi Market has high topic differences and low SDME. The SDME of an OTTP is high when the topic differences are small, which shows that the topic differences have a negative effect on the SDME of OTTPs.

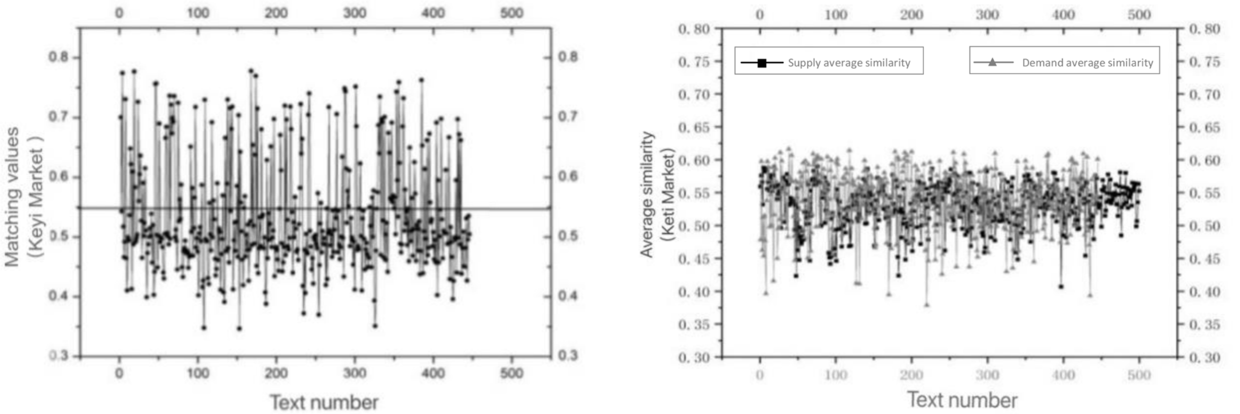
6. Conclusions and discussions

6.1. Conclusions

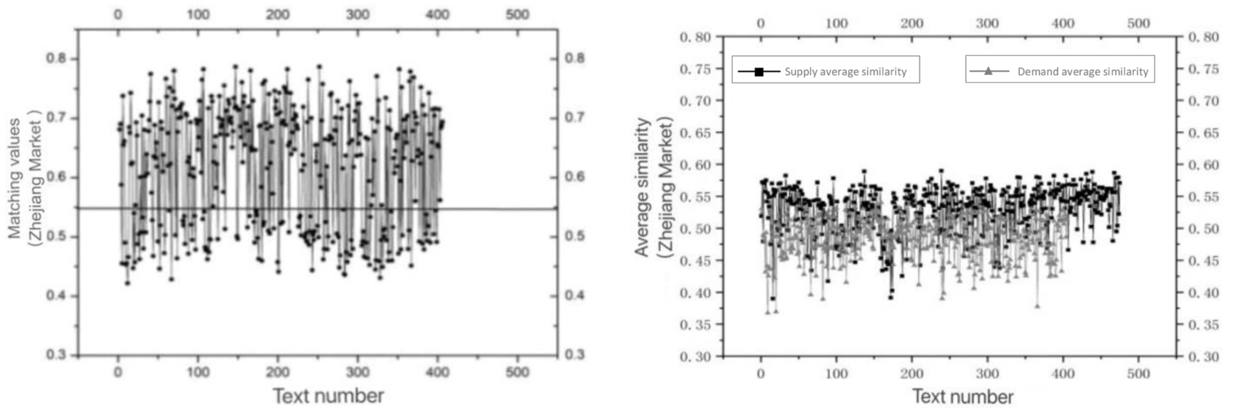
This paper uses the semantic similarity of technology supply and demand texts to calculate the matching values of technology supply and demand texts, and further proposes a new index named SDME to measure the matching efficiency of the OTTPs. Through the empirical research of technology supply and demand texts in the field of electronic information, we find that: First, government support has a positive effect in promoting the SDME. More concretely, government support can attract technology suppliers and demanders to participate in OTTPs trading activities, and standardize the expression of supply and demand information. Second, we compare the SDME and the latest announced signing rate of each OTTP, and find that the OTTP with high SDME also has high signing rate, and the changing trend of the two is consistent. Third, Technology E Market and Zhejiang Market have low topic differences and high SDME, while Keyi Market has high topic differences and low SDME, which indicated that the topic differences have a negative



