Technological Forecasting & Social Change 119 (2017) 170–183

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Identifying new business opportunities from competitor intelligence: An integrated use of patent and trademark databases

Mingook Lee [a], Sungjoo Lee [b,*]

[a] Future Information Research Center, Department of Technology Intelligence Research, Korea Institute of Science and Technology Information, Seoul, Republic of Korea
[b] Department of Industrial Engineering, School of Engineering, Ajou University, Suwon, Republic of Korea

ARTICLE INFO

Article history:
Received 10 August 2016
Received in revised form 17 March 2017
Accepted 20 March 2017
Available online 1 April 2017

Keywords:
Business opportunity
Patent
Trademark
Competitor intelligence
Technology-based

ABSTRACT

This study aims to analyze the position of technology-centered companies in complex market dynamics and discover new business opportunities from competitor intelligence. For this, we consider both technology and market characteristics in providing competitor intelligence by utilizing patent data as a representative proxy for a firm's technology, and trademark data as an information source for the firm's target goods and services. To analyze the two types of data, a collaborative filtering approach together with portfolio analyses and association mining techniques were adopted. Theoretically, this is one of the earliest attempts to combine patent data and trademark data to investigate corporate strategies. In practice, the research results are expected to be used as a decision criterion to diagnose the economic value that companies can obtain by entering the market, as well as the technological value to be passed onto their customers. Thus, the proposed approach can be useful to support effective technology and business strategies in a firm.

1. Introduction

With the increased development of technology and accelerated globalization, companies are encountering fierce competition (Archibugi and Iammarino, 2002). For these companies to not only survive but also thrive in a market, they need to continue to capture new business opportunities (Stevenson et al., 1989). Accordingly, discovering new business opportunities has attracted a great deal of interest in academia and industry (Venkataraman, 1997). It is a creative activity that involves combining various resources to offer superior value to a market (Casson, 1982; Schumpeter, 1994), which requires considering various factors and, accordingly, has been driven mainly by qualitative processes greatly affected by entrepreneurial alertness or prior knowledge (Ardichvili et al., 2003). However, there also exists a quantitative approach to identify new business opportunities, recognizing the value of information available to investigate the market, technology, and competitor trends. Among various sources of information used for this purpose, patent documents have been regarded as one of the most significant sources due to their amount and diversity of information. Related studies have tried to identify emerging technology (Daim et al., 2006) or vacant technology (Choi and Jun, 2014) from patent information, on which further R&D with business development is likely to take place. Other studies proposed benchmarking business areas by comparing corporate patent portfolios (Fabry et al., 2006). These approaches are valuable particularly for technology-based products and services.

Despite meaningful contributions, however, existing studies have two limitations. First, they tend to rely mostly on patent information, which restricts their ability to suggest new business opportunities directly. How to link new technology opportunities to new business opportunities has seldom been discussed in the literature. Second, most existing studies have concentrated on new technology-based, rather than application-based, business opportunities; few efforts have been made to find a way to apply existing technologies to new application areas. Business opportunities can result from applying existing technologies to new markets as well as identifying and developing new technologies, which is especially important for small and medium-sized enterprises (SMEs).

Of course, there exist some exceptions. For example, Lee et al. (2009a, 2009b) proposed a technology-based roadmap in which business opportunities are identified and linked to technology planning; they developed an algorithm to identify competitors having similar patent portfolios to a focal firm and argued that the business areas of these competitors can be a new business opportunity worth benchmarking. However, their study is limited in suggesting a new business opportunity directly. Yoon et al. (in press) suggested a method to identify new product opportunities using a collaborative filtering-based patent analysis. Though this is a meaningful attempt to directly guide product items, the study only investigates a particular technology; a new business opportunity may come from a combination of technologies rather

* Corresponding author at: Worldcup-ro 206, Woncheon-dong, Yeongtong-gu, Suwon-si, Gyeonggi-do 16499, Republic of Korea.
E-mail addresses: mingook@kisti.re.kr (M. Lee), sungjoo@ajou.ac.kr (S. Lee).

http://dx.doi.org/10.1016/j.techfore.2017.03.026
0040-1625/

than a single technology. Further analysis is therefore needed to search for new business areas that are suited to corporate technology capabilities, which should be evaluated on the basis of corporate technology portfolios.

Therefore, this study aims to propose a novel approach to identifying new business opportunities based on corporate technology capabilities, for which a collaborative filtering-based competitor analysis is adopted. Collaborative filtering is a process of finding items or users with similar information or patterns and is frequently used for making recommendations. In this study, collaborative filtering is applied to patent information for identifying firms having similar technology characteristics, with a focal firm among those with superior technology capabilities in the business field of interest. Then, association mining and text mining are applied to trademark information in order to analyze competitors with similar technology portfolios and furthermore to identify promising new business areas systematically. This is one of the earliest attempts to combine patent and trademark information for extracting meaningful business intelligence. Patent information has long been the focus of competitor analysis, while trademark information has gained little interest despite its potential advantages of revealing corporate ongoing business areas or business areas of interest. If patent information showing technological characteristics and trademark information indicating business characteristics are combined and well analyzed, they can be valuable sources of competitor intelligence.

The remainder of this paper is structured as follows. Section 2 introduces the existing studies and the characteristics of patent and trademark data. Section 3 explains the approach to identifying new business opportunities suggested in this study. The approach is applied to the cloud computing industry, which is a representative technology-based service industry, to illustrate how the approach can be used to create value; these case study results are described in Section 4. Finally, Section 5 discusses limitations and future research directions.

2. Relevant studies

2.1. New business opportunities

Studies on new business opportunities have been conducted mainly in two streams. The first is opportunity identification and exploitation in the context of *startups*. The focus of this stream has been on opportunity recognition capabilities; a number of researchers have identified the factors affecting those capabilities, which include creativity (Shane and Nicolaou, 2015), cross-cultural experience (Vandor and Franke, 2016), prior business ownership experience (Ucbasaran et al., 2009), diversity of information (Gielnik et al., 2012), and gender differences (Gupta et al., 2014). Other studies have emphasized the possibility of creating opportunities in the network of firms, addressing the importance of the ecosystem in business opportunities (Nordman and Tolstoy, 2016; Overholm, 2015; Palo and Tähtinen, 2013). The second stream investigates new business opportunities in the context of *established firms*, where a framework to analyze business opportunities or a supporting system to identify such opportunities has been proposed. This study is in line with the second research stream because established firms are more likely to benefit from the competitor intelligence approach suggested in this study.

In general, market-oriented management was regarded as a core driver of corporate performance. Hence, most existing frameworks for analyzing business opportunities have been based on market orientation. According to Narver and Slater (1990), market orientation perspectives consist of customer orientation, competitor orientation, and interfunctional coordination; companies can obtain long-term benefits from considering each of the three perspectives. Among them, this study takes the competitor orientation perspective. Competitor orientation highlights the activities of collecting information to understand and analyze the products and strategies of competitors and benchmarking them (Armstrong and Collopy, 1996). Awareness of competitor market strategies enables the effective management of target markets (Fahey, 2001; Peyrot et al., 2002; Wu and Olk, 2014). For this purpose, previous studies have developed an analytic framework for investigating competitor products and strategies (Porter, 1980), or a method of presenting a competitor landscape to be used for market segmentation (Söllner and Rese, 2001). These competitive intelligence approaches are regarded as a useful tool for finding new business opportunities under the circumstances of blurring industry boundaries (Bröring et al., 2006) and broadening the range of competitors (Hoopes et al., 2003).

The approaches to identifying new business opportunities from competitive intelligence can be divided largely into two categories – technology-based and market-based – according to the type of information used for intelligence activities. One of the most representative approaches to technology-based competitive intelligence is patent analysis (Ashton and Sen, 1988; Lee et al., 2009a, 2009b). It strives to open new business opportunities via technology, as technology-based companies dominate market and R&D capabilities that determine corporate competitiveness. In addition, innovative technology leads product development, which justifies the use of technology information for identifying business opportunities. Sometimes, patent data is used together with other types of data to gain a deeper understanding of technologies that may influence business decisions. For example, Geum et al. (2010) used both patents and publications to identify and evaluate strategic partners for collaborative R&D. Taking a holistic approach, Lee et al. (2012) analyzed the techniques and tools available for discovering competitive intelligence from various types of technical documents. At the same time, patent analysis methods have been diversified and used for technology opportunity analysis. For example, Seol et al. (2011) adopted two techniques – data envelop analysis and text mining – to search and evaluate new business areas. Lee et al. (2015) suggested novelty-focused patent mapping for investigating technology opportunities. Indeed, patent analysis is an asset to competitive intelligence. However, technology information is limited because the real situation of competition in a market regarding actual products and services – business areas – is rarely captured within technology information. Technology can be applied to multiple products and services, which are determined by commercialization environments as a collection of the microeconomic and strategic conditions a company encounters (Gans and Stern, 2003).

