ARTICLE IN PRESS

Technological Forecasting & Social Change xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

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

Identifying emerging Research and Business Development (R&BD) areas based on topic modeling and visualization with intellectual property right data

Yujin Jeong, Inchae Park, Byungun Yoon*

Dongguk University, Seoul, South Korea

A R T I C L E I N F O

Keywords:
Intellectual property management
Research and Business Development (R&BD)
Network analysis
GTM
Topic modeling

A B S T R A C T

Although investments of R&D by government and firms have enlarged and the amount of patents has increased rapidly, R&D almost fails to commercialize for various reasons. For the purpose of decreasing failure rate of technology commercialization, it is important to identify emerging business based on technology in advance and establish appropriate strategy, leading to surviving at the market. Therefore, this paper aims to explore emerging Research and Business Development (R&BD) areas, and establish a business strategy based on valuable patents by comprehensively analyzing IPRs - patent as well as design and trademark. First, unrevealed but potential R&BD areas are explored by analyzing the relation between patent and trademark through topic modeling and network analysis, which aims to preferentially find potential business opportunities that can be implemented by new technology. Potential R&BD areas are recognized as the hidden link in the network of patents and trademarks. Second, emerging R&BD areas are selected by considering the status of the competition and markets through trademark analysis based on generative topographic mapping (GTM) after finding potential R&BD areas with network analysis from the viewpoint of the applicant for a trademark. Finally, new opportunities and strategies for successful R&BD are suggested by analyzing design patents that are representative of the appearance of a product in detail. The result of this study provides more concrete R&BD strategies within the framework of product and business development, based on relations between IPRs, which can be regarded as an initial study that comprehensively utilizes diverse kinds of IPRs.

1. Introduction

The economic paradigm is shifting from an asset-based economy to a knowledge-based economy, which makes intellectual property rights (IPRs) management more important. The IPRs involving patent, design, and trademark serve as the driving force of national and economic development (Alikhan, 2002; Gould and Gruben, 1996). If firms or governments neglect IPRs, this brings about the unnecessary and enormous expense, as well as hindering national development (Kanwar and Evenson, 2003). Among others, the patent has been in the limelight within the framework of technology management, because it contains significant information for decades (Ernst, 2003; Wu et al., 2015). Although the basic function of the patent is to protect the owners' rights and prevent infringements (Nam and Barnett, 2011), it can be also utilized to monitor competitors, assess technology, examine mergers and acquisitions (M&A) options, and manage human resources, based upon the basic function (Ernst, 2003). The trademark has also received a lot of attention, in that it is able to protect marketing assets within firms against competitors, as well as to link a firm with its customers (Sandner and Block, 2011). In this way, patent and trademark have played important roles in strategic planning and making decisions, and so it is necessary to analyze and manage IPRs comprehensively.

As the importance of IP management increases, investments in R&D have rapidly expanded worldwide. While this has caused many patent applications and paper publications, some arguments have been made about the efficiency of R&D investment. Most outputs produced by R&D activity, such as patents and research papers, have not connected with the product or service correctly. Namely, technology with novelty and creativity has often failed to be commercialized, for different reasons, like insufficient support policy, and asymmetric information between technology and business. Thus, Research and Business, Development (R&BD) should be emphasized that integrates both innovation and market, to pursue both market-oriented technology and technology commercialization. To do this, the importance of discovering new business opportunity by considering technological aspect in advance is continuously growing for successful commercialization.

* Corresponding author at: Department of Industrial & Systems Engineering, School of Engineering, 3-26, Pil-dong 3-ga, Chung-gu, Seoul 100-715, South Korea.
E-mail address: postman3@dongguk.edu (B. Yoon).

https://doi.org/10.1016/j.techfore.2018.05.010
Received 1 January 2018; Received in revised form 18 April 2018; Accepted 9 May 2018

Great attention has been shown to the question of success and failure of technology commercialization. The factors for success and failure of commercialization have been derived from a variety of perspectives, such as resource and capability (Kassicieh et al., 2002; Lin et al., 2015; Sohn and Moon, 2003; Zahra and Nielsen, 2002), firm strategy (Duke, 1995; Slater and Mohr, 2006), firm structure (Duke, 1995), or commercialization environment (Gans and Stern, 2003), using regression analysis, factor analysis, or structural equation modeling. Since there has been a huge amount of R&D investment in university and research institutions, few studies have focused on technology transfer by university-industry or university-government relations (Hsu et al., 2015). Also, the technology commercialization served as a mediator between organizational resources, innovative capabilities, and venture performance (Chen and Chen, 2007). After identifying determinants for the success of technology commercialization, there were attempts to evaluate the possibility of success or assess the priorities of factors by using Delphi, Analytic hierarchy process, or data envelopment analysis (Cho and Lee, 2013; Sohn and Moon, 2004). In addition, much has been said about exploration for promising research and development areas through patent analysis, since IPRs have been regarded as one of the outcomes of R&D activity, and include numerous items of information relevant to new technology. However, relatively little research has investigated technology planning and forecasting through design patents and trademarks or all kinds of IPRs in an integral manner. This is because they have a relatively smaller amount of information that is applicable to analysis for planning or forecasting with quantitative methodology than patent information. Only a few studies were conducted with the aim of developing design patent maps by technical and design experts (C. Chen, 2009; R. Chen, 2009; Chen and Chen, 2007).

Therefore, the aim of this study is to provide an approach to identify emerging R&BD areas through comprehensively analyzing the IPRs – patent, trademark, and design patent – to gain as much bibliometric information as possible from IPRs for successful technology commercialization. In this paper, emerging R&BD area is defined as an area that has the possibility to enter a new market with innovative technology. There are representative characteristics of *emergence* - prominent impact, high growth rate, novelty, and development effort and developers' capabilities, which can be considered in each module and lead to explore emerging areas for R&BD. For the sake of identifying emerging R&BD areas, first of all, potential but unrevealed R&BD areas are explored by analyzing the relations between patents and trademarks that can represent technology and business, respectively, based on topic modeling and network analysis. The hidden links in the network become potential areas where is possible to make a business, or implement products, through utilizing new technology, and they are discovered by common neighbor-based link prediction which is based on network analysis. Second, promising R&BD areas are chosen with consideration of the market and competition status, relying on trademark analysis from the viewpoint of the applicant. This is because the information related to the applicants of IPRs provides competition intelligence that is useful to identify promising areas, conducted by network analysis either. The emergence of potential R&BD area can be evaluated at this step, the definition of emerging R&BD areas considered. Then, the specific product or business will be suggested through analyzing design patents that provide appearances related to goods. It helps decision-makers to come up with new concepts of business.

The suggested approach intends to reflect the unique features of each IPR, which has a different scope of protection and bibliographic information. The patent has lots of bibliographic information about new technology, and the design is related to the appearance of the product based on technology. The trademark can provide the industrial environment, for example, the scope of business of a specific firm. Thus, the IPR is able to provide patent and technology intelligence, as well as competitive intelligence, through IPR relation analysis. The proposed approach explores emerging business opportunities based on technological information, thus it is suitable to make technology-driven business planning including market and product.

The remainder of this study is divided into five sections. Section 2 provides the review related to IPRs and R&BD, and describes the methodologies utilized to investigate emerging R&BD areas, network analysis, and GTM. Section 3 shows the basic concept and overall process for integrated IPRs management; and then, Section 4 applies the process to OLED technology to illustrate the suggested approach. Section 5 provides insights based on the results that are induced by the proposed framework, and finally, Section 6 brings to a conclusion with the academic and managerial contribution of this study.

2. Background

2.1. Research and Business Development (R&BD)

R&BD is defined as the process of developing technology for value creation through integrating market and innovation, which aims at developing both market-oriented technology and technology commercialization (C. Chen, 2009). Although it is often called in various ways – technology commercialization, or Commercialization and Development (C&D), it is a common activity and process among diverse definitions that create value through the application, diffusion, and transfer of R&D outcomes. The R&BD seeks to maximize the performance of R&D by adjusting goal and direction step-by-step to enable research and commercialization, in order to examine feasibility from an early stage. Technology commercialization has become an activity of focal attention and a force to be reckoned with, because technology has been an important source of national development and growth, as well as crucial for survival in a competitive world (Kumar and Jain, 2003; Sohn and Moon, 2004). In contrast to these trends, the success rate of technology commercialization is relatively low, notwithstanding the great amount of technological output, such as papers and patents. Great attention has been shown to increase the efficiency of technology commercialization and success factors impacting on technology commercialization (C. Chen, 2009; R. Chen, 2009; Shibata et al., 2010; Sohn and Moon, 2003; Svensson, 2007). Sohn and Moon (2004) proposed a decision tree analysis of DEA, resulting in forecasting the degree of commercialization efficiency, considering the environmental characteristics of new technology. This acquired posterior probability for effective commercialization projects when it only had the information about the environmental factors. Conceição et al. (2002) described initiatives within an innovative firm, with the aim of improving the performance of technology commercialization, by explaining general policies and its entrepreneurial culture. C. Chen (2009) and R. Chen (2009) studied the effects of technology commercialization, incubator and venture capital support on new venture performance from a resource-based view, through regression analysis.

There were other attempts to identify business areas with a systematic approach (Lee et al., 2009; Seol et al., 2011). Lee et al. (2009) suggested a technology-driven roadmapping process that starts from capability analysis for technology planning, and ends with business opportunity analysis for market planning through text mining, network analysis, citation analysis, and index analysis. Implications from these various techniques were represented in four maps – actor-similarity map, actor-relations map, technology-industry map, and technology-affinity map. Seol et al. (2011) proposed the exploration of new business areas based upon the relative strength of firms' technologies among the potential business opportunities. They sought to reflect the weighted quality of patents by using data envelopment analysis, as well as considering competitors, in adding new businesses to the firm. Likewise, this paper aims to discover emerging R&BD areas through analyzing IPRs data related to both technology and business. It will find out new business opportunities and suggest directions for implementation of product or market entrance with regard to technological aspects.

Table 1
Available information and application of each IPR.

IPRs	Representation	Information		Classification	Application
Patent	Technology	Description of new inventions	· Technology classification, Keyword, Assignee, Inventor, Citation	IPC (International Patent Class)	Patenting strategy
Design	Product	Appearance and form of products applying technology requiring mass production	· Keyword related to color, materials, finishing Classification	Locarno International Classification	Design strategy
Trade-mark	Business	Applicable fields of goods and services similar group of product	· Classification, Owner, Similar group of product	NICE International Classification	Business plan

2.2. Intellectual property rights

With the advent of a knowledge-based economy, the role of IP has shifted from legal rights, which protect from imitation by competitors and the outcome of R&D, to the intangible assets to gain competitive advantage (Reitzig, 2004). In this way, the IPs have an important role to establish strategy from the viewpoints of technology, product, and business. In particular, each IP, such as patent, design, and trademark, has its own unique features as shown in Table 1, and these should be considered at the early stage of R&D planning, when planning technology or market for successful R&BD, like productization or servitization.

The patent is the best-known way to create competitive advantage by protecting and assigning exclusive power for new inventions (Bennett, 2002; Benson and Magee, 2015; Bhaduri and Mathew, 2003; Kim et al., 2016), since the patent has a relatively huge amount of information related to new technology, and has been utilized for a wide range of purposes, such as technology planning, forecasting, and development (Abraham and Moitra, 2001; Daim et al., 2006; Fabry et al., 2006; Joung and Kim, 2017; Yan and Luo, 2017). The patent can be classified into three types – utility patent, design patent, and plant patent, according to the object to protect. This paper focuses on the design patent, in that design is related to goods and services, which have the possibilities of mass-production through applying technology. Since a patent is granted to anyone who invents a new, original, and ornamental design for an article of manufacture, it is able to provide the critical information related to goods based upon the new technological invention. The trademark is an exclusive right for products and services designated by an applicant, and it can allow owners to gain profit or dispose of the trademark. A strong brand protected by trademark is able to generate promotional advantages, as well as technical advantages.

All sort of IPs have abundant information, but to develop a design patent map is more difficult than to establish a patent map, in that design tends to protect forms and appearances, and is a rather subjective determination (Chen and Chen, 2007; Gwak and Sohn, 2017). Among several attempts, C. Chen (2009) and R. Chen (2009) tried to generate a design patent map with design patent examiners and designers by using a multi-dimensional scale that was used to identify the distribution of design patents, through converting primary data to secondary data, to find the distance between patent samples. This was an innovative study to find implications through developing a design patent map, but subjective opinion was involved in the process of converting data. Furthermore, they proposed a novel design of a community knowledge-based patent map system with an efficient genetic algorithm-based dissimilarity visualization engine, which was able to transform the dissimilarity among patents into a two-dimensional patent map (Chen et al., 2013).

2.3. Bibliometric analysis and techniques

Bibliometrics is generally defined as the quantitative study of physically published units or of bibliographic units (Broadus, 1987; Hood and Wilson, 2001; White and McCain, 1989) and is regarded as the measurement of texts and information (Norton, 2001; Daim et al., 2006). While the main point of traditional method for bibliometric analysis was to trace back academic journal citations, it is extended to forecast potential future events by understanding the past event. Bibliometric analysis makes it possible to explore, organize and analyze huge amounts of historical data which assists in identifying hidden patterns for researchers (Ferrara and Salini, 2012).

