Technology Analysis & Strategic Management
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ctas20

A semantic analysis approach for identifying patent infringement based on a product-patent map

Inchae Park[a] & Byungun Yoon[a]
[a] Department of Industrial & Systems Engineering, Dongguk University, Seoul, Republic of Korea
Published online: 22 Apr 2014.

Technology Analysis & Strategic Management, 2014
Vol. 26, No. 8, 855–874, http://dx.doi.org/10.1080/09537325.2014.909926

A semantic analysis approach for identifying patent infringement based on a product–patent map

Inchae Park and Byungun Yoon*

Department of Industrial & Systems Engineering, Dongguk University, Seoul, Republic of Korea

Identifying patent infringement beforehand is highly important for reducing the risk of damages. However, as the approach to patents is changing from the technological and legal approaches of the past to the business approach, strategic patent infringement litigation and securement of patents can be important tools for companies. From various viewpoints, companies can consider an aggressive patent infringement litigation as well as evasion strategy of patent infringement litigation to keep competitors in check. Previous research on patent infringement identification only considered the possibility of infringement between patents. However, infringement cases between patent and product are prevalent. Thus, this research aims to suggest an automated method for identifying patent infringement between the patent and product in business. To this end, this paper proposes the subject–action–object (SAO) semantic technological similarity-based product–patent method for generating infringement maps. Several indices and subgrouping methods are suggested to interpret the map. As an exemplary case, data on technology and products related to the light emitting diode (LED) lamp are collected and exploited.

Keywords: product–patent map; patent infringement; semantic analysis; subject–action–object (SAO) structure; social network analysis

Introduction

Protecting intellectual property rights (IPR) is a highly important issue and an increasing priority for governments. While intellectual property (IP) plays in promoting innovation growth, it causes concern with the adverse effects that counterfeiting and piracy are bringing about on economies and society as a whole (OECD 2008). The increase of copying technology and insignia results from the globalisation trend, highly dispersed value creation chains and distribution of goods in international markets (Berger, Blind and Cuntz 2012). Various types of IPs such as trademarks, copyrights, patents and design rights are covered in previous IPR infringement researches (Schmiele 2013; OECD 2008; Berger, Blind and Cuntz 2012). Patents among them are generally available for any inventions, including both products and processes, in all fields of technology.

*Corresponding author. Email: postman3@dongguk.edu

Patent infringement litigations have been increasing between companies. This phenomenon occurs because companies aim for a more competitive advantage in the market by protecting the outcomes of their research and development. To make matters worse, as patent trolls or nonpracticing entity (NPE), which are firms whose business models focus on enforcing patents against infringers in order to receive damages or settlements payments (Golden 2007; Lemley and Shapiro 2007; Reitzig, Henkel and Heath 2007; Pohlmann and Opitz 2013), have intensively emerged, patent infringement litigations have rapidly increased (Soo et al. 2006).

As the results of the litigation case between Kodak and Polaroid in 1976, Kodak paid almost US$1 billion to Polaroid in 1991 (Hall and Ziedonis 2001) and withdrew from the instant-camera market with huge penalties, such as closure of their factory, disposal of inventories, consumer injury compensation and legal fees. Therefore, a company must identify patent infringement beforehand to reduce the risk of damages from litigation, as getting involved in patent infringement litigation causes significant losses in terms of time and costs.

Another important example of multinational patent infringement litigation between Apple and Samsung electronics, which began in 2011, is not only much bigger in scale than the previous case but also involves design patents, trade dress and trademark rights. Thus, the needs of companies to claim intellectual property rights are expanding in respect to both scale and scope.

As the approach to patents changes from the technological and legal approaches of the past to the business approach, strategic patent infringement litigation and securement of patents are important strategies for companies. In other words, patent infringement litigation can be newly recognised as a tool for holding business competitors in check.

The safest way to identify patent infringement is analysis by human experts (Durham 2004). However, human patent experts cannot detect all possible infringement cases (Wallerstein, Mogee and Schoen 1993; Majewski and Williamson 2004) because the number of patent applications has been increasing (Arundel 2001). The complexity of technologies and trend of technology convergence, which require more technologies for individual products, have also increased (Carree, Klomp and Thurik 2000).

When patent infringement litigation occurs, a patent expert determines patent infringement based on the text of the patent document and the technological right of the patent is protected by the contents found in claims. Thus, measuring technological similarities based on textual information in patent documents is suited for suggesting an automated methodology.