On the other hand, one of the most basic approaches to market-based competitive intelligence is to identify types of products and services and to analyze the market share of competitors and its trends (Hanssens, 1980). This information has been commonly used to understand the competitive position in a market and establish a resource management strategy, which can be achieved not only from quantitative analysis but also from the evaluation of customers or experts, and thus risks subjectivity. To mitigate the possible bias from such subjectivity, recent efforts have been made to use a large set of customer data available online for competitor analysis. For example, Jin et al. (2016) used product online reviews to analyze customer requirements for competitive products when implementing a new product design. Similarly, He et al. (2013) collected social media data and applied a text-mining technique for competitive analysis. Some researchers have addressed the significance of a dynamic framework for competitor analysis; Peng and Liang (2016) described the competitive dynamics of interfirm rivalry based on the framework they proposed to identify competitors.

Recognizing the value of technology-based and market-based intelligence, we suggest the use of both patent and trademark data – the former is for technology-based competitive intelligence while the latter is for market-based competitive intelligence – and develop a novel approach to identify business opportunities by investigating a company's technology characteristics, as well as the market characteristics of the products and services the company offers.

2.2. Patent and trademark information

As the number of companies placing a high value on intangible assets – such as brand recognition – increases (Reitzig, 2004), the potential value of patent and trademark data, which contain a diverse range of information regarding two representative intangible assets, increases, but is yet to be fully explored.

A patent is a means to secure economic benefit by preventing others' use of an invention. Thus, it shows the R&D investment in a firm and, consequently, the firm's technological capabilities and R&D strategies (Greenhalgh and Rogers, 2010; Jaffe, 1989). Patented inventions are classified based on the international patent classification (IPC) system and include data regarding patent applications, patent registrations, citations, claims, and family patents. Relying on these data, various analyses were conducted to investigate technology innovation patterns in a firm (Albert et al., 1991; Austin, 1993; Lanjouw et al., 1998; Tong and Frame, 1994). Moreover, the recent methodological advances in patent analysis contribute to the increased value of patent analysis results. Patent analysis methods have evolved from bibliometric analysis, using structured data, to keyword analysis, using unstructured data (Madani and Weber, 2016). Visualization methods, as well as analysis methods, have been of concern (e.g. Niemann et al., 2017). Various pattern analysis methods developed in data mining have been applied to patent data for investigating major technology trends or core R&D topics (e.g. Venugopalan and Rai, 2015).

On the other hand, a trademark, followed by patents with respect to significance as intellectual property, is used to protect market assets (Doern, 1999). Like a patent, a trademark is also classified by an international classification system, called the Nice Classification (NCL). According to the Nice agreements, a trademark is assigned to designated goods and services – types and detailed items – for which it will be used. Therefore, trademark data are worth analyzing to evaluate the marketing innovation with which goods and services are differentiated in a market (Sandner and Block, 2011). Despite abundant information in trademark documents, existing studies on business opportunity identification have mainly focused on patent documents. Patents and trademarks help indicate technological and marketing innovation activities in addition to their original purpose of granting exclusive rights (Zhou et al., 2016). While patents contain time-series information about technology development activities, trademarks include time-series information regarding direct or indirect market entrance activities. These two set of databases can link technology, markets, and companies, offering meaningful information about how a firm's interests in technologies and markets, as well as relevant activities, have evolved over time.

Therefore, this study adopts the two databases of patents and trademarks, aiming to explore new business opportunities considering the technologies that may influence markets together with the relevant strategies of competitors. The integrated use of the two databases is expected to have the following advantages over existing approaches. First, the two databases are easy to access and provide up-to-date information, increasing the feasibility and usability of the suggested approach. Second, by analyzing both technology-related and market-related information, a balanced view can be maintained to explore business opportunities, which improves the reliability of the suggested approach. Finally, when the two databases are integrated, they can provide a diverse variety of information, ranging from markets to technologies, and thus enable rich analysis results.

2.3. Collaborative filtering

Collaborative filtering is a technique to systematically predict the interest of a particular user based on preference information collected from a large number of users (Su and Khoshgortaar, 2009), assuming that users' past trends will continue in the future (Goldberg et al., 2001). One of the most distinguishing characteristics of this technique is that it uses data from a number of users rather than a particular user. That is, it identifies users having similar patterns with respect to preference or interest on the basis of their expressed inclinations. Such analysis results are useful for cross-selling, where a product or service that is not purchased by a particular user but is purchased by other users having similar tastes is recommended. They may also be used to recommend a product or service related to customer tastes or lifestyles classified by the technique.

In general, this user-based collaborative filtering process is conducted in two steps (Bhatnagar, 2016, p. 133). In the first step, the patterns of customers are used to classify them into several groups with similar patterns. In the second step, the behavior of new customers is predicted by that of other customers in the same group on the assumption that they will behave similarly in the future. In this study, the technique is used to identify new business opportunities. More specifically, competitors having similar innovation patterns with a focal firm, measured by the similarity of their patent-based technology portfolios, are analyzed. These firms are regarded as firms having similar technological interests. The business areas in which the selected firms are interested are then suggested as the most promising. Here, unlike the existing collaborative filtering that takes an exploratory approach to predict future behaviors, this study adopts a normative approach to predict "desirable" future behaviors; thus, instead of using information from every similar firm, this study selects firms with superior technological capabilities that are thus worth benchmarking among firms having similar technology portfolios, which requires another step. Hence, a three-step collaborative filtering rather than the conventional two-step collaborative filtering is suggested in this study.

3. Research framework

In this section, the overall research process, which integrates patent and trademark data to identify promising business areas for a particular firm, is explained.

3.1. New business opportunities

To develop a research framework, it is necessary to define a new business opportunity. In this study, we focus on two types of business opportunities. The first type is *market-penetration* – business areas in which a firm has not been interested but is likely capable of entering with its current technology portfolio. That is, these are business areas where firms with similar technology portfolios operate but a focal firm does not. However, these areas tend to be already occupied by competitors; a market has already been established or is even mature and there might be dominant market leaders. The second type is *market-expansion* – business areas in which a focal firm has not been interested but can enter with its technology strength over its competitors. These are business areas that are closely related to existing areas in which a focal firm has a comparative advantage over its competitors with similar technology portfolios but does not operate. If these

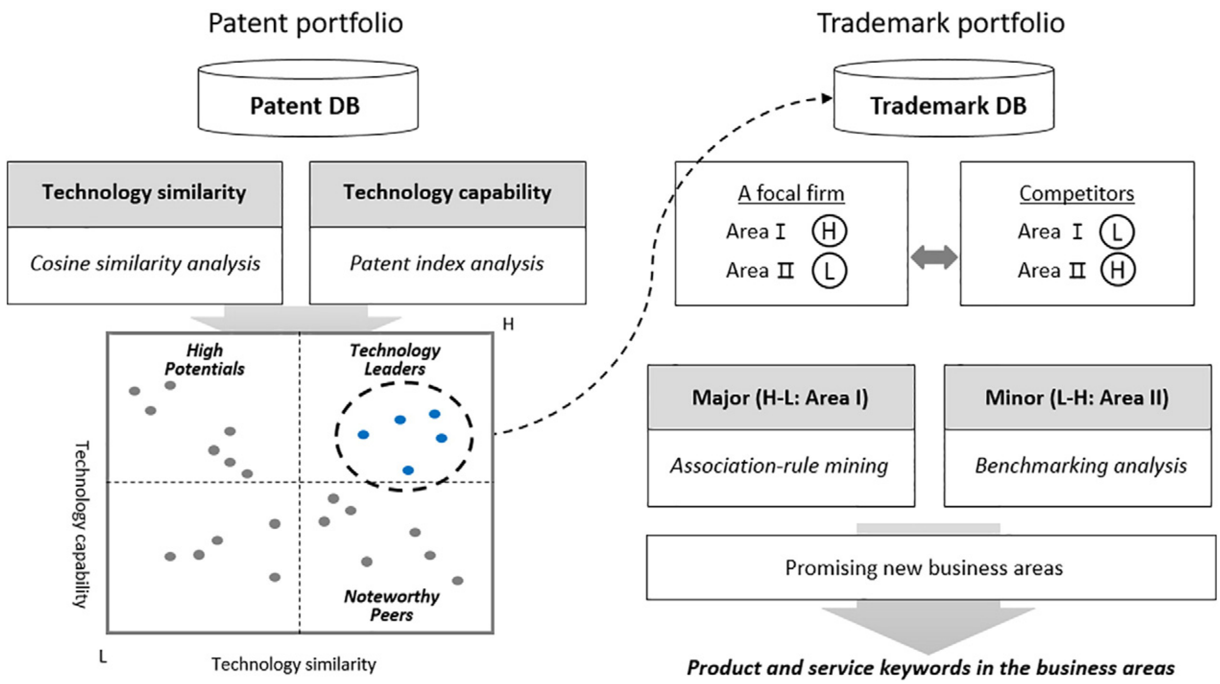

Fig. 1. Research framework.

business areas have not been explored by competitors in the current industry sector, they are worth considering as new business opportunities.

Fig. 1 shows the overall research framework based on the concept of new business opportunities in this study. Here, two types of portfolios – patent portfolio and trademark portfolio – are developed in a successive manner. The first portfolio is established based on patent data; it identifies competitors for further analysis, where the focus is on a company with *superior technology capabilities* and *overlapping technology interests* from the perspective of a focal firm. A group of such companies is called *Technology Leaders*. These companies are likely to be technology leaders in the areas of concern, and thus are the most significant source of competitor intelligence. Concentrating only on this group, the second portfolio using trademark data is developed to identify market-penetration and market-expansion opportunities.