According to bibliographic data and method, it is possible to achieve the goal of bibliometric analysis. For example, the bibliometric analysis is able to identify similar profiles of scholars through cluster analysis based on a collection of patents or publications. Bibliometric data analysis helps identifying and representing associations between bibliometrics variables, or studying the dependency/relationship between bibliometric variables by association rules (Li et al., 2009), tree-based models (Coronado et al., 2011), network analysis (Ding, 2011; Kajikawa and Takeda, 2009). It serves as a tool for summarizing the collection of bibliographic data by extracting representative topics under an assumption that the collection of documents is described in terms of the topics. Likewise, our paper utilizes three techniques – LDA, network analysis, and GTM - which are suitable to analyze bibliographic data and to explore hidden patterns among big data (Table 2). The LDA assists in summarizing and deriving topics among a large number of documents, and network analysis is used to extract hidden patterns which are potential R&BD areas based on existing relations between technology (Patent) and market (trademark). Along with these techniques, the GTM also supports to investigate emerging and promising R&BD areas by identifying potential vacuums in the map.

2.3.1. Topic analysis

The topic modeling has become a popular approach to identifying future hidden topics from a corpus of text (J. Chang et al., 2009; S. Chang et al., 2009; Hall et al., 2008; Monemi and Rost, 2016). The first generation of topic models was able to capture different topics covered by a collection of documents, and the Latent Dirichlet Allocation (LDA) is the best-known model in this generation. It is based on a Bayesian statistical technique to infer what each word might mean on the basis of its neighboring or co-occurrence of words (Blei et al., 2003). The basic assumption of the LDA is that each document is a mixture of topics, where each topic is a distribution over words. Each word appeared in the document can be assigned to one of topics with some probability, and the meaning of the word might change with the association of other words inside the document (Monemi and Rost, 2016). In contrast with text mining based on occurrence frequency, the LDA has the merit of reflecting the context of sentences in documents (Pépin et al., 2017). The LDA has been utilized to derive topics as well as to identify the evolutionary process or development path of technology by analyzing trends of topics and keywords (Furukawa et al., 2015; Monemi and Rost, 2016). Venugopalan and Rai (2015) used topic modeling to map patents to probability distributions over real-world categories/topics. It presented a natural language processing based hierarchical technique that enables the automatic identification and classification of patent datasets into technology areas and sub-areas. Likewise, since the LDA is regarded as an even more statistically sophisticated technique for the conceptualization and identification of scientific topics (Furukawa et al., 2015; Mavridis and Symeonidis, 2012), it is appropriate to

Table 2
Description related to methodologies used in this study.

Methodology	Purpose	Benefit	Technical problems that our study desires to solve
Latent Dirichlet Allocation (LDA)	- Grouping patents depending on similar technological features	- Topic modeling based on a probabilistic model - Documents summarized as topics and detailed keywords - Clustering documents by topics based on the probabilistic distribution of words and topics	- Not relying on simple keywords - Considering co-occurrence of words - Enabling to make groups with consideration of context - Yielding more accurate
Network analysis	- Exploring potential but hidden R&BD areas	- Applicable to a bipartite network (two-mode network) - Easily identifying hidden links based on local similarity features	- Identifying and predicting hidden links between nodes based on links sharing same nodes - Solution based on a basic idea that the more common neighbors they have, the more likely they have same functions or markets - To find unexplored fields
Generative topographic mapping (GTM)	- Exploring emerging and promising R&BD areas	- Finding vacuum areas intuitively from the GTM-based map - Utilizing vector information representing vacuum cells - Using bibliographic data including both descriptions and owners (or authors)	- To assess and detect the emergence of hidden and potential areas derived from network analysis

grouping patents based on probabilistic distribution. Thus, this paper exploits the LDA for grouping patents whose technological functions are similar because the LDA is able to derive topics taking co-occurrence of words into consideration, and documents involved in same topics can be grouped respectively.

2.3.2. Network analysis

The social network is composed of a set of 'entities', like persons, groups, or organizations, and a 'relation' on those entities, such as friendship, inter-personal communication, or agonistic acts (Butts, 2008). The analysis not only seeks to predict the structure of relationships among social entities, but also investigates the impact of that structure on other social phenomena. Thus, social network analysis has been widely utilized to identify the spillover of knowledge flow based upon bibliographic data such as citation relationship between patents (J. Chang et al., 2009; S. Chang et al., 2009; Érdi et al., 2013; Kim et al., 2008; Park et al., 2005; Yoon and Park, 2004). Beyond this, there are many attempts to exploit network structure to examine the complex social relationship, by extension, the challenge of missing and hidden links that has arisen in recent times. To predict missing or hidden links aims to estimate the likelihood of the existence of links between two agents based on the observed links and the attributes of agents (Liu et al., 2013). This has attracted attention from researchers in social networks, because link prediction is helpful for understanding the evolution of real networks. Also, predicting links accurately in a network can offer crucial evidence about the rules that drive its evolution (Lü et al., 2015). The link prediction is conducted by calculating the similarity between nodes, under the assumption that the higher the similarity between nodes, the higher the likelihood of the existence of a hidden link. There were several studies on link prediction for non-bipartite graphs (1-mode network) with a variety of methodologies, which are generally classified as node-based and topology-based link prediction (Liben-Nowell and Kleinberg, 2007). The node-based similarity depends on the actions or attribute of nodes, such as Euclidean distance, cosine similarity, and vector space model. On the other hand, topology-based link prediction relies on topological information, and it is compartmentalized as (1) neighbors-based, (2) path-based, and (3) random walk-based link prediction. Among them, prediction based on neighbors has been widely utilized, and there are many forms of link prediction relying on neighbors, such as common neighbors, Jaccard coefficient, preferential attachment, and Salton cosine similarity. Among them, the common neighbor technique, which counts the number of neighbors that two nodes have in common, is selected in this paper. This approach is based on an idea that the more the number of common neighbors of two nodes u and v, the more the probability of a future relationship between node u and v (Gupta et al., 2015). For an undirected graph, similarity for an edge is defined as follows:

$$\mathrm{Similarity}(\mathrm{u},\mathrm{v}) = |N(u) \cap N(v)|$$

As referred to earlier, it is intuitive to figure out hidden links in a social network, and simple to apply to a bipartite network (2-mode network). We assume that two elements - patent and trademark in this paper - are likely to have same functions or markets if they share common neighbors. In other words, the more common neighbors they have, the more likely they have same functions. Under this assumption, hidden links between patents and trademarks are discovered through detecting non-connected links sharing common neighbors. These hidden links derived from common neighbor technique are regarded as a candidate of potential R&BD area and it will be further explained in Section 3.

2.3.3. Generative topographic mapping (GTM)

Another methodology is generative topographic mapping (GTM), which is a sort of non-linear latent variable model, for which the parameters of the model could be determined through the EM algorithm (Bishop et al., 1998a, 1998b). The GTM is widely used as a methodology for grouping and data visualization, because it offers more flexibility by adopting soft clustering through the responsibility of each data point. In addition, it is exploited for patent mapping and identifying vacuum information in an objective way, as compared to principal component analysis, which is quite qualitative when finding vacuums and clusters (Son et al., 2012). They proposed a GTM-based patent map that aimed to automatically detect and interpret technology vacuums, so a GTM-based patent map provides a grid-based two-dimensional map, in which each patent is mapped to the relevant grid. Jeong and Yoon (2013) suggested a method that derived essential patents through GTM-based standard and patent maps. It included a systematic process that identified vacuums on a standard map in a specific technological field, and enabled analysts to find candidates for promising essential patents, rather than relying on experts. Jeong et al. (2015) proposed the GTM-based patent roadmap, where GTM served to explore novel areas for patent development. Information included in the vacuums of GTM-based patent map was connected to nodes on the patent roadmap.

Likewise, the GTM makes it possible to develop a map that consists of keyword vectors to detect vacuums intuitively in this map, using bibliographic data such as patents or publications. Each cell including vacuum possesses information related to keyword vector, which enables to investigate vacuum cells in detail. It relies on one of the advantages of the GTM that the neighborhood-preserving nature of the GTM mapping is an automatic consequence of the choice of a continuous function. Vacuums of the latent space are automatically identified and the GTM offers a probabilistic framework for the automatic identification of latent variables by inversing mapping (Bishop et al., 1998a, 1998b). That is to say, the continuous function of the GTM represents a mapping from the latent space onto the data space, as a result, a grid vacuum can be reversely mapped into real data using this characteristic. The GTM has decisive advantages in visualizing multidimensional information as well as investigating the characteristics of unoccupied cells on a map (Jeong and Yoon, 2013).

Therefore, our study will apply the GTM technique to discover hidden but potential R&BD areas because they map objects (such as nodes) and visualize relations, leading to detecting unrevealed ones. While the neighbor-based link prediction based on network analysis is able to find out hidden nodes depending on similar relations, the GTM is used to define potential R&BD areas where specific firms are able to enter without high competence. Our approach will establish the GTM-based map by mapping firms having ownership of trademarks as compared to previous studies, which provides the market status and the scope of business for each company. From the viewpoint of bibliographic data, network analysis-based link prediction intends to focus on textual data and relationships between elements and GTM-based map observes information related to the owner of trademark for finding the promising R&BD areas.

3. Methodology

3.1. Basic concept

This research suggests an approach to comprehensively analyze IPRs, and to explore hidden but promising R&BD areas in the near future. In other words, we define these areas as 'emerging R&BD areas' that have several characteristics - potential impact, high speed of growth, novelty, development effort and developers' capabilities – in reference to previous study (Lee et al., 2018). For the purpose of discovering emerging R&BD areas, those features are dealt with in each module. After finding unrevealed R&BD areas through network analysis, the emergence of R&BD areas is evaluated by keyword analysis and market indicators. In particular, patents and trademarks are utilized as fundamental resources to search promising R&BD areas, because they are respectively representative of technology and business. From this viewpoint, the relationships between technology and business are analyzed by network analysis and topic modeling, focusing on the similarity of contents. The potential R&BD areas are identified in the form of the hidden and missing links in the established network, which is composed of two actors – patent and trademark – in advance. Then, promising R&BD areas are chosen with regard to the status of market and competition through network analysis from the viewpoint of relations between actors. The idea of exploration for emerging R&BD areas is to discover latent relationships between IPRs that may have been missed out. In order to provide more meaningful insight for successful R&BD, possible business cases are suggested by exploiting design patents that show the detailed appearances of goods. Fig. 1 shows our approach to explore potential and promising R&BD areas, and the detailed process is described in the next section.

3.2. Overall process

3.2.1. Exploration for potential R&BD areas by analyzing the relation between patent and trademark

This study desires to discover emerging market opportunities by considering the status of technology development for achieving successful commercialization. From this viewpoint, it is necessary to investigate just what technology is able to provide and implement novel items for business. Thus, the patent is utilized to examine technological functions, because it has specification and description related to the technology. Patents are collected from national patent databases, and then keywords that are able to represent the functionality of technology are extracted by LDA. The number of topics and keywords comprising a topic is defined by researchers, giving consideration the amount of data and purpose of analysis. Since the probability that a document is involved in a specific topic is also calculated, it is possible to cluster documents into a group (a topic). As a result, patents are grouped on the basis of similar features, and each cluster can be regarded as a function of technology.

For the purpose of examining the relationship between technology and business, this study applies 2-mode network analysis by using patents and trademarks. Each trademark has 'Goods and services', which means fields protected by the legal right of the trademark application in the form of words. Thus, keywords related to each class of trademark classifications are extracted in reference to explanatory notes, and they are defined as "business keywords". In order to identify which area is implemented by novel functions of technology, the occurrence frequency of business keywords in patent documents is counted, and a patent-trademark relationship matrix is established that consists of two axes – technology (patent cluster) and business (business keyword derived from trademark), as shown in Fig. 2. This matrix establishes a two-mode network that is composed of two nodes – patent cluster, and class of trademark. The edge is based on the occurrence of business keywords in patent documents; in other words, a specific patent can be implemented by specific keywords, because patent documents have information related to this embodiment of this technology or function.

After developing the technology–trademark network, potential R&BD areas are identified by the common neighbor-based similarity in network analysis. The common neighbor-based technique considers that two nodes are more similar if they have many common features, so it makes a new link between them, as shown in Fig. 3. The potential R&BD areas are first identified by the number of a common neighbor, because the objective of this paper is to discover unrevealed areas where similar technology is already implemented, or has entered, but specific technology has not yet been implemented. From this viewpoint, nodes with weak ties that have low centrality are primarily chosen, and then common neighbors are calculated, and hidden links are discovered.