Technology E Market: (a) matching values distribution; (b) similarity distribution



Keyi Market: (c) matching values distribution; (d) similarity distribution



Zhejiang Market: (e) matching values distribution; (f) similarity distribution

Fig. 1. Distribution of matching values, supply similarity, and demand similarity.

effect on SDME.

We mine the tacit knowledge and semantic information contained in the technology supply and demand texts, and propose a new index to study the supply and demand efficiency of OTTPs. On the one hand, our research provides a feasible idea to evaluate the matching degree of supply and demand before technology trade. On the other hand, it can assist OTTPs, government and small-, medium-, and micro-sized enterprises to make management decisions. Specifically, our research provides (i) methods and ideas for

Table 2
Topic distribution of technology supply and demand of Keyi Market.

Number	Supply classification	Supply texts topic	Number	Demand classification	Demand texts topic
1	Artificial intelligence software and equipment	Big data/Platform /Intelligent/Software/System/Smart/ Management/Service	1	Artificial intelligence software and equipment	Big data/Platform /Intelligent/Software/System/Smart/ Management/Service/Project /Internet/ IT
2	Optical communication equipment	LED/Communication/Technology/Wireless/Equipment/ Transmission/ Device/ Panel light/Lighting/Test/Signal /Optical fiber/Optics	2	Optical communication equipment	LED/Communication/ Technology/Wireless/ Equipment/ Same screen transmission/Antenna/Bluetooth/Lithiumn/ Battery/Microwave/Solution
3	Network equipment and electronic instrument	Sensor/Electronic/Network/Sensor-network/ Electromagnetic/Semiconductor/Storage	3	Network equipment and electronic instruments	Sensor/Electronic/Network/Internet of things/Circuit board/ Conductivity meter/Computer
4	Satellite mobile communication and navigation terminal	Vehicle/Recorder/Theft-proof/Geography/Location	4	Information terminal equipment	APP/Touch screen/Game/Multimedia/Mobile phone/Printer
5	Virtual reality technology	VR/AR/Simulation/ Three-dimensional/Emulation/ Environment	5	Efficient energy saving technology and equipment in electronic industry	Energy /Environmental protection/Material
6	Digital video monitoring system	Digital/Image/Video/Audio/3D/Detection/Monitoring/ Scanning/Supervisory control/Real-time/Remote	6	E-commerce and e-information talents	E-commerce/Consumption/Patent/Investment/ Entrepreneurship/Talent/Labor force /R&D/Innovation/ Application/Construction

Table 3
Topic distribution of technology supply and demand of Technology E Market.