However, it should be noted here that two other groups are also worth considering. One of these is a group of companies with *similar technology interests* but *low technology capabilities*, called *Noteworthy Peers*. Their current technology level may be relatively low but can be improved by acquiring new leading technology, recruiting experts, or merging with other companies. Since these companies can be perceived as direct competitors due to their technology similarity with the focal firm, their business strategies need to be monitored. The other is a group of companies with *high technology capabilities* but *low technological similarity*, called *High Potentials*. The core technology interests of these firms may be different from those of the focal firm. Nevertheless, their superior technology capabilities imply the possibility of broadening their technology scope to provide greater variety to the market. Thus, they can be a potential threat and demand attention. However, the scope of this study is limited to the Technology Leaders group because alternate approaches to monitoring are needed for the other groups.

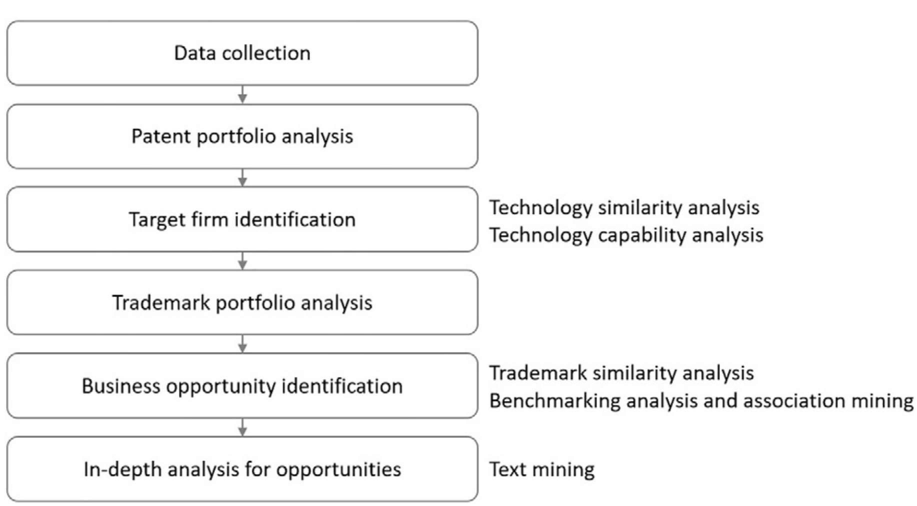

Fig. 2. Research process.

Table 1
Patent portfolio vectors.

Firm	IPC_1	IPC_2	…	…	IPC_n
Firm A	PA_1	PA_2			PA_n
Firm B	PB_1	PB_2			PB_n
…	…				…

3.2. Overall research process

The overall research process consists of the following steps (see Fig. 2). In the first step, patent and trademark data are collected for analysis from triadic patent offices, which include the United States Patent and Trademark Office (USPTO), European Patent Office (EPO), and Japanese Patent Office (JPO). Considering the significance of triadic patents, these three patent offices are expected to provide the most significant information regarding corporate innovation and business activities. Accordingly, technology-oriented firms possessing patents and trademarks in those three patent offices can benefit the most from the proposed approach. In the second step, the collected patent data are used to identify target firms for benchmarking. Competitors having similar and superior technologies are selected as targets for further analyses. In the third step, focusing only on the target firms, their trademark data are collected to investigate a firm's major and minor business areas. For minor business areas, a firm may find a way to explore opportunities through *market penetration strategies*, as other firms having similar technology capabilities are commonly interested in those areas. For major business areas, a firm may find a way to expand its market based on its strength, exploiting *market expansion strategies*. Here, the association rule mining technique is used to help identify new business opportunities that are related to a firm's main business areas. The technique identifies closely-related items and is thus appropriate for finding new business opportunities that are related to a particular business area. Finally, further analysis is conducted on the business areas by applying a text-mining technique to the products and services described in their trademarks. The analysis results show more detailed information regarding products and service items that a firm can consider as new business opportunities.

3.3. Detailed procedures

This section provides a detailed description of each step of the process.

3.3.1. Patent portfolio analysis and target firm identification

A patent portfolio analysis is conducted to identify firms with similar patterns in the process of collaborative filtering. Here, the IPC system is used to measure *technology similarity* between firms: a technology vector that has IPC codes as vector dimensions and the number of patents corresponding to the IPC codes as vector values is constructed, which is referred to as a patent portfolio in this study (see Table 1); vector similarity indicates technology similarity. As shown in Eq. (1), cosine similarity is used due to the focus on the similarity of technology portfolios.

$$Similarity(Firm\ A, Firm\ B) = \cos(\theta) = \frac{A \bullet B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} PA_i \times PB_i}{\sqrt{\sum_{i=1}^{n} (PA_i)^2} \times \sqrt{\sum_{i=1}^{n} (PB_i)^2}} \quad (1)$$

In addition to the technology similarity analysis, *technology capability* is evaluated by four patent indexes as a process of normative collaborative filtering. To do this, the indexes developed to measure (a) technology profitability, (b) technology impact, (c) technology applicability, and (d) technology competiveness were adopted from the work by Ernst (2003) but modified for this study.

First, the *technology profitability* (*TP*) index measures the degree of profits expected from the technology and is evaluated by the number of family patents. The number of family patents can be a proxy of technology economic quality (Ernst, 2003). According to the territorial principle, a firm should grant a patent right for all countries in which it wants to protect its invention. Considering that international patent applications are costly (Harhoff et al., 2003), a firm tends to apply for a patent only in countries expected to provide a commercial benefit or threaten technology competition. Therefore, a large number of family patents indicate either great necessity or the possibility of overseas market development (Grupp and Schmoch, 1999), and relatively great profits expected from the technology.

$$TP(k) = \frac{\text{the average number of family patents for patents granted by firm K}}{\text{the average number of family patents for all patents}} \quad (2)$$

Second, *technology impact* (*TI*) is measured by patent citation frequency. Citation information provides meaningful knowledge on technological significance and innovation impact (Verspagen et al., 2005). Unlike the number of patents, which indicate the quantity of technological innovation activities, citation information indicates the quality of those activities. If a patent is frequently cited by other patents, it indicates that the patent has made a significant contribution to subsequent technology development activities. Hence, if a firm possesses a number of highly cited patents, the firm is likely to have high-quality core technologies, influencing other technologies (Breitzman and Thomas, 2002).

$$TI(k) = \frac{\text{the average citation frequency for patents granted by firm K}}{\text{the average citation frequency for all patents}} \quad (3)$$

Table 2
Trademark portfolio vectors.

Firm	NCL_1	NCL_2	…	…	NCL_M
Firm A	TA_1	TA_2			TA_M
Firm B	TB_1	TB_2			TB_M
…	…				…

Third, *technology applicability* (*TA*) is measured by the number of IPCs. In general, the cost of patent applications increases with the number of IPCs. Thus, a firm is likely to assign IPCs that are directly related to its inventions. Despite the high cost, if a firm has patents covering diverse IPCs, its technology tends to be applied to various areas (Lerner, 1994).

$$TA(k) = \frac{\text{the average number of IPCs for patents granted by firm K}}{\text{the average number of IPCs for all patents}} \tag{4}$$

Finally, the *technology competitiveness* (*TC*) index is determined by a firm's patent share and growth rate. A firm with a high share of patents in a particular sector indicates that it concentrates on innovation activities in that business area (Chen and Chang, 2010). In a similar vein, a high increasing rate of patents by a firm signifies that the firm is likely to play a relatively significant role in the sector in the future. Therefore, the firm has a competitive technology advantage in the market if it has high values for both patent share and patent increase.

$$TC(k) = \left(\frac{\text{growth rate (firm K)- Min (growth rate)}}{\text{Max (growth rate)- Min (growth rate)}} + 1\right) \times \frac{\text{number of patents by firm K}}{\text{total number of patents}} \tag{5}$$

where

$$\text{growth rate} = \text{compound annual growth rate} = \left(\frac{\text{end value}}{\text{begin value}}\right)^{\frac{1}{\text{number of years}}}$$

Once index values are obtained, overall technology capability should be evaluated based on the values. In this study, an integrated index value is calculated using the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), which is a multi-criteria decision-making technique (Hwang and Yoon, 1981), by assigning identical weights of 25% to each index. With these technology similarity and technology capability values, all firms can be mapped onto a two-dimensional space. Here, the focus of this study is the first quadrant consisting of firms with great technology similarity and technology capability values. The firms on this quadrant will be the target firms for further analysis.

3.3.2. Trademark portfolio analysis and business opportunity analysis

A trademark portfolio analysis aims to predict a firm's behavior based on the behavioral patterns of firms having similar characteristics with the firm of interest during the collaborative filtering process. That is, the basic assumption is that a firm will be interested in the business areas in which similar firms showed great and common interest. However, in this study, we applied a modified collaborative filtering process taking into account the characteristics of new business opportunities and, accordingly, not only market penetration but also market expansion strategies are considered. Much like the development of a patent portfolio, a trademark portfolio is developed for the target firms using the NCL system as a vector dimension (see Table 2). Here, each NCL code is regarded as a business area.