3.2.2. Exploration for promising R&BD areas by analyzing trademarks

The potential R&BD areas at the prior step are assessed by trademark analysis, in order to find emerging and promising R&BD areas. The trademark has information about 'goods and services' and 'owners', so it is useful to identify market statuses, such as competitors, and fields dominated by other companies. The first module intends to find unexplored fields, whereas this module aims to assess and detect the emergence of hidden and potential areas (which are derived from the first module) with considering characteristics of emergence. Aforementioned, emerging R&BD areas have a few of features – (1) prominent impact, (2) growth rate, (3) novelty, and (4) development effort and developers' capabilities. Three of them (without development effort and developers' capabilities) are reflected as the indicator for evaluating the value of each cell in the GTM-based firm-trademark map. Another feature is also identified on this map and assessed by market and competitor status.

To do this, at first, trademarks relevant to potential areas derived by the relation between technology and business in the previous module are collected from the national trademark database, which provides bibliographic data of trademarks, along with the patent database. Second, business keywords are again extracted from trademark documents, especially in the part of 'goods and services', by term frequency-inverse document frequency (TF-IDF) which means the importance of each keyword compared to the number of document. The business keywords are limited to the potential R&BD areas defined at the prior step. Third, a firm-business keyword matrix is developed for the

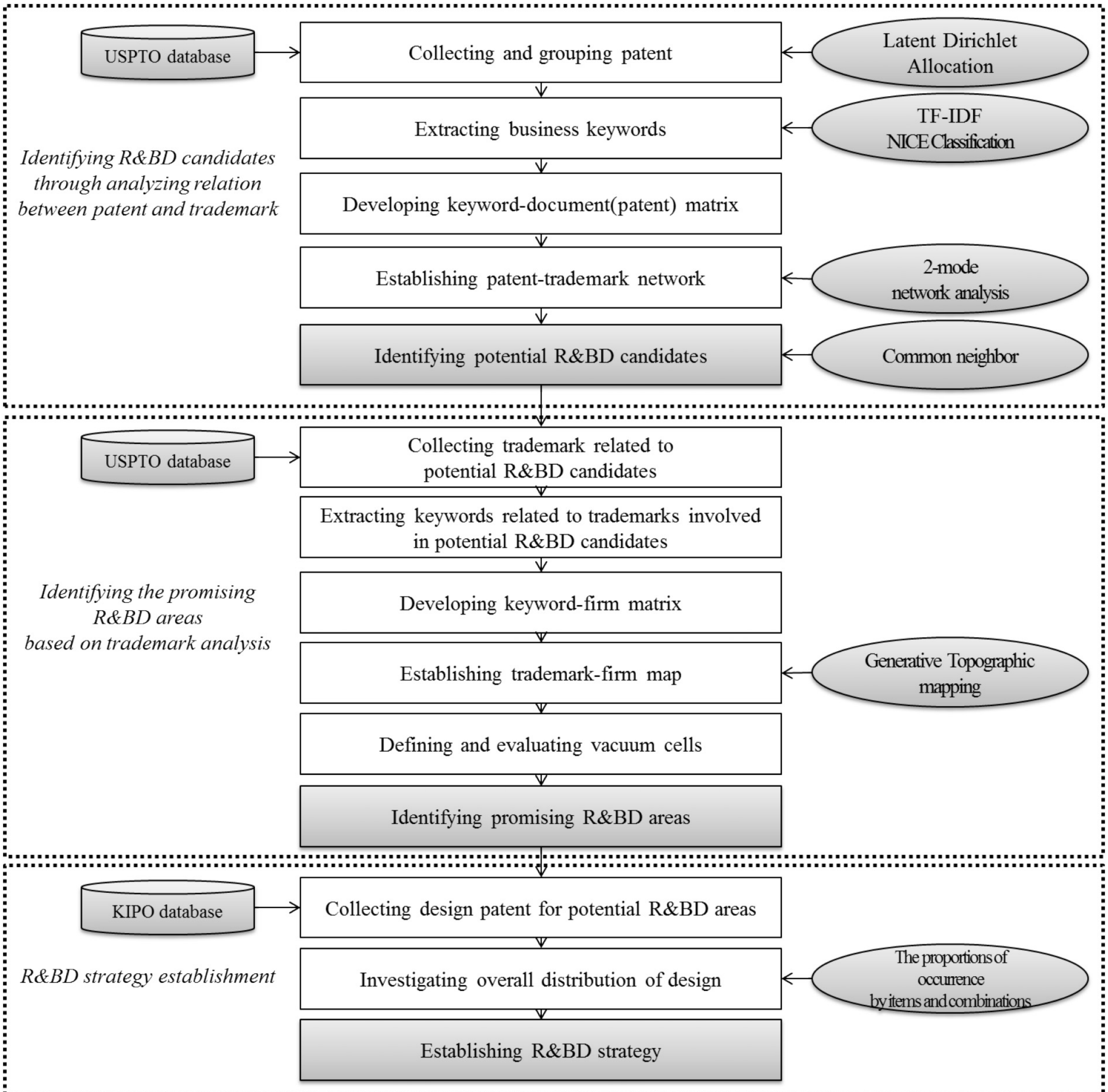

Fig. 1. Research framework.

Patent documents involved in each cluster

Keywords derived from trademark

	Patent cluster 1	Patent cluster 2	Patent cluster 3	…	…	Patent cluster n
Business keyword 1	1	0	0	…	…	1
Business keyword 2	0	1	0	…	…	0
…	…	…	…	…	…	…
Business keyword m	1	1	1	…	…	1

Fig. 2. A patent-trademark relationship matrix.

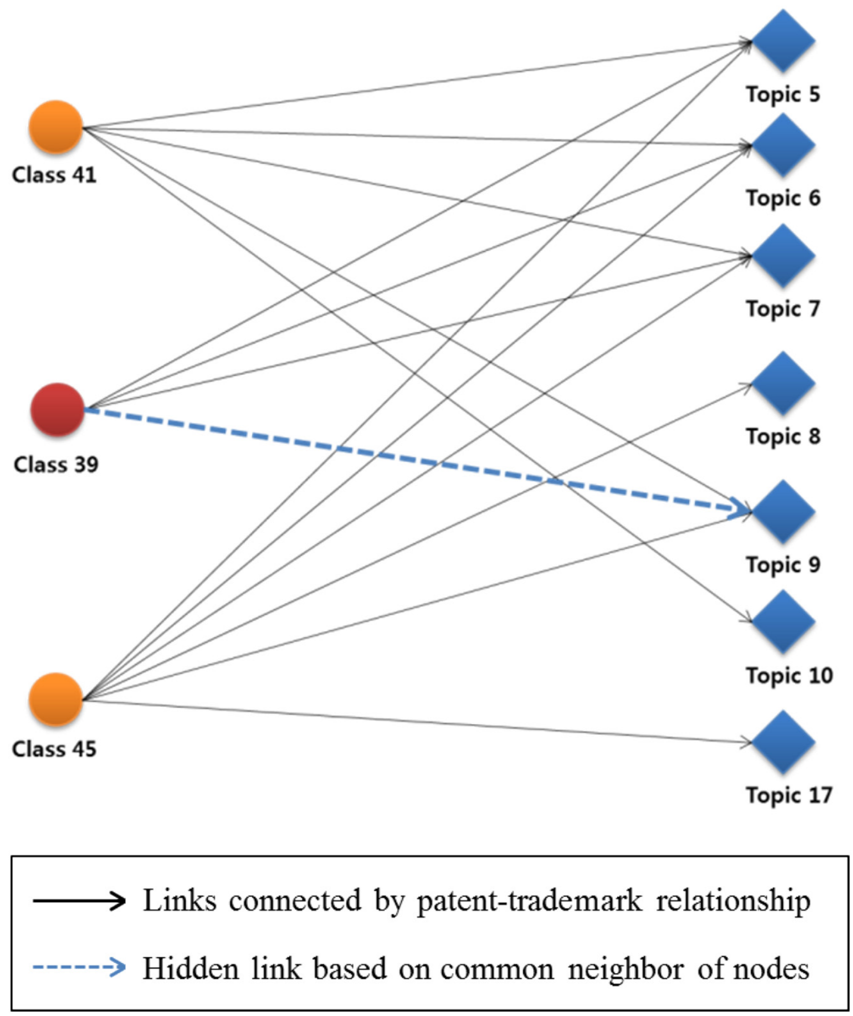

Fig. 3. Identifying potential R&BD areas based on common neighbors.

purpose of analyzing market status; in other words, this step examines which firm carries on the business. This is identified by whether the business keywords occur in the trademark document. Finally, a firm-business map is established by generative topographic mapping (GTM), which is useful to discover vacuums, as well as to visualize data, as shown in Fig. 4.

In particular, since the GTM is able to identify detailed information about vacuums, a GTM-based firm–business map provides the businesses that a firm carries out, and empty areas for doing a business in the form of vacuum. The vacuum is regarded as areas where do not a business yet, and no trademarks are mapped in these vacuums. Thus, they might be considered as a relatively novel area where other competitors have not yet entered into. However, it is unclear that the vacuum may be really novel and promising but hidden until now, or might be not a valuable area, making other competitors give up launching a business. Thus, it is necessary to identify whether the business areas corresponding to vacuums are really promising or not, with the aim of exploring hidden and promising R&BD areas while avoiding worthless areas. The evaluation process for finding promising R&BD areas is as follows. In order to evaluate the degree of emergence of vacuums, the information of keyword vector, whether the keyword is included in the vacuum, is utilized. It is because the GTM has an advantage that the continuous function defines a mapping from the latent space into the data space; this vacuum feature on a grid can be reversely mapped into real data. The process for evaluating the emergence of vacuum cell is based on an assessment of business keywords that is calculated by the weighted sum of score by keyword. Using the pre-defined list of business keywords from the trademark documents to develop the GTM-based trademark-firm map, each keyword is evaluated by five-point scale for each indicator (Table 3). Two indicators – technological competence and marketability – are defined, and they are divided sub-indicators which consider different characteristics of 'emerging' area such as prominent impact, growth rate, novelty, and development effort and developers' capabilities from both technology and market aspects (Table 3). The business keywords defined at the prior step were evaluated by a 5-point scale based on these indicators. At this time, quantitative indicators such as market size, growth rate are scaled by cut-off value as 5-point scale, and qualitative indicators which require to subjective judgment are also scaled by the degree of correspondence with the definition of indicators. The evaluation for a business keyword is conducted by relying on reviews on market or technology. Finally, the degree of emergence for each vacuum is calculated by the weighted sum, and the vacuum cell with the highest degree of emergence was defined as an emerging and promising R&BD area. The weighed sum of the vacuum is called the '*degree of emergence*' in this paper, and the cells with the highest degree of promise become emerging R&BD areas. Then, detailed information is suggested by extracting and assessing unique keywords of each vacuum that do not exist in other cells, based on the keyword portfolio where two axes – technological competitiveness and marketability – exist. Finally, keywords with a high value of the two indicators become definitive R&BD areas.

3.2.3. R&BD strategy establishment based on design patent

After identifying promising R&BD areas, the R&BD strategy is

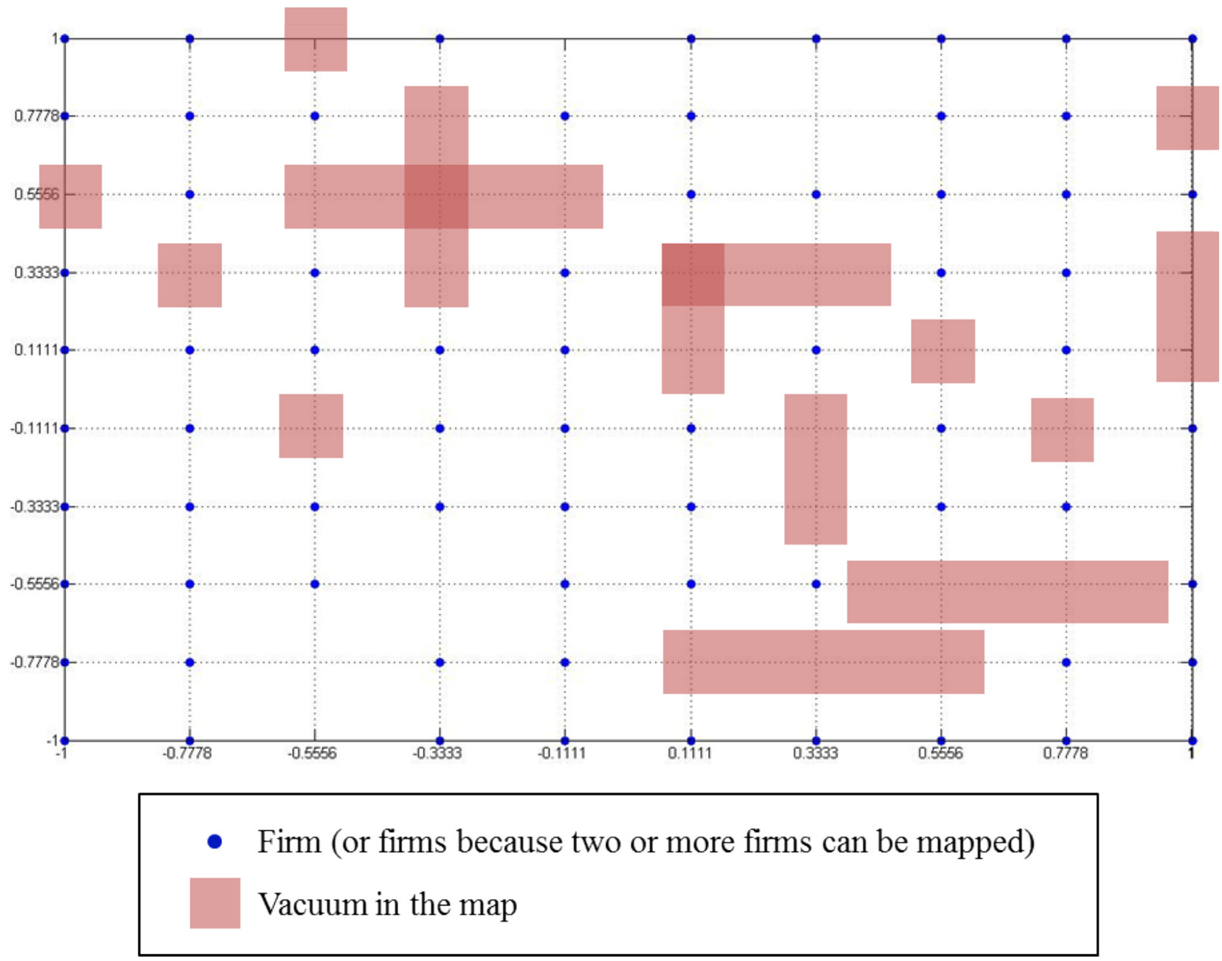

Fig. 4. The example of GTM-based firm-business map.

