ICIC Express Letters ICIC International ©2018 ISSN 1881-803X
Volume **12**, Number **6**, June **2018** pp. **599–606**

IDENTIFYING TECHNOLOGY OPPORTUNITIES BASED ON INTERNAL CAPABILITIES AND TECHNICAL SUITABILITY

Junhyeok Park[1], Seung-Jun Shin[2], Pyunghoi Koo[1] and Wonchul Seo[1,*]

[1]Division of Systems Management and Engineering
Pukyong National University
45 Yongso-ro, Nam-gu, Busan 48513, Korea
holicy92@pukyong.ac.kr; phkoo@pknu.ac.kr; *Corresponding author: wcseo@pknu.ac.kr

[2]Graduate School of Management of Technology
Pukyong National University
365 Sinseon-ro, Nam-gu, Busan 48547, Korea
sjshin@pknu.ac.kr

Received November 2017; accepted February 2018

Abstract. *As global technology competition intensifies, identifying technology opportunities in a rapidly changing market environment has become an indispensable factor in the survival of enterprise. There have been many studies on technology opportunities, but most of the studies did not consider internal capabilities of the enterprise or exclude entry barriers to the recommended technology opportunities. Therefore, this study proposes a methodology for analyzing internal capabilities through a patent portfolio held by the enterprise and recommending feasible technology opportunities in terms of technology themes and judging suitability for recommended technology themes by considering entry barriers and technical trends. The results of this study are expected to be helpful for business decision making such as expansion of business area or change of business area by recommending relevant and new technology opportunities.*
Keywords: Topic modeling, Collaborative filtering, Topic trends, Technology themes, Technology opportunity discovery

1. **Introduction.** Recently, due to the intensification of global technology competition, it has become necessary to discover new technology opportunities in order to survive in a rapidly changing market environment and to gain competitive advantage in the market [1]. Since these efforts are the basis for continuous growth as well as for the survival of enterprises, various methods have been proposed by many researchers to discover new technology opportunities. Previous studies have focused on the products owned by the enterprise to identify the technology opportunities [2,3]. However, in order to realize technological innovation, priority should be given to technology that can produce products rather than the products. In this sense, patent data is the most useful document to describe the characteristics of the technology. However, even in the case of using patent data, some previous studies identify technology opportunities by utilizing only International Patent Classification (IPC) or patent citation information [4]. However, IPC and patent citation information are so wide that it is difficult to identify specific technology opportunities. In addition, although identifying technology opportunities, they have not verified whether the discovered technologies are suitable as technology opportunities [5]. If so, there is a risk that the discovered technology field is already in the red ocean, or that it has been turned over by the global market trends. Therefore, in this study, we propose a methodology for recommending new technology opportunities in terms of technology themes through quantitative analysis in consideration of internal capabilities by conducting text mining based on the patent portfolio of enterprise. In addition, we verify that the technology is suitable as an opportunity and finally recommend it to the target

enterprise. The results of this study are expected to contribute to the decision making of business expansion and business area changing by discovering new technology opportunities which are related but not possessed by the enterprise. The organization of this study is as follows. We present an overview of the groundwork in Section 2. The methodology is described in Section 3 and a case study is conducted in Section 4. Finally, Section 5 concludes and summarizes this study, and presents future research directions.

2. Groundwork.

2.1. Latent Dirichlet Allocation (LDA).

Topic modeling is a probabilistic model that estimates topics that are latent in a document. The topic is estimated by clustering words having similar meanings in consideration of the occurrence frequency and relevance of the words in the document. The most widely used model of topic modeling is the LDA model. The LDA model is a topic modeling method that extracts a set of words constituting a topic by calculating the probability that words will be included in a specific topic using the Dirichlet distribution based on the words forming the document, assuming that all the documents are composed of a set of topics [6].