The next step is to conduct a gap analysis by comparing the ratio of trademarks in each business area for a focal firm with the average of the target firms. Another index is introduced to show the differences in business areas between a focal firm and the target firms. Gap*j* (A) indicates the differences in the importance of business area j between focal firm (A) and the target firms ($t = 1, \ldots, T$), where a positive value indicates that the focal firm has a stronger emphasis on the area while a negative value indicates the opposite. When calculating the average importance value for target firms, a weighted average method was applied, where the degree of technology similarity was given a weight value on the premise that information from firms with highly similar technology portfolios should be regarded as more valuable in decision making. The following equation shows how to measure the degree of gap between Firm A and its competitors for benchmarking regarding business area j.

$$Gap_j(A) = \frac{TA_j}{\sum_{m=1}^{M} TA_m} - \frac{1}{\sum_{t=1}^{T} Similarity\ (Firm\ A, Firm\ t)} \sum_{t=1}^{T} \left(Similarity\ (Firm\ A, Firm\ t) \times \frac{Tt_j}{\sum_{m=1}^{M} Tt_m} \right) \tag{6}$$

where
TA_j: the number of trademarks firm A has in j area (NCL code)
Tt_j: the number of trademarks firm t has in j area ($t = 1, \ldots, T$)
T: the total number of firms for benchmarking
Similarity (Firm A, Firm t): cosine similarity of a trademark portfolio vector for Firm A and Firm t

Fig. 3 shows how a gap analysis can be used to identify a firm's major and minor business areas. In the figure, the solid line indicates the concentration ratio of each business area for a focal firm, while the dotted line indicates the same for the target firms. The figure shows that D is a business area that a focal firm has concentrated more highly on compared to its competitors, while A is a business area that the firm has less interest in than the others. A different strategy is required to explore business opportunities in those two different areas.

For market penetration, it is relatively easy to explore new business opportunities. The business areas in which the focal firm has ease of entry become candidates for new business opportunities. If a firm has a differentiation strategy and the areas are expected to be promising in the future, these can be business opportunities.

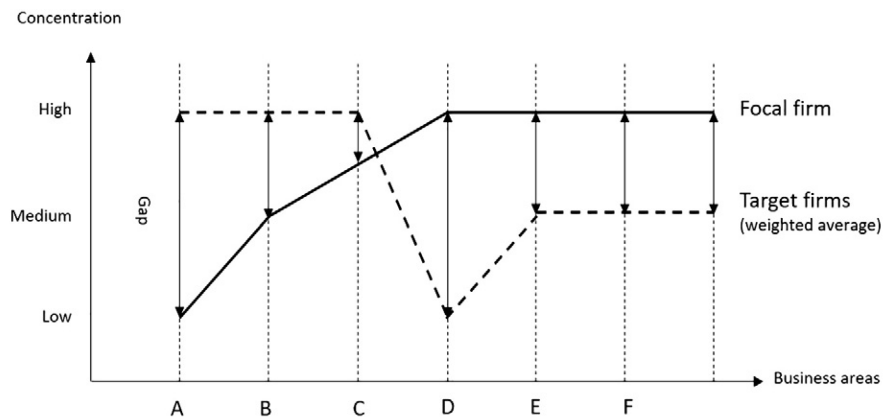

Fig. 3. An example of gap analysis results.

For market-expansion, an association mining technique is adopted. Association mining is a data-mining technique used to discover items that co-occur frequently within a dataset (Huang et al., 2011). These association rules are expressed by "(item set A: conditions) ➔ (item set B: results)", "if A, then B", or "A ➔ B." This association-mining algorithm is an unsupervised learning method for finding such association rules in the form of if-then relationships between items in a dataset. In the context of business opportunities, the association rules identify such relationships that if a company is doing business in the area of A, then it is likely to do business in the area of B. Though this association mining has focused on the transaction of products, which is referred to as market basket analysis (Blattberg et al., 2008), this study adopted the technique to analyze the association rule between business areas denoted by NCL codes (see Table 4). The three most commonly used indexes to develop the association rules include support, confidence, and lift.

First, *support* for business areas A and B is operationalized by the share of companies having both business areas together from the total number of companies used for analysis.

$$Support(A,B) = P(A \cap B) = \frac{\text{The number of companies with both business areas A and B}}{\text{The number of total companies}} \tag{7}$$

Second, *confidence* (A ➔ B) is measured by the share of companies operating in both business areas A and B from the total number of companies operating in business area A. Support and confidence values must be analyzed to find meaningful association rules.

$$Confidence(A,B) = \frac{P(A \cap B)}{P(A)} = \frac{\text{The ratio of companies with both business areas A and B}}{\text{The ratio of companies with business area A}} \tag{8}$$

Third, *lift* complements support and confidence. The lift value is measured by the share of companies having both business areas A and B on the condition that all companies having business area A are divided by the share of companies having business area B from the total number of companies. If the lift value is >1, the two business areas are positively related. If the lift is close to 1, they are likely to be independent. Similarly, if the value is <1, the two business areas tend to be negatively related to each other. As we are interested in identifying new business opportunities that can be expanded from other business areas, we will focus only on the positive relationships with a lift value >1.

$$Lift(A,B) = \frac{P(A \cap B)}{P(A) \times P(B)} = \frac{P(B|A)}{P(B)} = \frac{\text{The ratio of companies with both business areas A and B}}{\text{The ratio of companies with business area A} \times \text{The ratio of companies with business area B}} \tag{9}$$

3.3.3. In-depth analysis for opportunities

In the final step, specific business opportunities that a focal firm should concentrate on are suggested by applying text-mining techniques to the trademarks applied to the business areas of interest. Trademark documents indicate the product and service categories in which companies want to protect their brands, and therefore provide useful information about the products and services mainly offered in the business areas. For the text-mining analysis, an open source statistical program, "R," was used with "RWeka" and "tm" packages. The keyword extraction token was set to a range of two to four, aiming to obtain identifiable and meaningful keywords as product and service items.

4. Illustrative example

4.1. Data collection

A sector that can greatly benefit from the suggested approach would be one where technology and marketing both play a significant role in its innovation. Thus, companies in the sector would be characterized by active innovation activities and tend to protect their innovation via intellectual property rights, including both patents and trademarks. The cloud computing industry, which is a representative technology-based service sector, meets these conditions and was adopted as an illustrative example in this study. For this analysis, we collected patents

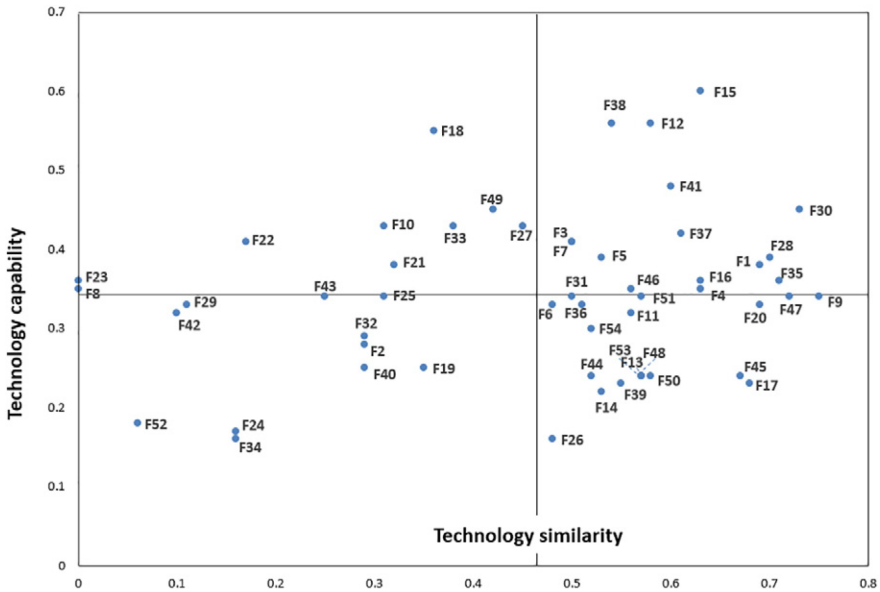

Fig. 4. Patent analysis results identifying target competitors.

published in the USPTO from 2005 to 2014 having the three keywords "cloud," "computing," and "service" in their titles, abstracts, or claims in H04 and G05 codes, which indicate IT services. Based on the number of patent applications, 55 companies were selected for further analysis. The companies had 1,827 patents and 22,215 trademarks. In this study, we substituted F1 to F55 for the real names of the companies. The list of companies is provided in Appendix 1.

4.2. Patent portfolio analysis and target firm identification results

We set Rackspace (denoted as Firm R) as a focal firm, which is positioned in the middle of the patent (13 patents) and trademark (135 trademarks) numbers. The core business areas of the company are cloud computing and cyber security. Then, centered on this company, technology similarity and technology capability values were calculated for its 54 competitors and mapped onto the two-dimensional space according to the two values, as shown in Fig. 4. In the figure, based on the mean values, the firms are classified into four groups, among which our target competitors are the first quadrant group. Detailed analysis results are provided in Appendix 2.