Number	Supply classification	Supply texts topic	Number	Demand classification	Demand texts topic
1	Artificial intelligence software and equipment	Intelligent/Service/ System/Data/Management/ Software/ Platform/Robot/Early warning system	1	Artificial intelligence software and equipment	Intelligent/Service/ System/Data /Management/Software /Platform /Application software
2	Optical communication equipment	Wireless/Optical fiber/Frequency band/Device/Equipment/ Base station/Load/ Ultrasonic/Cable/Night vision sight/ Transmitter/Solution/Laser/Component/ Subgroup/Power supply	2	Optical communication equipment	Wireless/Optical fiber/ Frequency conversion/ Device/ Equipment/Communication/LED/Liquid crystal/Backlight/ Cold-cathode/Grating/ Fluorescent lamp/Battery/ Diode
3	Network equipment and electronic instrument	Computer/Controller/Sensor/ Electric machinery/Module/ Router/Laser/Detector/Heat metering/ Network	3	Network equipment and electronic instruments	Computer/Control technology/Sensor/Electric machinery/Flat/ Control system/Internet of things/RFID
4	Satellite mobile communication and navigation terminal	Location/Region/Resolution	4	Satellite mobile communication and navigation terminal	Automatic location/Monitoring system
5	Digital video monitoring system	Image/Video/Product/Digit/Remote/Detection	5	Digital video monitoring system	Image/Video/Product/Digit/ Monitor/Remote monitoring /Image recognition/Detection technology
6	Virtual reality technology	Virtual reality/Scene/Target	6	Cloud computing device	Cloud computing/Algorithm
7	Network and information security software	Database/Data center/Information	7	Integrated circuit	Electrode/Resistance/Chip/Circuit

Table 4
Topic distribution of technology supply and demand of Zhejiang Market.

Number	Supply classification	Supply texts topic	Number	Demand classification	Demand texts topic
1	Artificial intelligence software and equipment	Artificial intelligence/Internet/Management system/Data/Platform/Service/Intelligent/Decision/Smart /Management/Software/ Robot	1	Artificial intelligence software and equipment	Internet/Internet/Management system/Data/Platform/Service/Management software/Smart/Service platform/Data collection/Robot/Software development
2	Optical communication equipment	Wireless/Optical fiber/Sensor/Antenna/Microwave/Laser/Sensing/Broadband/Test/Equipment	2	Optical communication equipment	Wireless/Optical cable/Sensor/Communication/Signal/Equipment
3	Network equipment and electronic instrument	Network/Electronic/ Control/Internet of things/Module/Electromagnetic flowmeter	3	Network equipment and electronic instruments	Network/Electronic/ Controller/Internet of things/Module/Server/Broadband/Control system/Data storage
4	Satellite mobile communication and navigation terminal	Mobile/Automatic/Position system	4	Satellite mobile communication and navigation terminal	GIS/Position system
5	Digital video monitoring system	Digital/Video/Monitoring system/Remote/Monitor/Image detection/Online/Recognition system/Product	5	Digital video monitoring system	Digital/Video/Monitoring system/Remote/Monitor/Test/Real-time/Video monitoring
6	Information terminal equipment	Mobile phone/Terminal/Multimedia	6	Information terminal equipment	Mobile phone/Terminal/WeChat
7	Network and information security software	Information/Blockchain	7	Network and information security software	Information/Network security/Database/Program/Backstage/Network monitoring
8	E-Commerce	E-Commerce/Enterprise/Logistics/Customer	8	E-Commerce	E-Commerce/Business management/Supply chain

Table 5
The topic differences of each OTTP.

Index OTTPs	Technology E Market	Keyi Market	Zhejiang Market
topic differences (%)	75.00	84.00	70.00

OTTPs to predict signing rate, evaluate trade performance, formulate trade promotion strategy and develop intelligent recommendation function of supply and demand, (ii) data support for government to evaluate operation efficiency of each OTTP and formulate management policies, (iii) online retrieval technology services for many small-, medium-, and micro-sized enterprises to help enterprises quickly find trading partners and reduce the search cost of both the suppliers and demanders.

6.2. Discussions

In order to improve the SDME of OTTPs and promote supply and demand docking and trading, we put forward the following considerations:

First, the OTTPs should further focus on the service field to improve the professional services. At the moment, the service field is relatively scattered, the classification is confused and lack of standards, which leads to low efficiency in searching supply and demand information. Moreover, effective technology supply and demand resources in the same field are constructed repeatedly, and domain experts are connected to many platforms, but the level of service personnel specialization is weak. For example, the key service fields of BTG in Britain only focus on medicine, natural science, biological science, electronics, and communications.