established by analyzing design patents, which is different from patents, in that the design patent aims to protect the appearances of products or services (e.g., graphic user interface). In particular, the Korea Patent and Trademark Office provides a database –the 'design map', where the design patents are collected from the United States, Japan, Europe, and Korea patent databases, and are classified by the products and features of each product. The design map developed by KIPO categorizes design patents into eight classifications, such as home appliances; digital devices including computers; furniture and interior; outside construction; household items and stationery; fashion accessories and cosmetics; health, sports and leisure; transportation; and graphic design. Each category has many products, and there are a total of ninety product classifications. Because each product has common features by product lines, but are distinct from other products, the map is structuralized according to formal features, such as formal shape, and location of operation key. The combination of features represents a product, and it will be evidence for checking whether the combination can be technically implemented or not, through investigating the status of design applications.

Consequently, the R&BD strategy is established with the goal of avoiding previous appearances, as well as technically checking the feasibility for combinations of each component or feature. It is suggested in two ways, and the first one is to offer the overall distribution of design patents. All kinds of design patents are gathered, and the number and proportion of each feature are calculated, leading to the overall distribution of design patents within a specific category of product. Another viewpoint for establishing R&BD strategy is to suggest frequently occurring product designs, which mean the combinations of

Table 3
Indicators for evaluating keyword.

Criteria	Indicators	Description
Technological competitiveness	Novelty	The extent to which the technology is new in the market
	Technology life cycle	The stage in the life cycle of technology which consists of four stage – introduction, growth, maturity, and decline
	Type of technology	– Basic technology, Applied technology, Development (within the framework of technology) – Material, component or process, product or service (within the framework of product and service)
	Applicability	The extent to which the technology can be applied to other fields of technology
	Possibility of substitution	The extent to which the technology is substituted by other competitive technology
Marketability	Market size	The size of market relevant to product/service keyword
	Growth rate	The growth rate of market relevant to product/service keyword
	Ease of commercialization	The extent to which the product or service can be implemented and commercialized with exploiting technology
	Regulation	The extent to which regulation is strict when going into the market
	Extensibility	The extent to which the scope of market is extended by applying to other markets

features. This makes it possible to automatically and quickly identify previous appearances of products, and to offer considerations or directions for R&BD when implementing the novel technology.

4. Results

4.1. Data collection

In order to illustrate the suggested process for identifying potential and promising R&BD areas, this paper selected Organic light-emitting diode (OLED) technology. This has been a backbone of IT devices, and continues to thrive, pursuing oncoming generations of displays. In particular, new technology, such as transparent displays and rollable displays, has infinite possibilities of applying to completely novel areas, as well as the existing industry, while growing the display industry. Thus, patents were collected from the United States patent database (USPTO), and the period for patent collection is limited to the last five years, because the objective of this study is to find promising R&BD areas based upon new patent (technology or functions). As a result, 771 patents that were registered from 2012 to 2016 were collected.

4.2. Exploration of potential R&BD areas by analyzing the relations between patent and trademark

Twenty topics and 15 keywords by each topic related to technology were extracted, and patents were grouped through LDA, as shown in Table 4. The number of topic is defined by two ways - statistic model and domain knowledge of researcher. In the statistical model, there are four well-known metrics proposed by Griffiths and Steyvers (2004), Arun et al. (2010), Cao et al. (2009), and Deveaud et al. (2014) and they are provided by R package '*ldatuning*'. Since the statistical models enable to narrow the range of choices for topic number, a simple approach to deciding the topic number is to find extremum by integrating the results from these metrics in order. The statistical methods suggested by the four studies are excellent to find optimal number of topics even though they were based on different models respectively. First, Griffiths and Steyvers (2004) selected Bayesian models and computed an estimate of posterior probability by varying the topic values through running Markov chains. Cao et al. (2009) utilized an idea of clustering process based on density, assuming that the similarity will be as large as possible in the intra-cluster while as small as possible between interclusters. This study integrated the clustering approach based on density into LDA and it enables to adaptively select the appropriate number of topics. Arun et al. (2010) and Deveaud et al. (2014) focused on maximizing the information divergence, especially Kullbaack-Leibler divergence which is used to measure the difference between two probability distributions and between all pairs of topics. The other metric is perplexity, which is the most common way to evaluate a probabilistic model is to measure the log-likelihood of a held-out test set. It measures how well a probability distribution or probability model predicts a sample. They considered different ideas and models when integrating them to LDA for finding optimal number of topics with different characteristics, resulting in excellent performance when finding optimal topic numbers. Although these metrics suggest the optimal number of

Table 4
Topics and keywords related to OLED technology.

Topic	Title	Keywords
1	Process by vacuum thermal evaporation	Deposit, material, mask, substrate, pattern, evaporation, source, measure, use, surface, amount, process, align, method, chamber
2	Structure of OLED	Electrode, drive, transistor, film, thin, layer, gate, connect, capacitor, semiconductor, form, emission, drain, control, switch
3	Organic materials for OLED	Group, compound, substitute, formula, prefer, use, unsubstitute, atom, organ, mmol, wherein, material, alkyl, option, independent
4	Structure of OLED display panel	Data, image, value, display, control, generator, block, method, luminance, signal, correspond, plural, pixel, accord, panel
5	OLED display and interface	Voltage, transistor, signal, scan, power, supply, data, control, electrode, node, drive, source, connect, coupling, ole
6	Switch drive technique and circuit	Transistor, circuit, terminal, signal, drive, connect, control, switch, voltage, ole, gate, capacitor, node, phase, program
7	TFT array substrate	Layer, electrode, form, insulation, gate, pattern, metal, conduct, TFT, active, drain, substrate, source, contact, oxide
8	OLED display with spalled semiconductor	Glass, substrate, layer, block, device, OLED, temperature, metal, structure, material, heat, polymer, packaging, semiconductor, transparent
9	System and method for remote display	Display, device, user, system, computer, screen, electron, control, communication, sensor, operator, least, data, use, receiver
10	Driving circuit and electro-optical device	Voltage, drive, TFT, data, sensor, OLED, refer, display, switch, supply, connect, device, accord, threshold, circuit
11	Touch panel and display device with touch panel	Display, touch, panel, surface, device, terminal, connect, embodiment, board, cover, circuit, side, form, member, pad
12	Electronic device and method for manufacturing	Film, substrate, layer, form, use, method, barrier, bond, device, laser, resin, silicon, beam, adhesive, manufacturing
13	Organic electro luminescent display and method	Layer, form, organ, display, substrate, OLED, film, material, encapsulation, electrode, inorganic, thin, surface, light emit, member
14	OLED device architecture	Light, emit, organ, diode, display, accord, exemplary, OLED, source, emission, refer, plural, relation, claim, position
15	Method of manufacturing for OLED array substrate	Color, subpixel, display, OLED, red, blue, green, filter, light emit, light, white, device, unit, structure, arrange
16	Flexible electronic display	Display, plural, direct, substrate, OLED, line, form, pixel, open, edge, connect, wherein, arrange, dispose, adjacent
17	Organic devices having a fiber structure	Layer, electrode, emission, form, organ, hole, material, OLED, transport, electron, auxiliary, inject, device, thick, anode
18	Retardation compensation element for liquid crystal display	Display, optic, image, light, transparent, reflect, polar, crystal, liquid, lenses, panel, mirror, device, surface, system
19	Nozzle-droplet combination techniques to deposit fluids in substrate locations within precise tolerances	Nozzle, print, use, target, droplet, volume, process, head, combine, deposit, ink, substrate, scan, represent, select
20	Flexible display devices with bridged wire traces/chamfered polarization layer/gate-in-panel circuit	Layer, flexible display, bend, trace, conduct, wire, provide, base, circuit, stress, support, active, design, allow

topics, it is limited to the range of optimal number and there is at risk for forcing noise information into a few of topics. At this point, this paper utilized four methods comprehensively rather than selecting one method among them, leading to identify the coarse range of topic numbers. As a result, the approximate scope of topic numbers ranged from 20 to 35 which was the extreme value (maximum and minimum) among other values based on excellent four models proposed by previous studies using R packages '*ldatuning*'. Then, the actual point was decided by reviewing distribution of words and topics to best account for the set of documents. In other words, the final work conducted by researchers after identifying coarse range of topic numbers based on the statistic models is to adjust the number of topics by considering whether a topic is able to account for the words used in a set of documents. The decision on the number of topics needs to be made by combining the statistical model and the domain knowledge of researchers, taking account of possibility of interpretation and feasibility, and utility for research questions. Thus, we determined the number of topic from this perspective, especially focusing on which number of topics is able to summarize and explain overall corpus extracted from patent documents.

Based on the statistic metrics, the range of optimal topics was 20 to 35, and then the number of topics was defined as 20 on the evidence of explanation ability of topic for documents and feasibility. Thus, in similar with the number of topics, the number of keyword is decided within the framework of how many words are able to represent a topic, because the title of topic is named after the keywords contained in each topic after extracting topics and keywords. At the same time, the business keywords were chosen on the basis of the explanatory note of NICE international classification that is commonly used as a worldwide classification for trademark. Then, whether the business keyword occurred in patent documents by each topic (cluster) was identified, which led to the network between patent and trademark, as shown in Fig. 5.

From the network established, nodes with a low degree of centrality were chosen, because they were open to connect with other nodes from two viewpoints – topic and trademark class. There was a node 'topic 3' with low centrality from the view of topic, and there were low centralities of nodes 'class 39', 'class 45', 'class 26', 'class 35', 'class 12', 'class 18', and 'class 27'. After selecting nodes with a low degree of centrality, nodes with high similarity based on common neighbor were selected, as shown in Tables 5 and 6. For example, 'Topic 3' was similar to 'Topic 4' and 'Topic 18', in that they shared common neighbors with 'Topic 3'. In order to find hidden links, other 'class' nodes that were connected with 'Topic 4' and 'Topic 18' were respectively investigated, because vertices between nodes only existed in between the patent topic and trademark class. As a result, common neighbors between topic 3 and topic 4 were trademark classes 2–7, 9–12, 16, 18, and 21. Similarly, common neighbors between topic 3 and topic 18 were trademark classes 1–7, 9–12, 18, and 21. Based on these common neighbors, potential vertices were determined as class 8, class 14, class 16, class 30, and class 35 in trademark, which can be interpreted at this phase as potential R&BD candidate areas.

4.3. Exploration of promising R&BD areas by analyzing trademarks

For R&BD candidates derived from network analysis, relevant trademarks were collected from the USPTO database. Because the technology related to Topic 3 has applied to business defined as Class 8 (Hand tools and implements (hand-operated); cutlery; side arms; razors), Class 14 (Precious metals and their alloys; jewelry, precious stones; horological and chronometric instruments), Class 16 (Paper and cardboard; printed matter; bookbinding material; photographs; stationery; adhesives for stationery or household purposes; artists' materials; paintbrushes; typewriters and office requisites (except furniture); instructional and teaching material (except apparatus); plastic materials for packaging; printers' type; printing blocks), Class 30 (Coffee, tea, cocoa and artificial coffee; rice; tapioca and sago; flour and preparations made from cereals; bread, pastries and confectionery; edible ices; sugar, honey, treacle; yeast, baking-powder; salt; mustard; vinegar, sauces (condiments); spices; ice), and Class 35 (Advertising; business management; business administration; office function), the searching query for trademark collection was composed of keywords included in Topic 3 and business keywords related to the five classes of trademark. As a result, 550 trademarks were collected, and 277 firms owned these trademarks. Then, core business keywords only included in the potential R&BD areas (the abovementioned five classes of trademark) were defined by using TF-IDF, and whether they were related to specific firms was identified as the matrix of business keyword–firm matrix, like a term-document matrix. From this matrix, the trademark-firm map was established through GTM, and vacuum cells were identified as shown in Fig. 6. Each cell in this map shows that each developers' (firms') efforts for specific business areas. The number of cell shows the range of market or product, and the number of firms mapped on each cell represents developers' effort for the relevant area. Thirty-six vacuums were derived from 81 cells, and then they were evaluated by weights sum for each keyword. This is because all of the vacuums have the possibility of commercialized technology caused by trouble in implementation to product or business, or low value for customers, as well as developers. Consequently, the top five vacuums – Cell 80, Cell 81, Cell 2, Cell 8, and Cell 17 (in descending order of degree of promise) - with high technological competitiveness and marketability were deducted as promising R&BD areas, and keywords involved in these cells are described in Table 7.