4.3. Trademark portfolio analysis and business opportunity analysis results

The first quadrant group included 17 companies; these firms have similar technology portfolios with Firm R and possess relatively superior technological capabilities. Restricting our focus only to the 17 companies with 3,469 trademarks, a trademark portfolio analysis was conducted. Market similarities of competitors with Firm R were calculated based on the trademark portfolio vectors as presented in Table 3. In the table, F1 and F4 possess the most common business interests with Firm R, while F35 has the least common business interests despite technology similarity with Firm R.

A gap analysis was then conducted to compare the business areas of Firm R with its 17 competitors. The weights applied to each competitor are summarized in Appendix 3 and detailed information regarding business areas is described in Appendix 4. The gap analysis results identifying Firm R's major and minor business areas are presented in Table 4.

The table shows that Firm R has focused more on business areas 35, 38, and 42, and less on business areas 9, 16, and 36 compared to its competitors having similar technological assets. Investigating the business areas in detail, we conclude that Firm R has less interest in traditional ICT-based business areas (e.g., information processing devices or hardware) as well as financial and insurance services, whereas it specialized in technological services and software development. In an attempt to understand its business areas in more detail, we applied a text-mining technique to the trademark documents, which led to identifying the specific products or service items in which Firm R is interested, as shown in Table 5. According to the table, we observe that Firm R's main products and services include "cloud computing," "computer

Table 3
Market similarity analysis results.

Ranking	Firms	Cosine similarity value	Ranking	Firms	Cosine similarity value
1	F1	0.86	10	F38	0.55
2	F4	0.73	11	F46	0.53
3	F9	0.65	12	F51	0.51
4	F30	0.64	13	F5	0.45
5	F41	0.63	14	F15	0.38
6	F12	0.60	15	F37	0.16
7	F16	0.60	16	F7	0.15
8	F28	0.59	17	F35	0.04
9	F3	0.56			

Table 4
Gap analysis results identifying major and minor business areas for Firm R.

Business areas	Gap	Business areas	Gap	Business areas	Gap
NCL_1	−0.06	NCL_16	−3.65	NCL_31	−0.02
NCL_2	−0.04	NCL_17	0.00	NCL_32	−0.02
NCL_3	−0.01	NCL_18	−0.41	NCL_33	0.00
NCL_4	−0.01	NCL_19	0.00	NCL_34	0.00
NCL_5	−0.02	NCL_20	−0.09	NCL_35	24.71
NCL_6	−0.08	NCL_21	−0.22	NCL_36	−4.02
NCL_7	−1.12	NCL_22	−0.02	NCL_37	−1.22
NCL_8	−0.12	NCL_23	−0.02	NCL_38	5.63
NCL_9	−35.79	NCL_24	−0.13	NCL_39	0.45
NCL_10	−0.28	NCL_25	−0.76	NCL_40	−0.07
NCL_11	−1.15	NCL_26	−0.06	NCL_41	2.84
NCL_12	−0.20	NCL_27	0.00	NCL_42	19.05
NCL_13	0.00	NCL_28	−1.28	NCL_43	−0.08
NCL_14	−0.55	NCL_29	−0.02	NCL_44	−0.15
NCL_15	−0.05	NCL_30	−0.03	NCL_45	−0.94

Table 5
Text-mining analysis results identifying Firm R's main products and services.

No	Keywords	Frequency
1	Cloud computing	76
2	Computer software applications	54
3	Databases digital content business	52
4	Internet web software applications	52
5	Websites Internet web software	52
6	Hosting managed server hosting	20
7	Assistance customers business problem	19
8	Managing deploying cloud computing	15
9	Platforms creating managing deploying	14
10	Operating systems computer	12

software applications," "databases digital content business," "Internet web software applications," and "websites Internet web software."

4.4. In-depth analysis of opportunities results

4.4.1. Market-penetration business opportunities

For Firm R, the three business areas of 9, 16, and 36 are worth investigating as new business opportunities from the market-penetration perspective. Here, in relation to these relatively minor business areas, the products and service items that competitors with similar technology portfolios are offering or intend to offer were analyzed using trademark data. The text-mining analysis results of the products and service items trademarked by competitors in business area 36 are demonstrated in Table 6. The table shows that the most frequently observed keywords in competitors' trademark documents include "credit card," "mobile devices," and "banking related financial services." These keywords indicate the new product and service candidates that Firm R can offer with its current technology assets. Such products or services can be new business opportunities if enough profits are projected because relevant markets are still at the early stage or show a high potential for growth, or Firm R has a competitive advantage in the market due to its technological strength. In our illustrative example, Firm R needs to explore further possibilities regarding the development of financial transaction networks or mobile payment services because the firm specializes in database construction and cloud service management.

4.4.2. Market-expansion business opportunities

The three business areas of 35, 38, and 42 were selected to explore new business opportunities from the market-expansion perspective. These are the business areas on which Firm R has concentrated, and business opportunities that are closely related to these three areas were identified using association rule analysis. The association rule analysis results indicate that business areas 16, 18, 25, and 37 have lift values >1 and thus positively relate to the three main business areas of Firm R (see Table 7).

Table 7
Association rules.

No	IF	Then	Support	Confidence	Lift
1	{35,38,42}	16	0.56	0.83	1.25
2		18	0.50	0.75	1.35
3		25	0.50	0.75	1.35
4		37	0.50	0.75	1.50

These are the business areas in which firms in the 35, 38, or 42 business areas are also likely to be. The areas of 16, 18, 25, and 37 might have relatively low technological similarity but high customer or market similarity with business areas 35, 38, and 42, and thus can be pursued together. Among the four business areas, 37 has the greatest lift value and therefore text-mining analysis was conducted to explore more detailed product and service items. Table 8 shows the keywords identified from trademarks concerning business area 37. The table shows that the most frequently observed keywords in the trademark documents regarding the 37 NCL code include "service field," "installation repair," "technical support services," "maintenance repair," and "data processing." These keywords indicate that new product and service items opportunities may lie in the maintenance of cloud services or repair of equipment for cloud services.

4.4.3. Discussions

The suggested approach allows a systematic exploration of new business opportunities based on the analysis of patent and trademark databases, which are among the largest and most reliable databases offering competitive intelligence. Compared to patent data, trademark data has attracted little attention from both academics and practitioners; this study has shown the potential of trademark data as a source of competitive intelligence. In particular, patent and trademark data show not only the past but also strategies for the future because companies are likely to acquire intellectual property rights before they start relevant businesses. Once competitors' strategies are clearly understood, it is easy for a firm to formulate its own strategy, either by benchmarking the competitors' or by building on a fundamentally different position strategy.

The approach has distinguishing advantages over existing approaches in identifying new business opportunities. Business opportunities may come not only from technology but also from markets. Considering both technology and market factors will allow for a more holistic approach in exploring business opportunities; integrating the two databases – patents and trademarks – enables a comprehensive understanding of a competitor strategy, which deliberates both business and technology strategies. The analysis results also indicate that companies with similar patent portfolios may have a diverse set of trademarks; technology and business strategies are not always aligned. Thus, considering both strategies is essential to gain competitor intelligence.

Moreover, the approach in this study suggests specific product and service items rather than broad business areas or general technology areas. Since the opportunities can be expressed in concrete terms such as "installation repair," "technical support services," "updating maintenance," and "servers storage" (as shown in Table 8), practical guidelines can be provided for the users. Furthermore, recent advances in big data

Table 6
Keywords identified from trademarks in the 36 code.

No	Keywords	Frequency
1	Credit card	57
2	Mobile devices	53
3	Banking related financial services	47
4	Consulting services	25
5	Bill payment	16
6	Algorithmic trading	15
7	Insurance agency brokerage services	14
8	Networks global communication networks	14
9	Brokerage services security	13
10	Electronic communications	13

Table 8
Keywords identified from trademarks in the 37 code.

No	Keywords	Frequency
1	Services field	59
2	Installation repair	56
3	Technical support services	47
4	Maintenance repair	46
5	Data processing	43
6	Services namely troubleshooting	33
7	Updating maintenance	28
8	Services information	25
9	Software managing	25
10	Servers storage	21

analytics are expected to expand our capability to draw competitor intelligence from patent and trademark databases towards more elaborated opportunity identification. Competitor intelligence activities will adopt more systematic approaches, whereas system development will help reduce the human efforts required to derive such intelligence.

However, the suggested approach may be limited in its use for those industries or firms where patenting and trademark application activities are not active. Innovation outputs can be protected in several ways including patents, trademarks, trade secrets, utility models, design rights, and copyrights. Different industry characteristics require different protection mechanisms, which affects the usability of patents and trademarks as a source of competitor intelligence. For example, SMEs competing over process technologies are more likely to use trade secrets to protect their innovations. Their technologies may not be patented. This phenomenon is strengthened with the size of companies; smaller companies are less likely to protect their innovation via legal protection mechanisms. Therefore, the most effective application of the suggested approach is expected in technology-based products or services and firms with technological and market strength that are not too small to use such legal protection mechanisms.

Finally, it is worth discussing that new business opportunities can be defined various ways and thus other available approaches need to be developed with continuous efforts. For example, the context of identifying new business opportunities was set to “multiple technologies” and “within a business area.” That is, the competitors investigated were limited to those in the same business areas, where a technology portfolio was considered to determine the final set of companies to be examined. However, new business opportunities can come from converging technologies; companies with a similar technology portfolio operating in other business sectors are another valuable source of technology intelligence, although they are not direct competitors in the current business areas. Similarly, instead of using a technology portfolio, using a single technology (patent) for identifying new business opportunities is also feasible and promising, as addressed by Yoon et al. (in press). A further study is needed to define business opportunities from various perspectives, adopting a taxonomical or typological approach, which enables better use of patent and trademark databases for competitor intelligence.