Second, it is necessary to construct the technology domain knowledge graphs to improve the accuracy of semantic matching. Specifically, accurate semantic matching of technology supply and demand texts needs the support of domain dictionaries and domain knowledge. For example, knowledge base construction includes a technology component library, technology process library, technology performance library, and relationship graph among components, technology, and performance. Moreover, we feel that selecting effective technology supply and demand information can improve the SDME and docking success rate. On this basis, artificial intelligence, big data, natural language processing, and knowledge graph technologies can be combined to develop an effective platform for automatic retrieval of technology supply and demand information and intelligent recommendations, so as to promote accurate and intelligent matching of technology supply and demand based on knowledge inference.

Third, standardize the supply and demand information, and enrich the multi-dimensional attributes and relationship features of suppliers and demanders. Based on the technology domain knowledge graphs, it is necessary to standardize the description of unstructured technology texts, reduce the interference of non-technical information. Moreover, we should avoid the lack of multi-dimensional attributes of supply and demand subjects, such as the type and scale of supply and demand subjects, price range and maturity of supply and demand technology, and cooperation mode of supply and demand, so as to improve the accuracy of supply and demand matching under multi-dimensional semantic feature fusion.

Finally, small-, medium-, and micro-sized enterprises should be encouraged to participate in online technology trading and innovation activities. These enterprises are weak in innovation and risk resistance due to their small-scale and limited capital. However, their technology demand is easier to find a mature technology supply in the market. Therefore, in order to encourage these enterprises to participate in online technology trade, appropriately reducing information release fees, the proportion of transaction fees and value-added service fees charged by a platform, as well as providing subsidies and incentives to enterprises that actively participate in online technology trading and innovation will promote the improvement of their innovation ability.

However, there are still some limitations in our research. First, there are many OTTPs in China, and we only select three representative OTTPs for empirical research. As future work, we are planning to further expand the samples to evaluate the semantic matching efficiency and features of supply and demand texts in different regions and different technology fields. Second, we will continue to mine and refine the multi-dimensional tacit knowledge such as technology points, technology efficacy and technology field contained in the technology supply and demand texts, and construct the knowledge graph to explore the intelligent matching method of technology supply and demand texts, so as to improve the matching accuracy and assist in trade decision-making.

Funding

This work has been supported by the National Natural Science Foundation of China (grant numbers: 71974009)，Natural Science Foundation of Beijing Municipality (grant numbers: 9172002) and Educational Management Research Project of Beijing University of Technology (grant numbers: GL2017-A03).

References

Benedetti, F., Beneventano, D., Bergamaschi, S., & Simonini, G. (2019). Computing inter-document similarity with context semantic analysis. *Information Systems, 80*, 136–147.