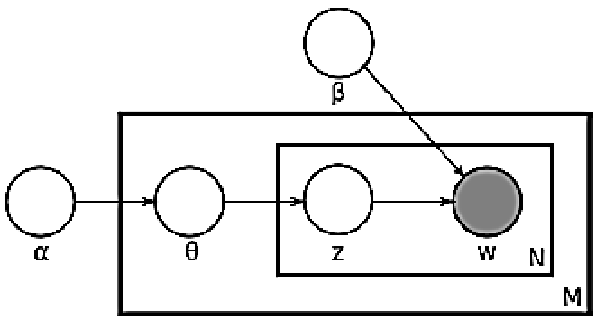

Figure 1. LDA model

Figure 1 shows the process of the LDA model. The LDA model is composed of the number of documents (M), the number of words per document (N), the word (w) in the document, topic (z), the ratio of a topic (θ), parameter (α) for the Dirichlet prior probability of topics distribution and parameter (β) for the Dirichlet posterior probability of words distribution in a topic [7]. Figure 1 shows a generative probabilistic model for creating a document by selecting topics and words from (α) and (β). Topic modeling is a technique for estimating a topic from a written document by going through this process in the reverse direction. LDA-based studies have mainly focused on analyzing research trends related to science and technology by utilizing patent data forecasting technology using topic based patent analysis [8] and identifying trends of convergence technology by using topic modeling and cross-impact analysis [9]. In this study, we use the LDA model to obtain technology information from patent data held by enterprise. Topics extracted from patent data using topic modeling are defined as 'technology themes'. Because the topics are key words in the documents, the technology themes are also the key words in patent documents. In other words, the technology theme is a concept that describes in detail the technology that an enterprise has. These technology themes contain more detail data about technology information than IPC or patent citation information. This study uses the LDA model that was performed by R to obtain technology information of the enterprise.

2.2. Collaborative filtering.

Collaborative filtering is a technique for predicting the preference of a target customer for a specific product by comparing the products preferred by customers using the matrix shown in Figure 2, assuming that customers with similar product preferences will have similar preferences for other products [10]. It suggests products suitable for the target customers through this assumption. It is also possible to convert preferences to binary values for product purchase, and to recommend products based on what products are purchased by customers. Therefore, many collaborative filtering-based studies have been conducted to recommend customized products or contents including developing a recommendation system that places or activities that utilize user-based mobile information [11] and proposing a personalized news recommendation system based on click behaviors [12].

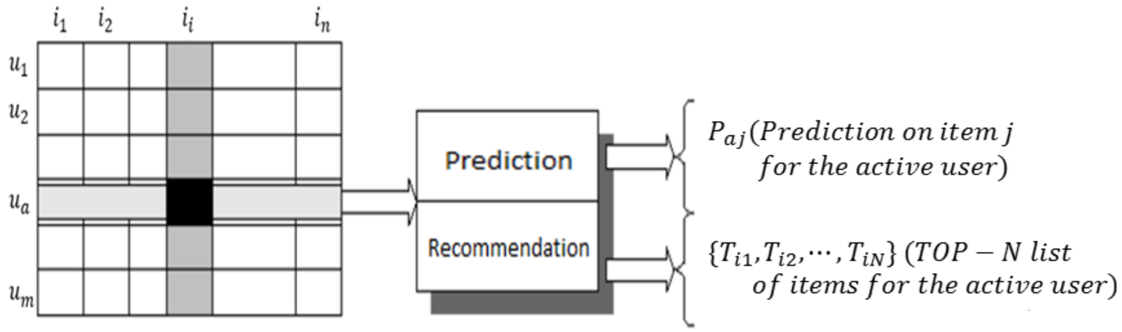

FIGURE 2. Collaborative filtering process

2.3. **Topic trends.** Topic trends is a method to analyze the change trend of topics by period, and a topic with a high rate of increase is defined as a 'hot topic' and a topic with a high rate of decrease is defined as a 'cold topic' [13] such as finding topic trends in digital libraries [14]. In this study, we analyze the topic trends by year and verify that the recommended technology themes are appropriate as technical opportunities. If the trend shows a decrease, the technology theme is excluded from the technology opportunity candidates.

3. **Methodology.** The study proposes a methodology to recommend technology themes as shown in Figure 3.