5. Conclusions

This study aims to develop a novel approach for systematically exploring new business opportunities from competitor intelligence. To achieve this aim, we suggest the integrated use of patent and trademark data along with a modified collaborative filtering method. In addition to this method, we used several other analyses, including patent index, association rule, and text mining. We first defined new business opportunities as *the business areas in which companies having similar technological assets offer products or services but a focal firm does not* or “*the business areas closely related to those in which the focal firm has strength*.” Patent documents, as representative technological information, were used to identify companies having similar technological assets to the focal firm as well as strong technological capabilities to benchmark. Trademark documents, as representative market information, were used to compare the business areas between the focal firm and competitors and further investigate the products and services competitors offer in the business areas in detail. Therefore, academically, this is one of the earliest attempts to integrate patent and trademark databases to extract meaningful implications; relatively less effort has been made to analyze the trademark data, which provide fruitful competitor intelligence if well analyzed. In practice, the suggested approach can be used to develop a decision-support tool based on objective and quantitative data, complementary to the existing qualitative data such as customers' opinions or experts' insights.

Despite these meaningful contributions, this study is subject to several limitations. First, there exist various types of new business opportunities, in addition to the market-penetration and market-expansion types adopted in this study, which we failed to consider. Second, this study conducted only static analyses. However, competitors' changing trends in business areas can also be valuable sources for identifying new business opportunities, necessitating the use of dynamic analyses. Particularly for dynamic analyses, a specific algorithm for trademark data needs to be developed. Trademarks are updated every 10 years perpetually, unlike patents that can claim exclusive rights only for 20 years. If an update was not made for a trademark in a particular business area, it can indicate changes in business strategies or marketing strategies. Further analyses will be needed to develop an algorithm to deduce intelligence from a trademark database considering the peculiar characteristics of trademarks as a mechanism to protect innovation in a market.

Acknowledgement

This work was supported by the Ajou University research fund (grant number: S-2015-G0001-00333).

Appendix 1. List of companies for analysis (alphabetical order)

F1. ACCENTURE GLOBAL SERVICES GMBH
F2. ADOBE SYSTEMS INC
F3. ALCATEL LUCENT INC
F4. AMAZON TECHNOLOGIES INC
F5. APPLE INC
F6. AT&T INTELLECTUAL PTY LLP
F7. BANK OF AMERICA CORP
F8. BOEING CO
F9. BOX INC
F10. CENTURYLINK INTELLECTUAL PROPERTY LLC
F11. CISCO TECHNOLOGY INC
F12. CITRIX SYSTEMS INC
F13. CLOUDFLARE INC
F14. CLOUDNEXA
F15. COMMVAULT SYSTEMS INC
F16. EGNYTE INC
F17. EMC CORP
F18. ENDURANCE INTERNATIONAL GROUP INC
F19. ETRI
F20. GOOGLE INC
F21. HEWLETT PACKARD CO
F22. HONEYWELL INTERNATIONAL INC

F23. HONG FU JIN PRECISION IND CO LTD|HON HAI PRECISION IND CO LTD
F24. HOPTO INC
F25. INCONTACT INC
F26. INFOSYS TECHNOLOGIES LTD
F27. INTEL CORP
F28. INTERNATIONAL BUSINESS MACHINES CORP
F29. LOCKHEED MARTIN CORP
F30. MICROSOFT CORP
F31. NASUNI CORP
F32. NEC CORP
F33. NET POWER AND LIGHT INC
F34. NETAPP INC
F35. NEXTBIT SYSTEMS INC
F36. NOKIA CORP
F37. NOVELL INC
F38. ORACLE AMERICA INC
F39. PALO ALTO NETWORKS INC
F40. PARALLELS
F41. RED HAT INC
F42. RICOH CO LTD
F43. RINGCENTRAL INC
F44. RIVERBED TECHNOLOGY INC
F45. SALESFORCE COM INC
F46. SAMSUNG ELECTRONICS CO LTD
F47. SAP SE
F48. SYMANTEC CORP
F49. TELEFONAKTIEBOLAGET LM ERICSSON
F50. VERIZON PATENT AND LICENSING INC
F51. VMWARE INC
F52. XEROX CORP
F53. YAHOO INC
F54. ZSCALER INC
RACKSPACE (a focal firm)

Appendix 2. Technology similarity and technology capability analyses

Table A1
Technology similarity analysis results.

Firm	Cosine similarity value		Firm	Cosine similarity value		Firm	Cosine similarity value	
	Rank	Value		Rank	Value		Rank	Value
F1	7	0.69	F19	39	0.35	F37	13	0.61
F2	43	0.29	F20	6	0.69	F38	24	0.54
F3	31	0.50	F21	40	0.32	F39	23	0.55
F4	10	0.63	F22	47	0.17	F40	45	0.29
F5	25	0.53	F23	54	0.00	F41	14	0.60
F6	34	0.48	F24	49	0.16	F42	51	0.10
F7	32	0.50	F25	42	0.31	F43	46	0.25
F8	53	0.00	F26	33	0.48	F44	28	0.52
F9	1	0.75	F27	35	0.45	F45	9	0.67
F10	41	0.31	F28	5	0.70	F46	22	0.56
F11	21	0.56	F29	50	0.11	F47	3	0.72
F12	16	0.58	F30	2	0.73	F48	17	0.57
F13	19	0.57	F31	30	0.50	F49	36	0.42
F14	26	0.53	F32	44	0.29	F50	15	0.58
F15	11	0.63	F33	37	0.38	F51	18	0.57
F16	12	0.63	F34	48	0.16	F52	52	0.06
F17	8	0.68	F35	4	0.71	F53	20	0.57
F18	38	0.36	F36	29	0.51	F54	27	0.52

Table A2
Technology capability analysis results (T-values and TOPSIS values).

Firm	Technology profitability	Technology impact	Technology applicability	Technology competitiveness	TOPSIS values
F1	51.94	55.28	47.63	57.13	0.38
F2	40.30	48.32	42.74	57.13	0.28
F3	52.17	47.08	54.43	62.43	0.41
F4	44.13	47.53	48.02	62.82	0.35
F5	73.28	43.30	48.17	45.84	0.39
F6	40.30	49.19	44.71	62.36	0.33
F7	45.96	50.69	75.20	34.26	0.41
F8	62.46	42.49	32.48	61.98	0.35
F9	46.46	47.39	62.01	40.72	0.34
F10	64.04	43.06	71.24	34.52	0.43
F11	44.45	53.92	48.83	51.50	0.32
F12	56.13	80.70	53.62	51.53	0.56
F13	40.30	43.79	52.21	40.06	0.24
F14	40.30	54.60	43.56	39.69	0.22
F15	53.87	92.31	59.38	45.42	0.60
F16	40.30	43.06	71.24	34.26	0.36
F17	41.81	41.83	43.23	51.66	0.23
F18	79.87	56.25	73.22	39.69	0.55
F19	49.35	41.83	47.51	45.92	0.25
F20	50.60	46.11	56.30	45.92	0.33
F21	44.10	42.60	45.22	73.50	0.38
F22	65.36	54.84	45.86	50.78	0.41
F23	50.38	41.88	43.56	67.44	0.36
F24	40.30	40.90	45.86	34.32	0.17
F25	48.22	48.83	57.40	45.12	0.34
F26	38.72	43.64	39.40	45.75	0.16
F27	66.92	45.56	57.40	51.53	0.43
F28	45.05	46.00	48.44	69.52	0.39
F29	40.30	46.53	46.32	61.98	0.33
F30	50.74	50.71	47.89	74.18	0.45
F31	56.13	63.26	29.71	50.55	0.34
F32	43.13	46.36	49.49	51.66	0.29
F33	65.63	55.76	49.09	50.87	0.43
F34	44.26	46.67	34.90	45.36	0.16
F35	56.13	42.50	63.55	34.26	0.36
F36	49.21	44.51	53.94	51.37	0.33
F37	44.89	65.54	47.93	57.19	0.42
F38	67.44	51.92	73.22	57.24	0.56
F39	49.80	44.80	46.32	39.76	0.23
F40	41.44	47.80	49.49	45.59	0.25
F41	40.95	76.71	50.90	56.65	0.48
F42	49.35	46.36	55.42	45.92	0.32
F43	70.37	42.49	46.32	34.44	0.34
F44	56.13	49.04	33.67	45.42	0.24
F45	40.30	48.33	47.71	45.57	0.24
F46	50.33	52.01	51.86	51.34	0.35
F47	43.18	58.10	42.72	56.92	0.34
F48	43.75	42.75	44.76	51.22	0.24
F49	62.29	50.28	60.48	51.43	0.45
F50	39.47	43.90	50.11	46.08	0.24
F51	47.16	61.14	43.56	51.34	0.34
F52	44.26	42.46	44.13	40.11	0.18
F53	40.30	42.70	38.37	56.77	0.24
F54	41.02	51.85	43.56	56.92	0.30

Appendix 3. Weights for gap analysis

Table A3
Adjusted weights based on technology similarity values.