After selecting vacuum cells having a high degree of promise, feature keywords that were differentiated from other cells were respectively extracted. In other words, feature keywords were also defined as unique and representative keywords of each cell that did not exist in other cells, as shown in Table 7. In particular, most vacuums were able to include two or more categories of keywords, because business keywords used to analyze trademarks were chosen in different classes of trademark classification. Among them, promising business keywords that have high technological competitiveness and marketability were finally chosen as promising areas for successful R&BD, through developing a keyword portfolio, as shown in Fig. 7. In the meantime, outdated keywords that might exist in vacuum cells but were not appropriate to the latest trend could be removed.

Topic 3 is related to organic electroluminescent materials that are placed in between two conductors of different work functions on the authority of keywords contained in this topic. The technology pertaining to topic 3 is executed not only as raw materials for comprising displays or other products, but also as merchandise or service, after producing displays or screens first. In this regard, the information involved in vacuum cells is interpreted, and led to both potential and promising R&BD areas, as below.

Vacuum cell 80 had the highest degree of promise, and possessed business keywords as below – diagnostic, smartphones, engines, hinges, brake, filtering, and preservatives. As mentioned earlier, this cell also had a string of themes, and there were three kinds of keywords in this cell – vehicle (hinges, brake, engines etc.), diagnosis, and preservatives. Since the OLED is thin, and has no limitation for viewing angle, it is possible to be applied and implemented in the field that requires more rigor, like diagnosis or filtering. In particular, the OLED is widely utilized in head-up displays in vehicles, so keywords associated with vehicles often occurred simultaneously. In addition, the OLED display has recently attracted great attention in the connected car and self-driving car industry, which uses sensors and displays to navigate public roads without a human driver, through always connected with a network. Since this industry will continue to attract attention, it is possible for the OLED technology to enter the automobile industry. On the other hand, one of the advantages of the OLED is derived from the ease of chemical manipulation to adjust color, so the organic materials utilized in OLED are various. The combination of electroluminescent organic

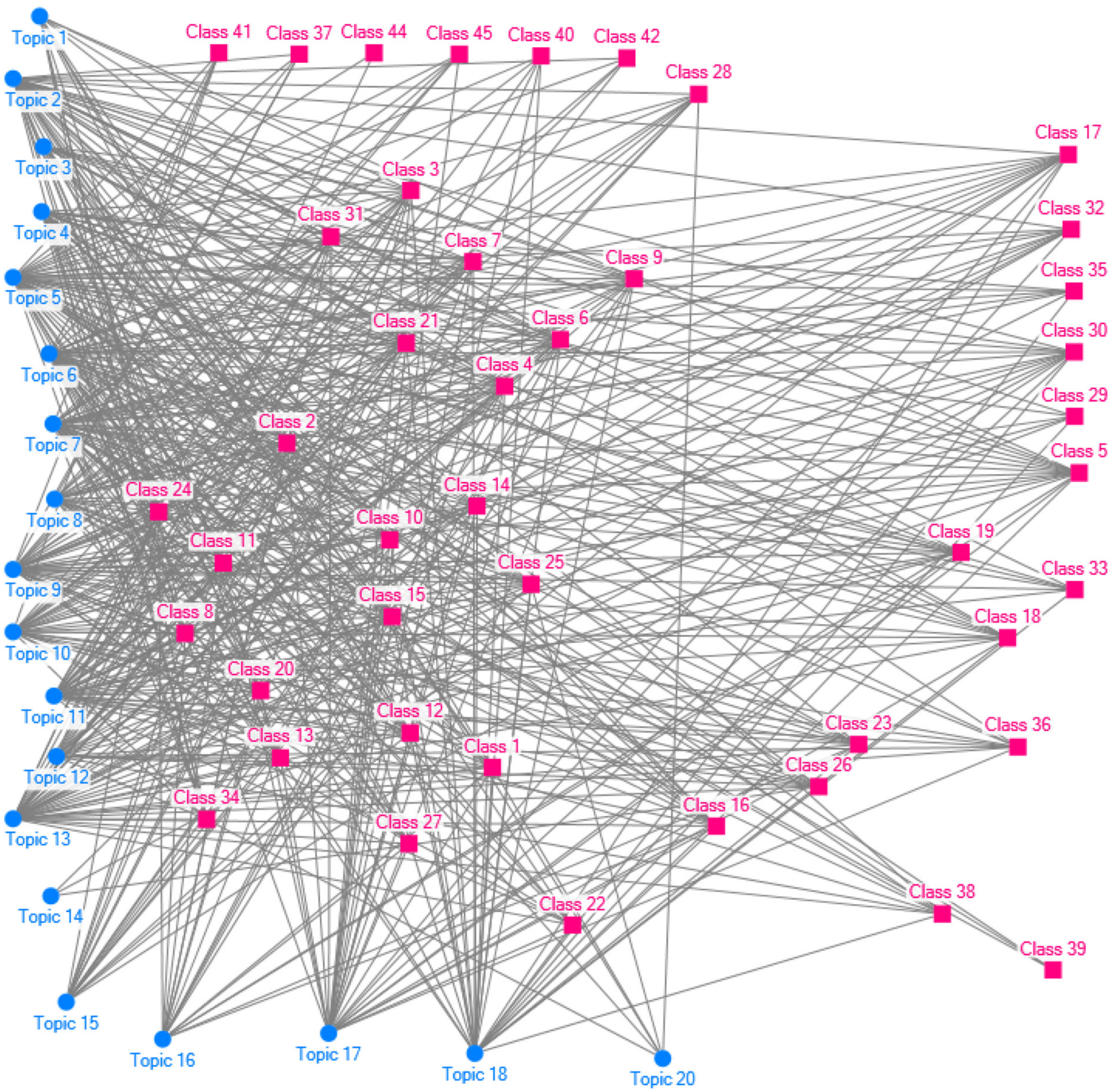

Fig. 5. The patent-trademark network in the OLED technology.

material is able to provide different colors of light. From this viewpoint, it has the possibility to expand the application scope of organic materials, such as producing preservatives.

Vacuum cell 81 contains three different themes - 'Internet', 'infant', and 'astronomy' representatively. As previously mentioned, the OLED has lower power consumption than Liquid crystal display (LCD), due to the lack of a backlight, which makes the display slimmer. This leads to the range of application being extended, like the keywords 'Internet' and 'astronomy'. Although the Internet has attracted attention in recent years, interest in the Internet or wireless network has lasted and accelerated due to smart devices, like wearable and monitoring devices within the framework of the 'Internet of things'. In particular, it is based on sensors and displays that perceive visual and tactile data, which is transmitted and connected to the system. The raw data transmitted by display and sensor is analyzed automatically, and then useful information and services are offered through the wireless network and

Table 5
Potential vertex in OLED technology based on network analysis (from viewpoint of topic).

Topic with a low degree of centrality	Topics with high similarity		Common neighbor (class of trademark)	Spurious vertex (class of trademark)	Potential vertex (class of trademark)
	Similarity	Connected classes of trademark			
Topic 3	Topic 4 (0.5)	10, 34, 6, 5, 8, 3, 4, 2, 24, 27, 20, 25, 21, 7, 11, 9, 30, 16, 18, 14, 19, 15, 12, 35, 26	10, 6, 5, 3, 4, 2, 21, 7, 11, 9, 16, 18, 12 10, 6, 5, 3, 4, 1, 2, 21, 7, 11, 9, 18, 12	34, 8, 24, 27, 20, 25, 30, 14, 19, 16, 35, 26	Class 8 Class 14 Class 16 Class 30 Class 35
	Topic 18 (0.5)	10, 13, 6, 5, 8, 3, 4, 1, 2, 22, 33, 25, 21, 38, 7, 11, 29, 9, 30, 16, 23, 17, 18, 14, 15, 12, 35, 40	10, 6, 5, 3, 4, 1, 2, 21, 7, 11, 9, 18, 12	13, 8, 22, 33, 25, 38, 29, 30, 23, 17, 14, 16, 35, 40	

Table 6
Potential vertex in OLED technology based on network analysis (from viewpoint of trademark).

Class of trademark with low degree of centrality	Class with high similarity		Common neighbor (patent topic)	Spurious vertex (patent topic)	Potential vertex (patent topic)
	Similarity	Connected patent topic			
Class 39	Class 41 (0.6)	5, 6, 7, 9, 10	5, 6, 7	9, 10	9
	Class 45 (0.5)	5, 6, 7, 8, 9, 17	5, 6, 7	8, 9, 17	
Class 45	Class 27 (0.6)	2, 4, 5, 6, 7, 8, 9, 13, 16, 17	5, 6, 7, 8, 9, 17	2, 4, 13, 16, 17	2, 13, 16
	Class 37 (0.5)	2, 5, 6, 8, 9	5, 6, 8, 9	2	
	Class 23 (0.5)	2, 5, 6, 7, 8, 9, 10, 11, 13, 16, 17, 18	5, 6, 7, 8, 9, 17	2, 10, 11, 13, 16, 18	
	Class 31 (0.5)	2, 5, 6, 7, 8, 9	5, 6, 7, 8, 9, 17	2	
	Class 19 (0.5)	2, 4, 5, 6, 7, 8, 9, 11, 12, 13, 16, 17	5, 6, 7, 8, 9, 17	2, 4, 11, 12, 13, 16	
	Class 41 (0.571429)	5, 6, 7, 9, 10	5, 6, 7, 9	10	
	Class 39 (0.5)	5, 6, 7	5, 6, 7	–	
Class 26	Class 19 (0.571429)	2, 4, 5, 6, 7, 8, 9, 11, 12, 13, 16, 17	4, 5, 7, 8, 9, 11, 12, 13	2, 10, 16, 17	2, 10, 16, 17
Class 35	Class 27 (0.545455)	2, 4, 5, 6, 7, 8, 9, 13, 16, 17	4, 5, 6, 7, 8, 13	2, 9, 16, 17	
Class 12	Class 27 (0.5625)	2, 4, 5, 6, 7, 8, 9, 13, 16, 17	2, 4, 5, 6, 7, 8, 9, 13, 17	16	16
	Class 23 (0.6875)	2, 5, 6, 7, 8, 9, 10, 11, 13, 16, 17, 18	2, 5, 6, 7, 8, 9, 10, 11, 13, 17, 18	16	
	Class 31 (0.6875)	2, 5, 6, 7, 8, 9, 11, 12, 13, 15, 16, 17	2, 5, 6, 7, 8, 9, 11, 12, 13, 15, 17	16	
	Class 19 (0.6875)	2, 4, 5, 6, 7, 8, 9, 11, 12, 13, 16, 17	2, 4, 5, 6, 7, 8, 9, 11, 12, 13, 17	16	
Class 18	Class 27 (0.571429)	2, 4, 5, 6, 7, 8, 9, 13, 16, 17	2, 4, 5, 6, 7, 9, 13, 17	8, 16	8, 16
	Class 31 (0.5)	2, 5, 6, 7, 8, 9, 11, 12, 13, 15, 16, 17	2, 5, 6, 7, 9, 11, 13, 17	8, 15, 16	
	Class 19 (0.6)	2, 4, 5, 6, 7, 8, 9, 11, 12, 13, 16, 17	2, 4, 5, 6, 7, 9, 11, 13, 17	8, 16	
Class 27	Class 23 (0.692307)	2, 5, 6, 7, 8, 9, 10, 11, 13, 16, 17, 18	2, 5, 6, 7, 8, 9, 13, 16, 17	10, 11, 18	10, 11, 18