Keyword-based text mining to measure technological similarity has been widely adopted for research related to patent analysis (Yoon 2008; Kim, Suh and Park 2008; Tsourikov, Batchilo and Sovpel 2000). However, keyword-based text mining utilises only the frequency of predefined keywords, so it cannot reflect specific, technological key findings or structural relationships among components (Lee, Song and Park 2013). To overcome the approach, this paper suggests a way to identify possible patent infringement using subject–action–object (SAO) based on semantic technological similarity (Park, Yoon and Kim 2012).

Research on the evaluation of patent infringement risk based on the SAO approach (Bergmann et al. 2008) exploits the method of measuring semantic technological similarity by using expert-based analysis. To overcome the limitations of this measurement method, Park, Yoon and Kim (2012) suggested an automated, SAO-based semantic method of measuring technological similarities to identify patent infringement using WordNet, which is a hierarchical thesaurus of English (Miller 1995). Although these two studies identified patent infringement by generating a patent map, they considered only the possibility of patent infringement between patents. Thus, methodology for identifying patent infringement between the patent and product is still necessary, as patent infringement litigation can occur between a patent and product in a real business situation.

This research aims to suggest an automated method for identifying patent infringement between a patent and product. To this end, we proposed the SAO semantic technological similarity-based product-patent infringement map generating method. Several indices and subgrouping methods are suggested for interpreting the map. As an exemplary case, data on technology and products related to the light-emitting diode (LED) lamp are collected and analysed.

This paper is organised as follows. First, as an introductory statement, the general theoretical background is presented. Second, the overall process of developing a product–patent infringement map and determining possible patent infringement are described. Third, an exemplary case is used to exhibit the process of analysis and to assure the utility of the method. Finally, conclusions are presented.

Background

Subject–action–object-based text mining

Extracting SAO structures is a text-mining technique in that it makes it possible to extract information from a huge amount of unstructured textual documents. An SAO structure consists of three elements, the syntactically ordered structure of the subject, action and object (Franzosi 1994). This structure provides the relationship between elements that appear in a patent document. SAO structures are fundamentally related to the concept of function, which is defined as 'the action changing a feature of any object' (Savransky 2000).

Recent studies related to patent analysis for R&D management, such as management of human resources (Moehrle et al. 2005), evaluation of patent risk (Bergmann et al. 2007, 2008), product forecasting (Gerken, Moehrle and Walter 2010), identification of evolving technological trends, patent infringement between patents, new technological opportunities and technological competition trends (Park, Yoon and Kim 2012; Yoon and Kim, 2011, 2012; Yoon, Park and Kim 2013), have started to exploit SAO-based text mining instead of keyword-based text mining.

Measuring semantic technological similarity using WordNet

WordNet is a hierarchically structured lexical database for the English language (Miller et al. 1990). In other words, WordNet can be used as a lexical ontology in the sense of computer science. It contains sets of synonyms, called synsets, for nouns, verbs, adjectives and adverbs. WordNet superficially resembles a thesaurus in that it groups words together based on their meaning. However, WordNet links not just word forms but also specific senses of words and labels the semantic relationships among them. Words that are found in close proximity to one another in WordNet are semantically disambiguated and provide some explicit patterns other than similarity of meaning available in a thesaurus (Fellbaum 2005). The latest version of WordNet, WordNet 3.0, contains 155,287 words organised in 117,659 synsets for a total of 206,941 word–sense pairs. Therefore, WordNet is considered to be the most suitable ontology to offer semantic information for tokens of a sentence and is a useful tool for computational linguistics and natural-language processing (Yoon and Kim 2011).

Social network analysis

For a set of actions, the interactive relationship among actors can generally be portrayed as a network (Gelsing 1992). Social network analysis, a quantitative technique derived from graph theory, facilitates the analysis of interactions (edges) between actors (nodes). Actors may be discrete individuals of any kind and a relation is a collection of ties among actors in the group. Usually exhibited in visual form, the structure of relations among actors and the location of individual actors in the network provide rich information on the behavioural, perceptual and attitudinal aspects of individual units and the system as a whole (Knoke and Kuklinski 1982; Marseden and Laumann 1984).

The applicability of network analysis is wide and diverse. Some typical topics include the world political and economic system (Synder and Kick 1979), interindustrial diffusion and adoption of innovations (Leoncini, Maggioni and Montressor 1996; Park and Kim 1999), and human networks in knowledge management (Cross, Borgatt and Parker 2001).

In the context of the network of product–patent infringement, individual patents and products account for nodes, and the relationships among patents and products represent edges in the network. A two-mode network is generated to represent the relationship between different sets of entities in this research. A one-mode network is generated to detect clusters after converting two-mode matrices to square one-mode matrices.