Ranking	Firms	Cosine similarity value	Ranking	Firms	Cosine similarity value
1	F9	0.07	11	F12	0.06
2	F30	0.07	12	F51	0.05
3	F35	0.07	13	F46	0.05
4	F28	0.07	14	F38	0.05
5	F1	0.07	15	F5	0.05
6	F4	0.06	16	F3	0.05
7	F15	0.06	17	F7	0.05
8	F16	0.06			
9	F37	0.06			
10	F41	0.06			

Appendix 4. Business areas based on NICE trademark classification

Table A4
NICE codes and descriptions (source: http://www.wipo.int/classifications/nivilo/nice/index.htm).

NICE codes	Descriptions
1	Chemicals used in industry, science and photography, as well as in agriculture, horticulture and forestry; unprocessed artificial resins, unprocessed plastics; manures; fire extinguishing compositions; tempering and soldering preparations; chemical substances for preserving foodstuffs; tanning substances; adhesives used in industry.
2	Paints, varnishes, lacquers; preservatives against rust and against deterioration of wood; colorants; mordents; raw natural resins; metals in foil and powder form for painters, decorators, printers and artists.
3	Bleaching preparations and other substances for laundry use; cleaning, polishing, scouring and abrasive preparations; soaps; perfumery, essential oils, cosmetics, hair lotions; dentifrices.
4	Industrial oils and greases; lubricants; dust absorbing, wetting and binding compositions; fuels (including motor spirit) and illuminants; candles and wicks for lighting.
5	Pharmaceutical and veterinary preparations; sanitary preparations for medical purposes; dietetic food and substances adapted for medical or veterinary use, food for babies; dietary supplements for humans and animals; plasters, materials for dressings; material for stopping teeth, dental wax; disinfectants; preparations for destroying vermin; fungicides, herbicides.
6	Common metals and their alloys; metal building materials; transportable buildings of metal; materials of metal for railway tracks; non-electric cables and wires of common metal; ironmongery, small items of metal hardware; pipes and tubes of metal; safes; goods of common metal not included in other classes; ores.
7	Machines and machine tools; motors and engines (except for land vehicles); machine coupling and transmission components (except for land vehicles); agricultural implements other than hand-operated; incubators for eggs; automatic vending machines.
8	Hand tools and implements (hand-operated); cutlery; side arms; razors.
9	Scientific, nautical, surveying, photographic, cinematographic, optical, weighing, measuring, signaling, checking (supervision), life-saving and teaching apparatus and instruments; apparatus and instruments for conducting, switching, transforming, accumulating, regulating or controlling electricity; apparatus for recording, transmission or reproduction of sound or images; magnetic data carriers, recording discs; compact discs, DVDs and other digital recording media; mechanisms for coin-operated apparatus; cash registers, calculating machines, data processing equipment, computers; computer software; fire-extinguishing apparatus.
10	Surgical, medical, dental and veterinary apparatus and instruments, artificial limbs, eyes and teeth; orthopedic articles; suture materials.
11	Apparatus for lighting, heating, steam generating, cooking, refrigerating, drying, ventilating, water supply and sanitary purposes.
12	Vehicles; apparatus for locomotion by land, air or water.
13	Firearms; ammunition and projectiles; explosives; fireworks.
14	Precious metals and their alloys and goods in precious metals or coated therewith, not included in other classes; jewelry, precious stones; horological and chronometric instruments.
15	Musical instruments.
16	Paper, cardboard and goods made from these materials, not included in other classes; printed matter; bookbinding material; photographs; stationery; adhesives for stationery or household purposes; artists' materials; paint brushes; typewriters and office requisites (except furniture); instructional and teaching material (except apparatus); plastic materials for packaging (not included in other classes); printers' type; printing blocks.
17	Rubber, gutta-percha, gum, asbestos, mica and goods made from these materials and not included in other classes; plastics in extruded form for use in manufacture; packing, stopping and insulating materials; flexible pipes, not of metal.
18	Leather and imitations of leather, and goods made of these materials and not included in other classes; animal skins, hides; trunks and travelling bags; umbrellas and parasols; walking sticks; whips, harness and saddlery.
19	Building materials (non-metallic); non-metallic rigid pipes for building; asphalt, pitch and bitumen; non-metallic transportable buildings; monuments, not of metal.
20	Furniture, mirrors, picture frames; goods (not included in other classes) of wood, cork, reed, cane, wicker, horn, bone, ivory, whalebone, shell, amber, mother-of-pearl, meerschaum and substitutes for all these materials, or of plastics.
21	Household or kitchen utensils and containers; combs and sponges; brushes (except paint brushes); brush-making materials; articles for cleaning purposes; steel wool; unworked or semi-worked glass (except glass used in building); glassware, porcelain and earthenware not included in other classes.
22	Ropes, string, nets, tents, awnings, tarpaulins, sails, sacks and bags (not included in other classes); padding and stuffing materials (except rubber or plastics); raw fibrous textile materials.
23	Yarns and threads, for textile use.
24	Textiles and textile goods, not included in other classes; bed covers; table covers.
25	Clothing, footwear, headgear.
26	Lace and embroidery, ribbons and braid; buttons, hooks and eyes, pins and needles; artificial flowers.
27	Carpets, rugs, mats and matting, linoleum and other materials for covering existing floors; wall hangings (non-textile).
28	Games and playthings; gymnastic and sporting articles not included in other classes; decorations for Christmas trees.
29	Meat, fish, poultry and game; meat extracts; preserved, frozen, dried and cooked fruits and vegetables; jellies, jams, compotes; eggs; milk and milk products; edible oils and fats.
30	Coffee, tea, cocoa and artificial coffee; rice; tapioca and sago; flour and preparations made from cereals; bread, pastry and confectionery; ices; sugar, honey, treacle; yeast, baking-powder; salt; mustard; vinegar, sauces (condiments); spices; ice.
31	Grains and agricultural, horticultural and forestry products not included in other classes; live animals; fresh fruits and vegetables; seeds; natural plants and flowers; foodstuffs for animals; malt.
32	Beers; mineral and aerated waters and other non-alcoholic beverages; fruit beverages and fruit juices; syrups and other preparations for making beverages.
33	Alcoholic beverages (except beers).
34	Tobacco; smokers' articles; matches.
35	Advertising; business management; business administration; office functions.
36	Insurance; financial affairs; monetary affairs; real estate affairs.
37	Building construction; repair; installation services.
38	Telecommunications.
39	Transport; packaging and storage of goods; travel arrangement.
40	Treatment of materials.
41	Education; providing of training; entertainment; sporting and cultural activities.
42	Scientific and technological services and research and design relating thereto; industrial analysis and research services; design and development of computer hardware and software.
43	Services for providing food and drink; temporary accommodation.
44	Medical services; veterinary services; hygienic and beauty care for human beings or animals; agriculture, horticulture and forestry services.
45	Legal services; security services for the protection of property and individuals; personal and social services rendered by others to meet the needs of individuals.

References

Albert, M.B., Avery, D., Narin, F., McAllister, P., 1991. Direct validation of citation counts as indicators of industrially important patents. Res. Policy 20 (3), 251–259.

Archibugi, D., Iammarino, S., 2002. The globalization of technological innovation: definition and evidence. Rev. Int. Polit. Econ. 9 (1), 98–122.

Ardichvili, A., Cardozo, R., Ray, S., 2003. A theory of entrepreneurial opportunity identification and development. J. Bus. Ventur. 18 (1), 105–123.

Armstrong, J.S., Collopy, F., 1996. Competitor orientation: effects of objectives and information on managerial decisions and profitability. J. Mark. Res. 188–199.

Ashton, W.B., Sen, R.K., 1988. Using patent information in technology business planning I. Res. Technol. Manag. 31 (6), 42–46.

Austin, D.H., 1993. An event-study approach to measuring innovative output: the case of biotechnology. Am. Econ. Rev. 83 (2), 253–258.

Bhatnagar, V., 2016. Collaborative Filtering Using Data Mining and Analysis. IGI Global, PA.

Blattberg, R.C., Kim, B.D., Neslin, S.A., 2008. Market basket analysis. Database Marketing: Analyzing and Managing Customers, pp. 339–351.

Breitzman, A., Thomas, P., 2002. Using patent citation analysis to target/value M&A candidates. Res. Technol. Manag. 45 (5), 28.

Bröring, S., Martin Cloutier, L., Leker, J., 2006. The front end of innovation in an era of industry convergence: evidence from nutraceuticals and functional foods. R&D Manag. 36 (5), 487–498.

Casson, M., 1982. The Entrepreneur: An Economic Theory. Barnes and Noble Books, NJ.

Chen, Y.S., Chang, K.C., 2010. Exploring the nonlinear effects of patent citations, patent share and relative patent position on market value in the US pharmaceutical industry. Tech. Anal. Strat. Manag. 22 (2), 153–169.

Choi, S., Jun, S., 2014. Vacant technology forecasting using new Bayesian patent clustering. Tech. Anal. Strat. Manag. 26 (3), 241–251.

Daim, T.U., Rueda, G., Martin, H., Gerdsri, P., 2006. Forecasting emerging technologies: use of bibliometrics and patent analysis. Technol. Forecast. Soc. Chang. 73 (8), 981–1012.