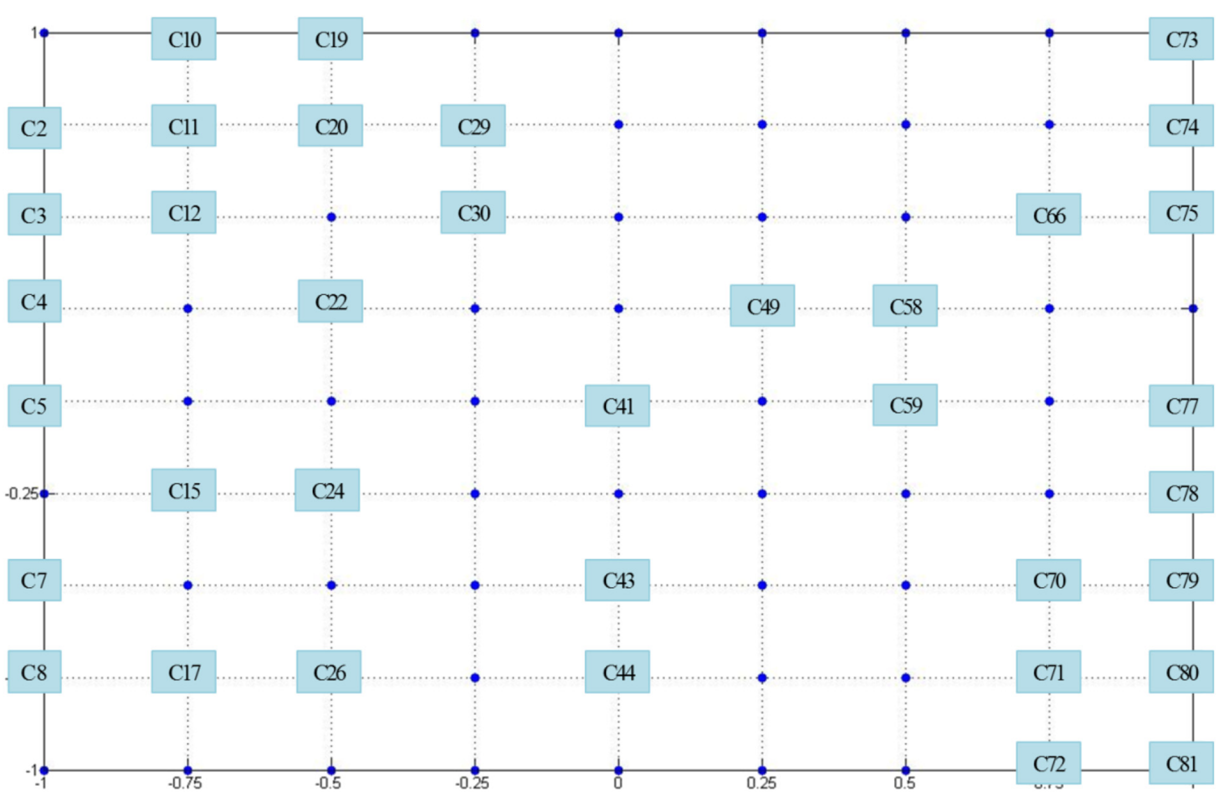

Fig. 6. GTM-based firm-business map related to OLED technology and vacuum cells.

Internet (Al-Fuqaha et al., 2015). In the meantime, display using OLED technology will play an important role in the IoT systems, like networked vehicles, intelligent traffic systems, and smart home or cities, from the viewpoint of visual recognition (Rose et al., 2015). Furthermore, OLED can be exploited for infants, as well as astronomy, such as wearable band, pulse oximeter for infants, and medical monitoring devices, because OLED technology makes the size of screen or display smaller with advanced quality.

Table 7
Promising R&BD areas and relevant product lines based on vacuum cell in the GTM-based map.

Vacuum cell number	Keywords	Promising R&BD areas	Related product line (design map of KIPO*)
Cell 80	Brake, hinges, engines, filtering, diagnostic, preservatives, smartphones	Head-up display Navigation Diagnosis system Preservatives Smartphones	Diagnosis devices for medical Navigation
Cell 81	Infants, internet, astronomy	Internet of Things (IoT) Wearable or monitoring device for infant Display for astronomy	GUI of devices for telecommunication
Cell 2	Entertainment, amusement, music, information, program, movie, promotional, microphones, semiconductor	Digital contents e-Publication	–
Cell 8	Navigation, content, automobile, screen, remote, measurement, processors, antennas, capacitors	Navigation and relevant contents	Navigation GUI of devices for telecommunication
Cell 17	Safety, warning, surveillance, gauges, thermometers, ceramic, circuit	Surveillance and safety system Thermometer	CCTV Camera Black box Medical measurement devices

* means that there is no product line related to cell 2, and it can be substituted "none".

In the case of vacuum cell 2, the business keywords were primarily associated with entertainment and education, such as ‘music’, ‘programs’, ‘journals’, ‘promotional’, ‘entertainment’ and ‘education’. This means that the OLED display is able to enter this area in the form of e-publication or contents for entertainment like music, movie, or magazines, which may be regarded as subsidiary areas, after implementing the display to screens or tablets.

Cell 8 included keywords such as ‘content’, ‘navigation’, ‘processors’, ‘capacitors’, and ‘antennas’ that are related to components and contents implemented by applying the OLED screen or display. Similar to vacuum cell 81, OLED has provided a high quality of navigation in automobiles, due to the wide viewing angle; but there are attempts to make flexible displays using OLED technology, which makes it possible to design elastic displays. Moreover, it is useful to adapt anywhere, with being free of location for attaching displays, so OLED technology may do business for navigation or relevant contents.

Vacuum cell 17 is relevant to safety and measurement, like ‘safety’, ‘warning’, ‘surveillance’, ‘gauges’, and ‘thermometers’. Since the display was originally applicable to a wide range of cameras, such as lenses, screens, and viewfinders, OLED technology was especially prevalent in the compact camera. More recently, it has also started to appear in DSLR and high-end cameras, by being acknowledged as tiny high-resolution displays that are magnified to be used as near-eye displays. In addition, the color gamut is wider than the LCD display, which allows more accurate reproduction of the color of the image. Due to totally unlit black pixels, the contrast is high, and makes a more realistic impression of how the images will look. That is why OLED is able to penetrate the surveillance market in the way of being implemented, such as video, camera, and CCTV. Most of all, the display applying OLED technology does not require as much power as equivalently bright screens, because the light itself is emitted from the display (Zeljkovic, 2013), and thus it is appropriate to the video surveillance system that requests monitoring in real-time, and for a long time without pause. In accordance with a surveillance system, it will be a promising business by extending alarm systems based on surveillance data, with the aim of strengthening the safety of society. In the case of ‘thermometer’ or ‘gauges’, they are adapted to the smart watch for measuring the environment when wearing this device in recent times. In this process, the OLED display is always mentioned in company with the thermometer, because it serves as the main component of the smart watch. Above this, OLED display is utilized as a principal part itself of thermometers, and thus the OLED technology can be commercialized to the measurement market, like thermometers and gauges.

4.4. R&D strategy establishment by analyzing design patent

Among the derived R&BD areas at the prior step, design patents related to navigation, devices for medical diagnosis, and watch were collected and analyzed. These product lines were chosen as representative promising R&BD areas, in which OLED technology is applicable.

Designs for navigation were classified according to five features – shape of front, shape of display, location of operation key, location of direction key, and installation type, as shown in Table 8. Most of the navigation designs were horizontally long rectangular-shaped, and had the same shaped display (respectively 72.18% and 83.89%). Also, the majority of navigation designs had no operation key or direction key, in that the functions of navigation, such as searching destination, and monitoring of current position, were usually operated with the touch screen involved in navigation. The distribution of design patents found that navigation was mostly installed on dashboards (52.74%) or within dashboards (25.99%), and some of them were allowed to be portable (18.90%). The combinations of each feature in navigation designs were mostly reported as ‘rectangular shape (horizontally long) – rectangular-shaped display (horizontally long) – no operation key – no direction key – installation on dashboard’ (139 designs of 931 designs). Since this type of navigation is helpful for drivers to operate and monitor while driving, both drivers and designers prefer this composite of navigation designs. Taking all these considerations into account, OLED technology will be applicable to enhance display on navigation units with the aim of higher visibility and sensitivity of the touch screen.

In the case of devices for medical diagnosis, they had four criteria – type of device, shape, display, and operation button (see Table 9). The largest percentage of medical diagnosis devices is diagnostic/test device (46.24%), which is all kinds of testing apparatus, except for special uses. Their overall shape is mostly hexahedral (35.40%) or freestyle (28.54%), and they have a display that is able to show the procedures of diagnosis or results (68.36%). The operation of the medical diagnosis device is mostly conducted by push-button (50%), and then a hybrid approach was used as operation key (34.07%). The results of composition were similar to the proportion of each feature, and ‘diagnostic/test device – hexahedron – no display – no operation button’ was a big part of design patents in medical diagnosis devices (49 designs of 1101 designs). At the device level, the ultrasonic diagnostic device was the next largest proportion, with hybrid-shaped, display and hybrid operation button (44 designs of 1101 designs).

The smart watch was also chosen as a promising R&BD area with applying thermometers, so design patents related to the watch were

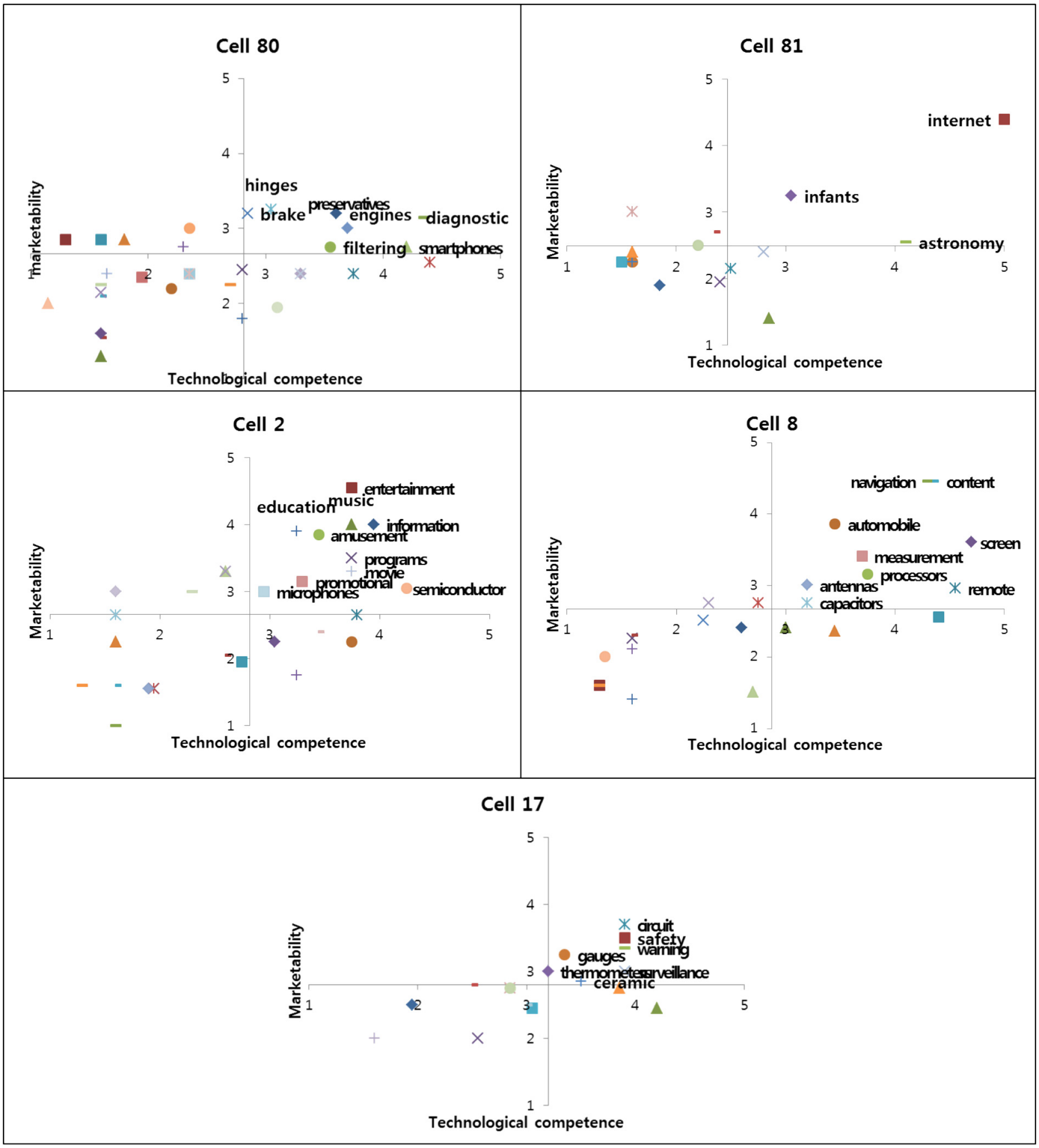

Fig. 7. Keyword portfolio of each vacuum.

analyzed. That is because the appearances of the smart watch depend on the existing watch, in which new functions are added, such as communication, and networking on the whole. The designs for watches are classified on the basis of display type, like analog, digital, and hybrid, and as a result, most watches were affiliated with analog type (91.67%). The shape of case and hour plate was circular (36.09% and 63.60% respectively) for the most part. In addition, the next largest proportion of case shape was left and right convex (27.38%), and rectangular (19.67%). In the case of the hour plate, rectangular was in second place (21.91%). In the type of time display, the hybrid type was the largest percentage of display method (33.89%), which displays time by combining two or more types among Arabic numerals, digital, dot/line/figure, or Roman alphabet. Looking at the combination of features, 'analog type – circle-shaped case – circle-shaped hour plate – hybrid approach for displaying time' was the representative appearance of the watch (1790 designs of 12,438 design patents). The designs related to the watch with the same shape, but displaying time as 'dot/line/figure', were 1188 out of 12,438 design patents. From this point of view, the smart watch applying OLED technology in the form of the above type can be developed, and furthermore advanced to this industry.