Chen, Y., Zhou, X, & Zhang, L. (2018). Research on the methods to identify online food safety rumor-related documents based on semantic co-occurrence matching. *Information Studies: Theory & Application, 41*(6), 130–136.
Cooke-Davies, T. (2002). The “real” success factors on projects. *International Journal of Project Management, 20*(3), 185–190.
Cui, X., Cai, S., & Feng, J. (2017). Ontology-based similarity measurement for online complaint cases of communications services. *Systems Engineering—Theory&Practice, 37*(6), 1638–1647.
Danquah, M. (2018). Technology transfer, adoption of technology and the efficiency of nations: Empirical evidence from sub Saharan Africa. *Technological Forecasting and Social Change, 131*, 175–182.
Disdier, A., & Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and Statistics, 90*(1), 37–48.
Duque, A., Stevenson, M., Martinez-Romo, J., & Araujo, L. (2018). Co-occurrence graphs for word sense disambiguation in the biomedical domain. *Artificial Intelligence in Medicine, 87*, 9–19.
Feng, J., & Zhang, Y. (2017). Research on the method of detecting and analyzing scientific fronts based on LDA and ontology. *Information Studies:Theory&Application, 40*(8), 49–54.
Gu, C., Xu, H., Zhou, H., et al. (2018). Texts similarity computing based on lexical semantic information. *Application Research of Computers, 35*(02), 391–395.
Hawashin, B., Alzubi, S., Kanan, T., & Mansour, A. (2019). An efficient semantic recommender method for Arabic text. *Electronic Library, 37*(2), 263–280.
He, W., & Li, M. (2014). Information asymmetry, transaction costs and use of enterprise mechanism on electronic commerce platform. *Technology Economics, 33*(6), 26–31.
He, X., Ma, S., & Wu, Y. (2018). Research on semantic matching for online technology supply and demand information based on ontology and SAO structure. *Information Science, 36*(11), 95–100.
He, X., Ma, S., Wu, Y., & Jiang, G. (2019). Research on the supply and demand matching of online technology trading platform based on multi-level feature fusion—Taking the data of Beijing-Tianjin-Hebei region as an example. *Journal of Intelligencem, 38*(6), 174–181.
Hu, Z., Tao, R., & Yuan, Y. (2018). Research on performance evaluation index system of scientific and technological achievement transformation based on regional perspective. *Science Technology and Industry, 18*(1), 57–62.
Hua, D., & Jiang, F. (2018). The study on the efficiency of technology transaction from the perspective of subject relationship. *Shanghai Economic Review, 7*, 98–107.
Hussain, M., Wasti, H., Huang, G., Wei, L., Jiang, Y., & Tang, Y. (2020). An approach for measuring semantic similarity between Wikipedia concepts using multiple inheritances. *Information Processing & Management, 57*(3), https://doi.org/10.1016/j.ipm.2019.102188.
Jia, B., Yang, H., Wu, B., & Xing, Y. (2020). Collective entity disambiguation based on hierarchical semantic similarity. *International Journal of Data Warehousing and Mining, 16*(2), 1–17.
Jiang, Y., Bai, W., Zhang, X., & Hu, J. (2017). Wikipedia-based information content and semantic similarity computation. *Information Processing & Management, 53*(1), 248–265.
Khatua, A., Khatua, A., & Cambria, E. (2019). A tale of two epidemics: Contextual Word2Vec for classifying twitter streams during outbreaks. *Information Processing & Management, 56*(1), 247–257.
Kumar, S., & Bhatia, K. (2020). Semantic similarity and text summarization based novelty detection. *SN Applied Sciences, 2*(3), https://doi.org/10.1007/s42452-020-2082-z.
Li, F., Huang, H., & Zeng, L. (2018). Global perspective on technology trading platform construction experience and its revelation and suggestion. *Forum on Science and Technology in China, 1*, 24–29.
Li, C., Lan, T., & Liu, S. (2015). Patent attorney as technology intermediary: a patent attorney-facilitated model of technology transfer in developing countries. *World Patent Information, 43*, 62–73.
Lichtenthaler, U., & Ernst, H. (2007). External technology commercialization in large firms: results of a quantitative bench- marking study. *R&D Management, 37*(5), 383–397.
Lin, Q., & Mao, S. (2019). Research on transformation efficiency of Chinese universities. *Forum on Science and Technology in China, 05* 144-151+162.
Lin, J., Miao, R., & Zhang, Z. (2019). Research on extraction methods of topic knowledge tuples in professional social media. *Library and Information Service, 63*, 101–110.