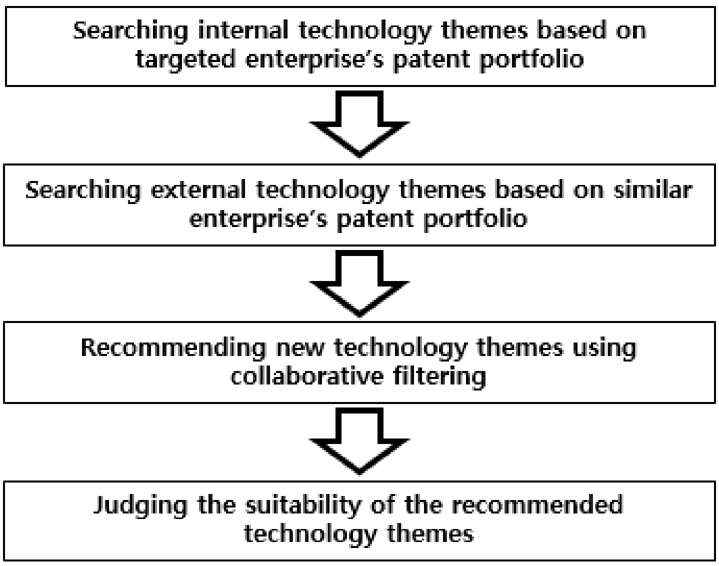

FIGURE 3. Conceptual framework for recommending technology themes

3.1. **Searching internal technology themes based on target enterprise's patent portfolio.** The first step is searching the internal technology themes based on the capabilities of the target enterprise. Patent data of the target enterprise is collected to search the internal technology information. After that, extract the summary of the patent data and construct the corpus to conduct the LDA. In this case, the topics extracted through the LDA are defined as the internal technology themes.

3.2. **Searching external technology themes based on similar enterprise's patent portfolio.** The second step is finding enterprises similar to the target enterprise and searching external technology themes related to the technologies possessed by the similar enterprises. First, we search patents based on the internal technology themes searched in Section 3.1, and the applicants of the patents are extracted and defined as similar enterprises. After that, we collect the summary of patent data held by the similar enterprises, construct the corpus, and perform the LDA for each enterprise. The topics extracted through the LDA are defined as external technology themes.

3.3. **Recommending new technology themes using collaborative filtering.** The third step is recommending external technology themes that are new to the target enterprise through collaborative filtering. In this study, we define a customer as a similar enterprise, an item as a technology theme, and assign 1 if the enterprise owns the corresponding technology theme, or assign a 0 to construct a transaction matrix. And the transaction matrix is used to find similar enterprises and get technology opportunities. In order to generate the transaction matrix, the similarity between the internal technology themes and the external technology themes is calculated, and the external technology themes with high similarity to the internal technology themes are regarded as the same as the internal technology themes. Then, the similarity between the external technology themes with low similarity to the internal technology themes is also calculated, and repeat the process of reducing the dimension by combining similar external technology themes to finally generate a formalized transaction matrix as shown in Figure 4.

If the collaborative filtering is performed using the transaction matrix shown in Figure 4, the technology themes that the target enterprise does not possess will be recommended.

Transaction Matrix												
	Internal Technology Themes				External Technology Themes							
Target enterprise	1	1	0	1	0	0	0	0	0	0	0	0
Similar enterprises	1	1	0	1	0	0	0	0	1	0	0	1
	0	0	0	1	0	0	1	0	0	0	1	0
	1	1	0	0	0	1	0	0	0	1	0	0
	0	1	0	1	0	1	0	1	0	0	0	0
	1	0	0	0	0	1	1	0	1	1	1	0
	0	1	0	0	0	1	1	0	0	1	0	0
	0	1	1	1	1	0	0	1	0	0	0	0
	1	1	1	1	0	0	0	1	0	0	0	0
	1	1	1	0	0	1	0	0	1	1	0	0
	1	1	0	0	0	0	0	1	0	0	1	0
	0	1	1	1	0	0	0	0	0	0	1	0
	0	1	0	1	0	0	1	0	0	0	0	1
	1	0	0	1	1	1	0	1	1	0	1	0

FIGURE 4. Transaction matrix

3.4. **Judging the suitability of the recommended technology themes.** The fourth step is analyzing the topic trends in order to find out which of the technology themes recommended through collaborative filtering in the third step are the suitable technology fields in the current global market trend. If the recommended technology theme does not follow the current global market trend, then the technology theme is not a technology opportunity. Therefore, we perform topic trend technique to verify if there is a possibility of growth of the recommended technology theme. In this study, we search the recommended technology theme by year and create a linear trend formula with the number of patents retrieved. Then, the linear trend coefficient of the linear trend formula is defined as a Trend Index (TI). Also, if TI is negative, the technology theme is not likely to grow and is excluded from the technology opportunity candidate by assigning 0 to TI. However, some of the technology opportunity candidates may be showing an increasing trend due to intensified competition in technology. In other words, even if the topic trend shows an increase, it may be because it is a field where competition is already overheated. Thus, it is necessary to identify the entry barriers of the recommended technology theme by using the transaction matrix created in the previous step. Therefore, in this study, we identify the number of enterprises that have recommended technology themes as compared to the total number of similarity enterprises in the transaction matrix, and the higher the ratio is, the higher the entry barriers are. Therefore, the reciprocal of the ratio described above is defined as Entry Barrier Index (EBI) as shown in Equation (1). Finally, the value obtained by multiplying TI by EBI is defined as Suitability Index (SI) as shown in Equation (2), and the higher the SI is, the more suitable it is as a technology opportunity for the target enterprise. Additionally, if SI is less than 1, it is determined that it is not suitable as a technology opportunity, and it is finally excluded in the candidate for technology opportunity.