5. Implication and discussion

In sum, OLED technology can be embedded into medical devices for diagnosis and measurement, navigation, CCTV, camera, black box, watches, and GUI for telecommunication devices, which were determined by network analysis, GTM, keyword portfolio, and analysis for a design patent. In other words, they are regarded as emerging R&BD areas where is possible to implement product and enter the market by exploiting OLED technology. The organic materials composing OLEDs

Table 8
R&BD strategy for navigations.

(1) Structure of product and proportion of each feature

D1: Form of front	D2: Display form	D3: Location of operation key	D4: Location of direction key	D5: Type of navigation
Modified rectangular (10.96%)	Rectangular (horizontal-long) (83.89%)	Bottom (2.77%)	Circular button (7.84%)	In dashboard (25.99%)
		Circular button (0.001%)	Crosshairs button (0.86%)	
Rectangular (horizontal-long) (72.18%)	Rectangular (vertical-long) (6.76%)	Left (9.13%)	Shuttle key (1.40%)	On dashboard (52.74%)
		Left and right (11.82%)	Square-shaped button (3.87%)	
Rectangular (vertical-long) (10.63%)	Square (4.30%)	No operation key (25.78%)	Unidentified (2.47%)	Portable (18.90%)
		Right (11.82%)	Up/down (4.73%)	
Square (3.44%)	No display (4.40%)	Top (3.65%)	Wheel or jog (7.84%)	Etc. (0.86%)
		Top and bottom (0.54%)	Etc. (5.27%)	
Etc. (2.79%)	Etc. (0.64%)	Etc. (9.45%)	No direction key (65.74%)	No type (1.50%)

(2) Top 5 combination of each feature

Rank	D1: Form of front	D2: Display form	D3: Location of operation key	D4: Location of direction key	D5: Type of navigation	Freq.
1	Rectangular (horizontal-long)	Rectangular (horizontal-long)	No operation key	No direction key	On dashboard	139
2	Rectangular (horizontal-long)	Rectangular (horizontal-long)	Bottom	No direction key	In dashboard	74
3	Rectangular (horizontal-long)	Rectangular (horizontal-long)	Bottom	No direction key	On dashboard	51
4	Rectangular (horizontal-long)	Rectangular (horizontal-long)	Left and right	No direction key	On dashboard	40
5	Rectangular (horizontal-long)	Rectangular (horizontal-long)	Left	No direction key	On dashboard	36

were able to apply to displays, as well as preservatives themselves, and screens embedding OLED technology are able to enter the market associated with medical, image and video processing, navigation, and a variety of automobile applications. Beyond this, it is possible to enter new services, such as content for navigation designs or smart watches, entertainment, or education through implementing OLED technology and convergence with other technology. While OLED was just considered in the way of screen or display, this approach discovered hidden and promising R&BD areas by analyzing trademark and design patents. They are not totally new areas but emergent areas according to recent advances in technological progress, and it may serve as a secondary pointer for decision making when entering a new market or dominating existing ones. OLED technology is possible to give body to three types of next-generation display – rollable, transparent, and 3D display under current technological capability. If they are integrated with each combination of features shown in Tables 8, 9, and 10, new types of product can be developed. For instance, the first combination with OLED technology may implement the navigation with horizontally long-shaped and rollable or foldable display. In particular, the results shown in Tables 8, 9, and 10 may be regarded as development patterns for the product until now, so thus similar technology can be applied to these patterns or substitute and create new business.

This paper aims at identifying technical emergence as well as market emergence. In the case of technical emergence in OLED technology, we collected patents over the past five years. It means that inventions registered as patents are regarded as emerging technology which is recognized for novelty and creativity compared previous technology such as flexible display, organic materials. Using the latest patent documents, the recent and newest trends of emerging OLED technology were identified in the form of topics by the LDA. In addition, based upon this, market emergence was investigated by the GTM and keyword analysis. The market emergence in this paper can be defined as emerging R&BD areas – devices for medical diagnosis, e-publication service or contents for entertainment, where are possible to enter and occupy rapidly by exploiting the latest OLED technology represented as the form of the topic. In the field, this hybrid approach is useful to explore emerging R&BD areas with consideration of both technical emergence and market emergence, but it is applicable by splitting this approach partially. It means practitioners can focus on identifying emerging technology with indicators for evaluating the degree of

Table 9
R&BD strategy for medical diagnosis devices.

(1) Structure of product and proportion of each feature

D1: Device	D2: Shape	D3: Display	D4: Operation key
Cardiac diagnostic device (0.01%)	Cylindrical (2.43%)	Display exist (68.36%)	Dial (3.98%)
Diagnostic/test device (46.24%)	Cylindrical (long type) (2.43%)		Hybrid (34.07%)
Endoscope (5.54%)		No display (31.20%)	Push button (50%)
Immunity diagnostic device (2.43%)	Hexahedron (35.40%)		Sliding type (1.77%)
Pregnancy diagnostic device (0.01%)			Switch (3.54%)
Ultrasonic diagnostic device (41.15%)	Freestyle (28.54%)	Etc. (0.44%)	Etc. (6.63%)
Etc. (3.10%)	Etc. (1.11%)		

(2) Top 5 combination of each feature

Rank	D1: Device	D2: Shape	D3: Display	D4: Operation button	Freq.
1	Diagnostic/test device	Hexahedron	No display	No button	49
2	Diagnostic/test device	Freestyle	No display	No button	47
3	Ultrasonic diagnostic device	Hybrid	Display exist	Hybrid	44
4	Diagnostic/test device	Freestyle	No display	No button	32
5	Ultrasonic diagnostic device	Freestyle	No display	Push button	30

Table 10
R&BD strategy for watches.

(1) Structure of product and proportion of each feature

D1: Functions	D2: Shape of case	D3: Shape of hour plate	D4: Type of displaying time
Analog (91.67%)	Circle (36.09%)	Circle (63.60%)	Arabic numerals (11.75%)
	Ellipse (1.77%)	Ellipse (0.19%)	Digital (3.26%)
Digital (4.16%)	Etc. (6.29%)	Etc. (2.90%)	Dot/line/figure (31.08%)
	Left and right convex (27.38%)	Left and right convex (3.92%)	
Hybrid (3.87%)	Polygon (except of rectangular) (4.11%)	Polygon (except of rectangular) (2.40%)	Hybrid (33.89%)
	Rectangular (19.67%)	Rectangular (21.91%)	
	Spherical (0.58%)	Spherical (0.15%)	Etc. (0.91%)
Etc. (0.30%)	Top and bottom convex (1.19%)	Top and bottom convex (0.12%)	No text (14.19%)
	Totally convex (2.92%)	Totally convex (0.19%)	Roman alphabet (4.91%)

(2) Top 5 combination of each feature

Rank	D1: Device	D2: Shape	D3: Display	D4: Operation button	Freq.
1	Analog	Circle	Circle	Hybrid	1790
2	Analog	Circle	Circle	Dot/line/figure	1188
3	Analog	Left and right convex	Circle	Hybrid	977
4	Analog	Left and right convex	Circle	Dot/line/figure	781
5	Analog	Rectangular	Rectangular	Dot/line/figure	700

emergence restrictively, and then exploring the potential and emerging R&BD areas. That is to say, the proposed approach is very flexible in accordance with usage and focus of users.

Furthermore, the proposed approach can be applied to other fields such as artificial intelligence and deep learning which have received a lot of attentions from almost all organizations. Since artificial intelligence and deep learning are regarded as emerging and disruptive technology recently, the technological ability has been advanced more and more. Even though the technological ability of a technology is improved, it may be not useful if the market or the industry that utilizes the emerging technology is not defined earlier. In other words, unsettled market or industry may have a negative effect to growth in artificial intelligence market. However, this study is able to solve this problem by predicting emerging R&BD areas based on relevant IPRs using machine learning, deep learning, and artificial intelligence.

6. Conclusion

This study suggested a novel approach for identifying emerging R&BD areas through comprehensively analyzing IPRs, including patent, trademark, and design patent. After finding all kinds of possible and potential R&BD areas based on the relationship between patent and trademark with LDA and keyword analysis, the promising R&BD areas were derived by vacuums in GTM-based trademark map, and additional assessment relying on keywords. Finally, keyword analysis for each vacuum offered possible product or service, and then design patents assisted to check the technical feasibility for the appearances of products in the form of a morphological matrix.

As referred to earlier, this paper attempted to detect emerging and promising R&BD areas by the hybrid approach integrating topic analysis (LDA), network analysis, and GTM. The proposed approach was composed of three techniques using bibliographic data to achieve a goal for each module. The hybrid approach made better use of bibliometrics as well as explored emerging R&BD areas structurally and intuitively through data visualization – especially network analysis and GTM. In addition, it took account of heterogeneous features and bibliographic data of IPRs. For example, patent documents involve title, abstract, owner, description while trademark possesses limited information such as owner, classification and targeting business. Patent documents contain a large amount of textual data related to the detailed description for an invention and what to insist and protect legally which is called claims. It represents technological attributes and functions while trademark represents the scope of business. Techniques for bibliometric analysis were chosen diversely to detect emerging R&BD areas, reflecting aforementioned different properties of IPRs.

Above all things, this study is meaningful, in that all kinds of IPRs were utilized to integrally discover new R&BD areas, giving sufficient consideration to the unique features of each IPR. The information about technological specification and embodiment involved in patent documents served to extract functions that can be implemented in goods and services. Trademarks contributed to determining the market status on the whole, and the scope of each firm that belongs to the industry. Then, just for promising R&BD areas, design patents were able to provide possible appearances and the chances for implementation from both technology and market perspectives. Moreover, the suggested approach was to explore new but hidden opportunities in the market, by relying on technological functions. The R&BD opportunities were also proposed at the product level in detail, in other words, a blueprint was suggested in the form of concrete appearance. Design patents served to check the feasibility of part composition for products, by asking whether the product can be technologically developed and implemented. From this perspective, the proposed approach will be useful to apply to high-tech industry and technology-driven industry.

From the viewpoint of practitioners belonging to the fields, this research helps all stakeholders included in departments of technology planning, as well as manufacturing, marketing, and advertising on the whole. When planning and developing new technology, the stakeholders in this phase may utilize this approach for the purpose of avoiding infringements of IPRs after terminating research and development, which leads to increase in the rate of success in technology commercialization. In the case of manufacturers, our study is able to support the minimization of the changeover of design for product and service, because the feasibility is already checked by IPR analysis – in particular, design patent analysis on the authority of emerging R&BD areas, depending on patent and trademark analysis.

Nevertheless, our study still has several limitations. First, the analysis for design patent was restrictively conducted, while patent and trademark were analyzed as much as possible. Because the design patent has a relatively small amount of information that is immediately applied to analysis with specific methodology compared with the patent and trademark, the application of design patent has had to be restricted. The patent has a large amount of bibliographic data, such as technical specification and application, and the trademark also has a range of application. In contrast, the design patent just includes the appearances of products or graphic user interface that is insufficient to analyze and examine R&BD areas. Secondly, subjective opinions by the researcher are still involved in the process of calculating and assessing the degree of promise, based on weights for keywords constituting cells in the GTM-based trademark map. Our study endeavored to minimize subjective opinions and effort or time by a human with a systematic approach, through calculating degree relying on objective data and information; but experts' opinions was additionally required, with the goal of more accurate forecasting and reflecting practical opinion.

In order to overcome these limitations, deep learning, which is proficient at automatically analyzing a huge amount of images and videos while minimizing human effort, can be applied to investigate design patents. Since deep learning is able to extract features by itself, and to learn from accumulated data, the future study for analyzing design patents based on deep learning may create more meaningful insights about promising R&BD areas. Moreover, to minimize the supervision of experts, Word-to-Vector (Word2vec), which is regarded as one of the machine learning techniques, is applicable to assess grading for each business keyword. It is effective to examine the importance or value of keywords, by analyzing the co-occurrences in text data, and then reflecting the context of documents. Since keywords can be evaluated though considering relations with other words and context, future study depending on Word2Vec will be able to procure the results of keyword-based analysis.

Acknowledgement

This work was supported by National Research Foundation of Korea (NRF-2017S1A5A2A01023966).