Doern, G.B., 1999. Global Change and Intellectual Property Agencies: An Institutional Perspective. Psychology Press.

Ernst, H., 2003. Patent information for strategic technology management. World Patent Inf. 25 (3), 233–242.

Fabry, B., Ernst, H., Langholz, J., Köster, M., 2006. Patent portfolio analysis as a useful tool for identifying R&D and business opportunities—an empirical application in the nutrition and health industry. World Patent Inf. 28 (3), 215–225.

Fahey, L., 2001. Competitor scenarios: projecting a rival's marketplace strategy. Compet. Intell. Rev. 10 (2), 65–85.

Gans, J.S., Stern, S., 2003. The product market and the market for "ideas": commercialization strategies for technology entrepreneurs. Res. Policy 32 (2), 333–350.

Geum, Y., Lee, S., Yoon, B., Park, Y., 2010. Identifying and evaluating strategic partners for collaborative R&D: index-based approach using patents and publications. Technovation 33 (6–7), 211–224.

Gielnik, M., Frese, M., Grf, J., Kampschulte, A., 2012. Creativity in the opportunity identification process and the moderating effect of diversity of information. J. Bus. Ventur. 27, 559–576.

Goldberg, K., Roeder, T., Gupta, D., Perkins, C., 2001. Eigentaste: a constant time collaborative filtering algorithm. Inf. Retr. 4 (2), 133–151.

Greenhalgh, C., Rogers, M., 2010. Innovation, Intellectual Property, and Economic Growth. Princeton University Press.

Grupp, H., Schmoch, U., 1999. Patent statistics in the age of globalisation: new legal procedures, new analytical methods, new economic interpretation. Res. Policy 28 (4), 377–396.

Gupta, V., Goktan, A., Gunay, G., 2014. Gender differences in evaluation of new business opportunity: a stereotype threat perspective. J. Bus. Ventur. 29, 273–288.

Hanssens, D.M., 1980. Market response, competitive behavior, and time series analysis. J. Mark. Res. 470–485.

Harhoff, D., Scherer, F.M., Vopel, K., 2003. Citations, family size, opposition and the value of patent rights. Res. Policy 32 (8), 1343–1363.

He, W., Zha, S., Li, L., 2013. Social media competitive analysis and text mining: a case study in the pizza industry. Int. J. Inf. Manag. 33 (3), 464–472.

Hoopes, D.G., Madsen, T.L., Walker, G., 2003. Guest editors' introduction to the special issue: why is there a resource-based view? Toward a theory of competitive heterogeneity. Strateg. Manag. J. 24 (10), 889–902.

Huang, Z., Lu, X., Duan, H., 2011. Mining association rules to support resource allocation in business process management. Expert Syst. Appl. 38 (8), 9483–9490.

Hwang, C.L., Yoon, K., 1981. Multiple Attribute Decision Making: Methods and Applications a State-of-the-art Survey. vol. 186. Springer Verlag, NY.

Jaffe, A.B., 1989. Characterizing the "technological position" of firms, with application to quantifying technological opportunity and research spillovers. Res. Policy 18 (2), 87–97.

Jin, Jian, Ji, Ping, Rui, Gu, 2016. Identifying comparative customer requirements from product online reviews for competitor analysis. Eng. Appl. Artif. Intell. 49, 61–73 (March).

Lanjouw, J.O., Pakes, A., Putnam, J., 1998. How to count patents and value intellectual property: the uses of patent renewal and application data. J. Ind. Econ. 46 (4), 405–432.

Lee, S., Yoon, B., Lee, C., Park, J., 2009a. Business planning based on technological capabilities: patent analysis for technology-driven roadmapping. Technol. Forecast. Soc. Chang. 76 (6), 769–786.

Lee, S., Yoon, B., Park, Y., 2009b. An approach to discovering new technology opportunities: keyword-based patent map approach. Technovation 29 (6–7), 481–497.

Lee, S., Mortara, L., Kerr, C., Phaal, R., Probert, D., 2012. Analysis of document-mining techniques and tools for technology intelligence: discovering knowledge from technical documents. Int. J. Technol. Manag. 60 (1/2), 130–156.

Lee, C., Kang, B., Shin, J., 2015. Novelty-focused patent mapping for technology opportunity analysis. Technol. Forecast. Soc. Chang. 90 (B), 355–365.

Lerner, J., 1994. The importance of patent scope: an empirical analysis. RAND J. Econ. 319–333.

Madani, F., Weber, C., 2016. The evolution of patent mining: applying bibliometrics analysis and keyword network analysis. World Patent Inf. 46, 32–48.

Narver, J.C., Slater, S.F., 1990. The effect of a market orientation on business profitability. J. Mark. 20–35.

Niemann, H., Möehrle, M., Frischkorn, J., 2017. Use of a new patent text-mining and visualization method for identifying patenting patterns over time: concept, method and test application. Technol. Forecast. Soc. Chang. 115, 210–220.

Nordman, E., Tolstoy, 2016. The impact of opportunity connectedness on innovation in SMEs' foreign-market relationships. Technovation 57 (58), 47–57.

Overholm, H., 2015. Collectively created opportunities in emerging ecosystems: the case of solar service ventures. Technovation 39 (40), 14–25.

Palo, T., Tähtinen, J., 2013. Networked business model development for emerging technology-based services. Ind. Mark. Manag. 42, 773–782.

Peng, Y., Liang, I., 2016. A dynamic framework for competitor identification: a neglecting role of dominant design. J. Bus. Res. 69 (5), 1898–1903.

Peyrot, M., Childs, N., van Doren, D., Allen, K., 2002. An empirically based model of competitor intelligence use. J. Bus. Res. 55 (9), 747–758.

Porter, M.E., 1980. Competitive Strategy. Free Press, NY.

Reitzig, M., 2004. Strategic management of intellectual property. MIT Sloan Manag. Rev. 45 (3), 35–40.

Sandner, P.G., Block, J., 2011. The market value of R&D, patents, and trademarks. Res. Policy 40 (7), 969–985.

Schumpeter, J., 1994. Capitalism, Socialism, and Democracy. Routledge, NY.

Seol, H., Lee, S., Kim, C., 2011. Identifying new business areas using patent information: a DEA and text mining approach. Expert Syst. Appl. 38 (4), 2933–2941.

Shane, S., Nicolaou, N., 2015. Creative personality, opportunity recognition and the tendency to start businesses: a study of their genetic predispositions. J. Bus. Ventur. 30, 407–419.

Söllner, A., Rese, M., 2001. Market segmentation and the structure of competition: applicability of the strategic group concept for an improved market segmentation on industrial markets. J. Bus. Res. 51 (1), 25–36.

Stevenson, H., Roberts, M., Grousbeck, H., 1989. New Business Ventures and the Entrepreneur. Irwin, Boston.

Su, X., Khoshgortaar, T.M., 2009. A survey of collaborative filtering techniques. Adv. Artif. Intell. 2009, 4.

Tong, X., Frame, J.D., 1994. Measuring national technological performance with patent claims data. Res. Policy 23 (2), 133–141.

Ucbasaran, D., Westhead, P., Wright, M., 2009. The extent and nature of opportunity identification by experienced entrepreneurs. J. Bus. Ventur. 24, 99–115.

Vandor, P., Franke, N., 2016. See Paris and… found a business? The impact of cross-cultural experience on opportunity recognition capabilities. J. Bus. Ventur. 31, 388–407.

Venkataraman, S., 1997. The Distinctive Domain of Entrepreneurship Research. Advances in Entrepreneurship, Firm Emergence and Growth 3(1) pp. 119–138.

Venugopalan, S., Rai, V., 2015. Topic based classification and pattern identification in patents. Technol. Forecast. Soc. Chang. 94, 236–250.

Verspagen, B., Criscuolo, P., Narula, R., 2005. The role of home and host country innovation systems in R&D internationalisation: a patent citation analysis. Econ. Innov. New Technol. 14 (5), 417–433.

Wu, J., Olk, P., 2014. Technological advantage, alliances with customers, local knowledge and competitor identification. J. Bus. Res. 67 (10), 2106–2114.

Yoon, J., Seo, W., Coh, B., Song, I., Lee, J., 2017. Identifying product opportunities using collaborative filtering-based patent analysis. Comput. Ind. Eng. http://dx.doi.org/10.1016/j.cie.2016.04.009 (in press).

Zhou, H., Sandner, P.G., Martinelli, S.L., Block, J.H., 2016. Patents, trademarks, and their complementarity in venture capital funding. Technovation 47, 14–22.

Mingook Lee received his master's degree in industrial engineering from Ajou University, Suwon, Korea in 2016. He is currently a research associate at Korea Institute of Science and Technology Information, Seoul, Korea. His research interests include technology capability analysis, R&D collaboration, and technology intelligence.

Sungjoo Lee received Ph.D. on industrial engineering from Seoul National University, Seoul, Korea in 2007. She is currently an associate professor of Industrial Engineering at Ajou University, Suwon, Korea. Before joining Ajou University, she spent a year as a visiting scholar at the Centre for Technology Management, Institute for Manufacturing, University of Cambridge. Her research interests include technology roadmapping, patent engineering, and technology opportunity discovery.