$$\frac{\text{Total number of similarity enterprises in the transaction matrix}}{\text{The number of enterprises that have recommended technology themes}} \tag{1}$$

$$\text{Suitability Index (SI)} = \text{Trend Index (TI)} \times \text{Entry Barrier Index (EBI)} \tag{2}$$

4. Case Study.

4.1. **Searching internal technology themes based on target enterprise's patent portfolio.** In this study, we perform a case study to explore the practical applicability of the proposed methodology. We select a target enterprise called 'A.motor'. And we collect A.motor's 4,752 patent summary data registered in United States Patent and Trademark Office (USPTO) during 2011 and 2016. We search the patent data from KIPRIS (http://www.kipris.or.kr). And then, by performing LDA, we extract 80 topics and create the patent retrieval query with combination of top 5 terms being high weight in 10 terms of constituting the topic. These topics are defined as internal technology themes and Table 1 is an example for showing a part of patent retrieval query.

Table 1. Part of the patent retrieval query

Topic 1	airbag	cushion	inflator	gas	passenger
	pad	bag	deployed	collision	vent
Topic 2	moving	units	clamping	cart	common
	jig	vehicles	clamp	types	running
Patent retrieval query	((airbag and cushion and inflator and gas and passenger) or (moving and units and clamping and cart and common))				

4.2. **Searching external technology themes based on similar enterprise's patent portfolio.** We collect patent data through the query created in step 4.1. And the data are classified by the applicant. Then, we retrieved a total of 727 applicants based on the query generated in 4.1, and extract 52 applicants with more than 10 common technology themes. These applicants are defined as similar enterprises. We retrieve 208,060 patent data from 52 similar enterprises and extract 4,160 external technology themes by performing a new LDA.

4.3. **Recommending new technology themes using collaborative filtering.** We use Python's Gensim library to calculate the similarity between internal and external technology themes. If the similarity is more than '0.69', it is considered to be the same as the internal technology theme. (0.69 is the average of similarities between the most similar technology themes.) After that, we repeat the process of calculating the similarity of the other external technology themes and merging similar external technology to reduce the number of dimensions of the transaction matrix. As a result, we compress 4,160 external technology themes to 225. The adjusted transaction matrix is shown in Table 2. And 10 external technology themes are recommended using collaborative filtering as shown in Table 3.

Table 2. Part of the transaction matrix

	ex220	ex221	ex222	ex223	ex224	ex225
se47	0	0	0	0	0	0
se48	0	0	1	1	0	0
se49	1	0	1	0	1	0
se50	0	0	1	1	0	1
se51	0	1	0	0	1	0
se52	0	0	0	1	0	1

Table 3. Recommended list

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
TE	ex37	ex57	ex44	ex12	ex07	ex23	ex22	ex18	ex42	ex43

4.4. **Judging the suitability of the recommended technology themes.** In order to judge whether the technology themes recommended in step 4.3 are suitable as technology opportunities, we identify the number of patent data searched for each year based on the recommended technology theme. Then, we calculate the TI and EBI of the recommended 10 external technology themes (Table 4). Finally, it is judged that 6 external technology theme with SI more than 1 is suitable as technology opportunities. As a result, it can be seen that 'A.motor' needs to develop technology for smart car by combining ex22, ex37 and ex57 and the combination of ex12, ex44, and ex07 confirms the need for eco-friendly car development.

5. **Conclusion.** In this study, we propose a new technology themes recommendation method based on internal capabilities of enterprise. By extracting technology themes using the LDA technique, we can get technology themes, which is more detailed information than IPC or patent citation information. Moreover, by using collaborative filtering technique, it is possible to calculate the cosine similarity by vectorizing the technology themes possessed by similar enterprises, so that a customized recommendation system can be applied. Also, it is possible to recommend unexpected external technology fields. It is a similar company, but it may have a completely different technology theme. In other