References

Abraham, B.P., Moitra, S.D., 2001. Innovation assessment through patent analysis. Technovation 21 (4), 245–252.
Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., Ayyash, M., 2015. Internet of things: a survey on enabling technologies, protocols, and applications. IEEE Commun. Surv. Tutorials 17 (4), 2347–2376.
Alikhan, S., 2002. The Role of Copyright in the Cultural and Economic Development of Developing Countries.
Arun, R., Suresh, V., Veni Madhavan, C.E., Narasimaha Murthy, M.N., 2010. On finding the natural number of topics with latent dirichlet allocation: Some observation. In: Zaki, Mohammed J., Yu, Jeffrey Xu, Ravindran, Balaraman, Pudi, Vikram (Eds.), Advances in Knowledge Discovery and Data Mining. Springer, Berlin Heidelberg, pp. 391–402.
Bennett, D., 2002. The Power of Patents and their Strategic Use (the 2002 Law Journal). Business North Carolina.
Benson, C., Magee, C., 2015. Quantitative determination of technological improvement from patent data. PLoS ONE 10 (4), 1–23 e0121635.
Bhaduri, N., Mathew, M., 2003. Intellectual property management practices. In: Paper Presented at the Management of Engineering and Technology, 2003. PICMET'03. Technology Management for Reshaping the World. Portland International Conference on.
Bishop, C.M., Svensén, M., Williams, C.K., 1998a. Developments of the generative topographic mapping. Neurocomputing 21 (1), 203–224.
Bishop, C.M., Svensén, M., Williams, C.K., 1998b. GTM: the generative topographic mapping. Neural Comput. 10 (1), 215–234.
Blei, D., Ng, A., Jordan, M., 2003. Latent Dirichlet allocation. J. Mach. Learn. Res. 3 (Jan), 993–1022.
Broadus, R.N., 1987. Toward a definition of “bibliometrics”. Scientometrics 12 (5–6), 373–379.
Butts, C.T., 2008. Social network analysis: a methodological introduction. Asian J. Soc. Psychol. 11 (1), 13–41.
Cao, J., Xia, T., Li, J., Zhang, Y., Tange, S., 2009. A density-based method for adaptive LDA model selection. Neurocomputing 72 (7–9), 1775–1781.
Chang, J., Boyd-Graber, J., Wang, C., Gerrish, S., Blei, D.M., 2009. Reading tea leaves: how humans interpret topic models. Adv. Neural Inf. Proces. Syst. 1–9.
Chang, S.-B., Lai, K.-K., Chang, S.-M., 2009. Exploring technology diffusion and classification of business methods: using the patent citation network. Technol. Forecast. Soc. Chang. 76 (1), 107–117.
Chen, C.-J., 2009. Technology commercialization, incubator and venture capital, and new venture performance. J. Bus. Res. 62 (1), 93–103.
Chen, R., 2009. Design patent map visualization display. Expert Syst. Appl. 36 (10), 12362–12374.
Chen, A., Chen, R., 2007. Design patent map: an innovative measure for corporative design strategies. Eng. Manag. J. 19 (3), 14–29.
Chen, R., Chen, C.-C., Dai, T.-T., 2013. Community knowledge-based design patent map system with efficient dissimilarity visualization engine: design, performance, and applications. Int. J. Comput. Inf. Syst. Ind. Manag. Appl. 5, 326–336.
Cho, J., Lee, J., 2013. Development of a new technology product evaluation model for assessing commercialization opportunities using Delphi method and fuzzy AHP approach. Expert Syst. Appl. 40 (13), 5314–5330.
Conceição, P., Hamill, D., Pinheiro, P., 2002. Innovative science and technology commercialization strategies at 3M: a case study. J. Eng. Technol. Manag. 19 (1), 25–38.
Coronado, R.A., Wurtzel, W.A., Simon, C.B., Riddel, D.L., George, S.Z., 2011. Content and bibliometric analysis of articles published in the journal of orthopaedic & sports physical therapy. J. Orthop. Sports Phys. Ther. 41 (12), 920–931.
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.
Deveaud, R., SanJuan, E., Bellot, P., 2014. Accurate and effective latent concept modeling for ad hoc information retrieval. Document numérique 17 (1), 61–84.
Ding, Y., 2011. Scientific collaboration and endorsement: network analysis of coauthorship and citation networks. J. Informet. 5 (1), 187–203.
Duke, C.R., 1995. Organizational conflicts affecting technology commercialization from nonprofit laboratories. J. Prod. Brand Manag. 4 (5), 5–13.
Érdi, P., Makovi, K., Somogyvári, Z., Strandburg, K., Tobochnik, J., Volf, P., Zalányi, L., 2013. Prediction of emerging technologies based on analysis of the US patent citation network. Scientometrics 95 (1), 225–242.
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.
Ferrara, A., Salini, S., 2012. Ten challenges in modeling bibliographic data for bibliometric analysis. Scientometrics 93, 765–785.
Furukawa, T., Mori, K., Arino, K., Hayashi, K., Shirakawa, N., 2015. Identifying the evolutionary process of emerging technologies: a chronological network analysis of World Wide Web conference sessions. Technol. Forecast. Soc. Chang. 91, 280–294.
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.
Gould, D.M., Gruben, W.C., 1996. The role of intellectual property rights in economic growth. J. Dev. Econ. 48 (2), 323–350.
Griffiths, T.L., Steyvers, M., 2004. Finding scientific topics. Proc. Natl. Acad. Sci. 101, 5228–5235.
Gupta, S., Pandey, S., Shukla, K.K., 2015. Comparison analysis of link prediction algorithms in social network. Int. J. Comput. Appl. 111 (16), 27–29.
Gwak, J.H., Sohn, S.Y., 2017. Identifying the trends in wound-healing patents for successful investment strategies. PLoS ONE 12 (3), 1–19 e0174203.
Hall, D., Jurafsky, D., Manning, C.D., 2008. Studying the history of ideas using topic models. In: In Proceedings of the conference on empirical methods in natural language processing. Association for Computational Linguistics, pp. 363–371.
Hood, W.W., Wilson, C.S., 2001. The literature of bibliometrics, scientometrics, and informetrics. Scientometrics 52 (2), 291–314.
Hsu, D.W., Shen, Y.-C., Yuan, B.J., Chou, C.J., 2015. Toward successful commercialization of university technology: performance drivers of university technology transfer in Taiwan. Technol. Forecast. Soc. Chang. 92, 25–39.
Jeong, S., Yoon, B., 2013. A systemic approach to exploring an essential patent linking standard and patent maps: application of generative topographic mapping (GTM). Eng. Manag. J. 25 (1), 48–57.
Jeong, Y., Lee, K., Yoon, B., Phaal, R., 2015. Development of a patent roadmap through the Generative Topographic Mapping and Bass diffusion model. J. Eng. Technol. Manag. 38, 53–70.
Joung, J., Kim, K., 2017. Monitoring emerging technologies for technology planning using technical keyword based analysis from patent data. Technol. Forecast. Soc. Chang. 114, 281–292.
Kajikawa, Y., Takeda, Y., 2009. Citation network analysis of organic LEDs. Technol. Forecast. Soc. Chang. 76 (8), 1115–1123.
Kanwar, S., Evenson, R., 2003. Does intellectual property protection spur technological change? Oxf. Econ. Pap. 55 (2), 235–264.
Kassicieh, S.K., Walsh, S.T., Cummings, J.C., McWhorter, P.J., Romig, A.D., Williams, W.D., 2002. Factors differentiating the commercialization of disruptive and sustaining technologies. IEEE Trans. Eng. Manag. 49 (4), 375–387.
Kim, Y.G., Suh, J.H., Park, S.C., 2008. Visualization of patent analysis for emerging technology. Expert Syst. Appl. 34 (3), 1804–1812.
Kim, M., Park, Y., Yoon, J., 2016. Generating patent development maps for technology monitoring using semantic patent-topic analysis. Comput. Ind. Eng. 98, 289–299.
Kumar, V., Jain, P., 2003. Commercialization of new technologies in India: an empirical study of perceptions of technology institutions. Technovation 23 (2), 113–120.
Lee, S., Yoon, B., Lee, C., Park, J., 2009. Business planning based on technological capabilities: patent analysis for technology-driven roadmapping. Technol. Forecast. Soc. Chang. 76 (6), 769–786.
Lee, C., Kwon, O., Kim, M., Kwon, D., 2018. Early identification of emerging technologies: A machine learning approach using multiple patent indicators. Technol. Forecast. Soc. Chang. 127, 291–303.
Li, F., Li, C., Tian, Y., 2009. Applying association rule analysis in bibliometric analysis – a case study in data mining. In: Proceedings of the 2nd Symposium International Computer Science and Computational Technology, Huangshan, China, pp. 431–434.
Liben-Nowell, D., Kleinberg, J., 2007. The link-prediction problem for social networks. J. Assoc. Inf. Sci. Technol. 58 (7), 1019–1031.
Lin, Y., Wang, Y., Kung, L., 2015. Influences of cross-functional collaboration and knowledge creation on technology commercialization: evidence from high-tech industries. Ind. Mark. Manag. 49, 128–138.

Liu, H., Hu, Z., Haddadi, H., Tian, H., 2013. Hidden link prediction based on node centrality and weak ties. Europhys. Lett. 101 (1), 18004.
Lü, L., Pan, L., Zhou, T., Zhang, Y.-C., Stanley, H.E., 2015. Toward link predictability of complex networks. Proc. Natl. Acad. Sci. 112 (8), 2325–2330.
Mavridis, T., Symeonidis, A.L., 2012. Identifying Webpage Semantics for Search Engine Optimization. WEBIST 2012, 272–275.
Monemi, A., Rost, K., 2016. Identification and monitoring of possible disruptive technologies by patent-development paths and topic modeling. Technol. Forecast. Soc. Chang. 104, 16–29.
Nam, Y., Barnett, G.A., 2011. Globalization of technology: network analysis of global patents and trademarks. Technol. Forecast. Soc. Chang. 78 (8), 1471–1485.
Norton, M.J., 2001. Introductory Concepts in Information Science, New Jersey.
Park, Y., Yoon, B., Lee, S., 2005. The idiosyncrasy and dynamism of technological innovation across industries: patent citation analysis. Technol. Soc. 27 (4), 471–485.
Pépin, L., Kuntz, P., Blanchard, J., Guillet, F., Suignard, P., 2017. Visual analytics for exploring topic long-term evolution and detecting weak signals in company targeted tweets. Comput. Ind. Eng. 112, 450–458.
Reitzig, M., 2004. Improving patent valuations for management purposes—validating new indicators by analyzing application rationales. Res. Policy 33 (6–7), 939–957. https://doi.org/10.1016/j.respol.2004.02.004.
Rose, K., Eldridge, S., Chapin, L., 2015. The internet of things: an overview. The Internet Society (ISOC) 1–50.
Sandner, P.G., Block, J., 2011. The market value of R&D, patents, and trademarks. Res. Policy 40 (7), 969–985.
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.
Shibata, N., Kajikawa, Y., Sakata, I., 2010. Extracting the commercialization gap between science and technology—case study of a solar cell. Technol. Forecast. Soc. Chang. 77 (7), 1147–1155.
Slater, S.F., Mohr, J.J., 2006. Successful development and commercialization of technological innovation: insights based on strategy type. J. Prod. Innov. Manag. 23 (1), 26–33.
Sohn, S.Y., Moon, T.H., 2003. Structural equation model for predicting technology commercialization success index (TCSI). Technol. Forecast. Soc. Chang. 70 (9), 885–899.
Sohn, S.Y., Moon, T.H., 2004. Decision tree based on data envelopment analysis for effective technology commercialization. Expert Syst. Appl. 26 (2), 279–284.
Son, C., Suh, Y., Jeon, J., Park, Y., 2012. Development of a GTM-based patent map for identifying patent vacuums. Expert Syst. Appl. 39 (3), 2489–2500.
Svensson, R., 2007. Commercialization of patents and external financing during the R&D phase. Res. Policy 36 (7), 1052–1069.
Venugopalan, S., Rai, V., 2015. Topic based classification and pattern identification in patents. Technol. Forecast. Soc. Chang. 94, 236–250.
White, H.D., McCain, K.W., 1989. Bibliometrics. Annu. Rev. Inf. Sci. Technol. 24, 119–186.
Wu, M., Chang, K., Zhou, W., Hao, J., Yuan, C., Chang, K., 2015. Patent deployment strategies and patent value in LED industry. PLoS ONE 10 (6), 1–10 e0129911.
Yan, B., Luo, J., 2017. Filtering patent maps for visualization of diversification paths of inventors and organizations. J. Assoc. Inf. Sci. Technol. 68 (6), 1551–1563.
Yoon, B., Park, Y., 2004. A text-mining-based patent network: analytical tool for high-technology trend. J. High Technol. Managem. Res. 15 (1), 37–50.
Zahra, S.A., Nielsen, A.P., 2002. Sources of capabilities, integration and technology commercialization. Strateg. Manag. J. 23 (5), 377–398.
Zeljkovic, V., 2013. Video Surveillance Techniques and Technologies. IGI Global.

Yujin Jeong is currently a Ph.D. candidate in Department of Industrial & Systems Engineering of the Dongguk University. Her theme of study has involved SMEs, quality management, open innovation technology forecasting, technology intelligence, data mining and patent analysis.

Inchae Park is currently a postdoctoral associate in Massachusetts Institute of Technology (MIT), MA. He has a PhD with the Department of Industrial & Systems Engineering, Dongguk University, Seoul Rep. of Korea. His research interests include technology forecasting, technology intelligence, data mining, and patent analysis.

Byungun Yoon is a Professor in Department of Industrial & Systems Engineering of the Dongguk University. His theme of study has involved patent analysis, new technology development methodology and visualization algorithms. His current interest is in enhancing technology road mapping, R&D quality and product designing with data mining techniques.