Link, A., & Hasselt, M. (2019). On the transfer of technology from universities: The impact of the Bayh–Dole Act of 1980 on the institutionalization of university research. *European Economic Review, 119*, 472–481.
Liu, X., Bi, K., & Ye, H. (2016). A study on risk management of radical innovation in low-carbon technologies in the global value chain: A case of the Chinese Manufacturing Industry. *China Soft Science, 11*, 152–166.
Martinez-Gil, J., & Chaves-Gonzalez, J. (2019). Automatic design of semantic similarity controllers based on fuzzy logics. *Expert Systems with Applications, 131*, 35–39.
Mikolov, T., Sutskever, I., Chen, K., et al. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems, 26*, 3111–3119.
Ning, J., & Liu, J. (2016). Using Word2Vec with TextRank to extract keywords. *New Technology of Library and Information Service, 6*, 20–27.
Racharak, T., Suntisrivaraporn, B., & Two, S. (2018). Personalizing a concept similarity measure in the description logic ELH with preference profile. *Computing and Informatics, 37*(3), 581–613.
Rule, A., Cointet, J., & Bearman, P. (2015). Lexical shifts, substantive changes, and continuity in state of the union discourse,1790-2014. *Proceedings of the National Academy of Sciences, 112*(35), 10837–10844.
Shajalal, M., & Aono, M. (2019). Semantic textual similarity between sentences using bilingual word semantics. *Progress in Artificial Intelligence, 8*(2), 263–273.
Soares, V., Campello, R., Nourashrafeddin, S., Milios, E., & Naldi, M. (2019). Combining semantic and term frequency similarities for text clustering. *Knowledge & Information Systems, 61*, 1485–1516.
Sousa, R., Silva, S., & Pesquita, C. (2020). Evolving knowledge graph similarity for supervised learning in complex biomedical domains. *BMC Bioinformatics,* 26. https://doi.org/10.1186/s12859-019-3296-1.
Sun, Y., & Grimes, S. (2017). The actors and relations in evolving networks: The determinants of inter-regional technology transaction in China. *Technological Forecasting and Social Change, 125*, 125–136.
Sun, Y. T., & Liu, K. (2016). Proximity effect, preferential attachment and path dependence in inter-regional network: A case of China's technology transaction. *Scientometrics, 108*(1), 201–220.
Tien, N., Le, N., Tomohiro, Y., & Tatsuya, I. (2019). Sentence modeling via multiple word embeddings and multi-level comparison for semantic textual similarity. *Information Processing & Management, 56*(6), https://doi.org/10.1016/j.ipm.2019.102090.
Vinig, T., & Lips, D. (2015). Measuring the performance of university technology transfer using meta data approach: The case of Dutch universities. *Journal of Technology Transfer, 40*(6), 1034–1049.
Wang, L. (2014). A comparative study on the operation mode of European Innovation Relay Centers and traditional technology transfer services. *Researches in Library Science, 14*, 96–101.
Wu, Z., Zhu, H., Li, G., Cui, Z., Huang, H., Li, J., Chen, E., & Xu, G. (2017). An efficient Wikipedia semantic matching approach to texts document classification. *Information Sciences, 393*, 15–28.
Xiang, F. (2013). Zhejiang Model for the construction of online technology market. *China National Conditions and Strength, 02*, 31–33.
Xiong, H. (2017). Is the transformation rate of scientific research achievements in China really low. *Development Research Center of the State Council*. http://www.drc.gov.cn/xsyzcfx/201703/4-4-2893667.htm.
Xu, H., Dong, K., Liu, C., Wang, C., & Wang, Z. (2017). A review on topic identification of scientific texts files. *Information Science, 35*(1), 153–160.

Xu, Z., Yang, L., & Luan, S. (2019). Research on evaluation of transformation ability and efficiency of scientific and technological achievements in universities based on relay innovation. *Science and Technology Management Research, 39*(24), 8–14.

Yang, D., Xia, Q., & Ma, C. (2017). Efficiency analysis of online technology transfer platform and supply and demand matching. *Management Sciences in China, 30*(6), 104–112.

Ye, J., Yang, Y., Han, J, & Li, C. (2015). Comparison of technology transfer efficiency between Chinese and American universities from the perspective of patent. *China Science and Technology Forum, 1*, 150–155.

Zhou, X., Zhang, Y., & Tang, Z. (2016). Enterprise Europe network collaborative innovation system and its implications to enhance the innovation capability of SMEs in China. *Reform of Economic System, 4*, 174–179.