TABLE 4. Result of calculating TI, EBI and SI

Recommended external technology theme	TI	EBI	SI	Remarks
ex22	32.43	2.60	84.32	
ex07	13.26	2.36	31.33	
ex57	7.89	3.47	27.34	
ex12	2.74	2.89	7.92	
ex44	2.80	2.36	6.62	
ex37	1.51	2.60	3.94	
ex23	0.31	1.93	0.61	
ex18	0.00	2.08	0.00	(−0.0857)
ex42	0.00	2.36	0.00	(−1.6571)
ex43	0.00	2.74	0.00	(−1.7429)

words, by identifying the unique technology fields possessed by similar enterprises, we can identify unexpected technology opportunities. Therefore, in this study, we use collaborative filtering technique to find more suitable and potential technology opportunities. Finally, this study suggests using the topic trend technique to increase the reliability of recommended technology opportunities by analyzing the degree of competition for the recommended technology which is not reflected in the previous studies. As a result, it is anticipated that it will be possible to provide a way to reduce effort, time and cost in expanding a new business area or technical planning of an enterprise. Despite the contributions of this study, challenges for future research still remain. First, in order to construct the adjusted transaction matrix, we repeat the process of reducing the dimension by calculating the similarity between the internal technology themes and the external technology themes. In this case, as the number of similar enterprises increases, the number of times to calculate similarity increases exponentially. Therefore, future research should propose a way to appropriately limit the number of similar enterprises. Second, the element value of the transaction matrix for using collaborative filtering can be only 0 or 1 according to the existence of the technology themes. Therefore, even if an enterprise has a technology theme, it is difficult to understand what the technology focuses on, so it is naturally difficult to fully reflect the actual internal capabilities of the enterprise. Therefore, if the preference analysis is performed by reflecting the weights of technology themes in the future research, it will be possible to find more suitable technology opportunities based on the internal capabilities of the enterprise.

Acknowledgment. This work was supported by a Research Grant of Pukyong National University (2016 year; C-D-2016-1316).

REFERENCES

[1] R. A. D'Aveni, *Hypercompetition: Managing the Dynamics of Strategic Maneuvering*, Free Press, New York, 1994.

[2] W. Seo, J. Yoon, H. Park, B.-Y. Coh, J.-M. Lee and O.-J. Kwon, Product opportunity identification based on internal capabilities using text mining and association rule mining, *Technological Forecasting and Social Change*, vol.105, pp.94-104, 2016.

[3] J. Yoon, W. Seo, B.-Y. Coh, I.-S. Song and J.-M. Lee, Identifying product opportunities using collaborative filtering-based patent analysis, *Computers and Industrial Engineering*, vol.107, pp.376-387, 2017.

[4] H. Park, T. Ryu, J. Yoon, S. Koh and W. Seo, A study on evaluation model for traditional knowledge using patent portfolio analysis, *ICIC Express Letters*, vol.9, no.4, pp.1063-1068, 2015.

[5] J. Kim, W.-S. Lee, S. Choi and W. Seo, Identifying potential technology themes based on internal capabilities using topic modeling and association rule mining, *ICIC Express Letters, Part B: Applications*, vol.7, no.10, pp.2267-2273, 2016.

[6] D. M. Blei, A. Y. Ng and M. I. Jordan, Latent Dirichlet allocation, *The Journal of Machine Learning Research*, vol.3, pp.993-1022, 2003.
[7] J. Fu, N. Liu, C. Hu and X. Zhang, Hot topic classification of microblogging based on cascaded latent Dirichlet allocation, *ICIC Express Letters, Part B: Applications*, vol.7, no.3, pp.621-625, 2016.
[8] G. J. Kim, S. S. Park and D. S. Jang, Technology forecasting using topic-based patent analysis, *NISCAIR*, vol.74, pp.265-270, 2015.
[9] B. Jeong, J. Kim and J. Yoon, A semantic patent analysis approach to identifying trends of convergence technology: Application of topic modeling and cross-impact analysis, *The Journal of Intellectual Property*, vol.11, no.4, pp.211-240, 2016 (in Korean).
[10] Y. Park and J. Yoon, Application technology opportunity discovery from technology portfolios: Use of patent classification and collaborative filtering, *Technological Forecasting and Social Change*, vol.118, pp.170-183, 2017.
[11] V. W. Zheng, B. Cao, Y. Zheng, X. Xie and Q. Yang, Collaborative filtering meets mobile recommendation: A user-centered approach, *Association for the Advancement of Artificial Intelligence*, vol.10, pp.236-241, 2010.
[12] J. Liu, P. Dolan and E. R. Pedersen, Personalized news recommendation based on click behavior, *Proc. of the 15th International Conference on Intelligent User Interfaces*, Palo Alto, CA, USA, pp.31-40, 2011.
[13] T. L. Griffiths and M. Steyvers, Finding scientific topics, *Proc. of the National Academy of Sciences*, vol.101(suppl 1), pp.5228-5235, 2004.
[14] L. Bolelli, S. Ertekin, D. Zhou and C. L. Giles, Finding topic trends in digital libraries, *Proc. of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries*, Austin, USA, pp.69-72, 2009.