Research methodology

Overall framework

The overall research process consists of several steps shown in Figure 1, the first of which is data collection from patents and product documents. Second, SAO structures are extracted from each document. Third, technological similarity is measured between documents based on the comparison of SAO structures. Fourth, a map of product–patent infringement is developed using social network analysis. In this step, a two-mode network and a one-mode network are generated. Finally, a potential infringement product and patent are identified by interpreting the map

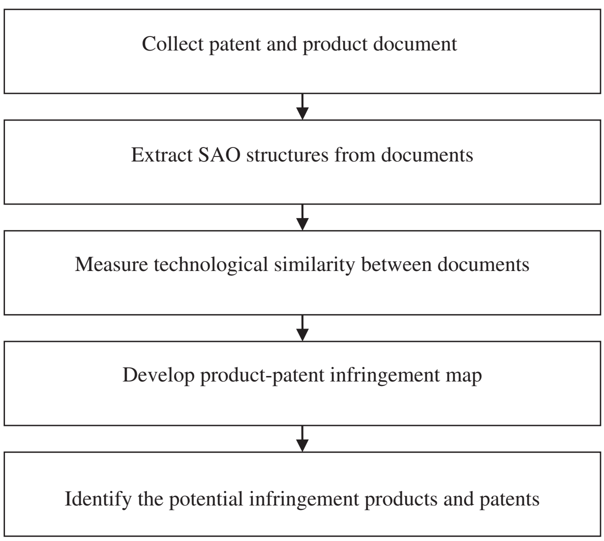

Figure 1. Overall framework of this research.

of product–patent infringement. In this step, both static analysis and dynamic analysis are conducted. In addition to this, indices for identifying patent infringement are suggested. Furthermore, subgrouping for extracting characteristics is conducted.

Data sources

In this research, it is assumed that we are in a specific company and the purpose of this research is to search for potential infringement between patents and products. To extract SAO structures from patents, the patent documents are collected from the US Patent and Trademark Office (USPTO) database.

To extract SAO structures from the product, a document related to the competitor's product such as the product manual, specification or product catalog, needs to be collected. However, it cannot be guaranteed that product documents contain technological content because the purpose of these documents is different from that of a patent. Product manuals offer users instructions for using a product. Specifications consist of numerical values with assembly drawings or tables of criteria. Product catalogs are for sales or marketing purposes and largely consist of product images. The format of product documents is also different for each company and it is difficult to obtain product documents.

To remedy these difficulties, it is assumed that the description section of a patent document is substituted for the product document, as the description section of a patent document contains product features. The description explains the invention and consists of the background, summary, drawing and detailed description of the invention. This description requires enough detail in clear, concise and exact terms (Bryant 1997). It is reliable, as most companies' patents are written by patent attorneys. Since patents related to products are collected by searching product keywords, patents include terms concerning the product, such as 'device' and 'apparatus', and terms concerning the physical structure of the product, such as 'body' and 'frame'. Furthermore, products were identified in previous research by extracting product dimensions from patent documents (Verhaegen et al. 2011). Therefore, we utilised the description section of patent documents to extract products' SAO structures. It is assumed that the group of patents that a competitor owns represents a product since the product can be considered an aggregate of many technologies. Before aggregating a product's patents, it is necessary to define the International Patent Classification (IPC) of the product. Patents that do not include a pre-defined IPC are screened, and the rest of the patents are aggregated.

SAO structure extraction

To extract SAO structures from the documents, this paper used Knowledgist 2.5, a system for processing textual language that has been used in several recent content-based patent analyses (Bergmann et al. 2008; Moehrle et al. 2005). Patent claims are especially considered to represent the inventor's expertise since the claims include intensive knowledge that requires legal protection (Fujii, Iwayama and Kando 2007). The claim is used to extract technological content from the companies that own the technology. In contrast, the description section of the patent is used to extract content related to the product from the sum of competitor's patents, as the product is an assemblage of related technologies.

Measurement of semantic technological similarity

In this step, a semantic patent–product matrix is generated by measuring semantic sentence similarity. Since the similarity between two documents is based on the similarity between SAO structures, sentence similarity between two SAO structures must be measured. The procedure for measuring the semantic similarity of two sentences consisted of (1) tokenising the sentences, (2) stemming words, (3) tagging parts of speech, (4) determining the most likely sense of each word in each sentence and (5) computing the similarity of the sentences based on the similarity of the pairs of words (Simpson and Dao 2005). First, the measure of similarity between two concepts in a concept hierarchy (Resnik 1999) is defined as follows:

$$\mathrm{sim}\,(c_1, c_2) = \frac{2 \times \mathrm{depth}\,(lcs\,(c_1, c_2))}{\mathrm{depth}\,(c_1) + \mathrm{depth}(c_2)}, \tag{1}$$

where lcs is the lowest common subsume of two concepts c_1 and c_2 in the concept hierarchy, and depth is the distance from a concept node c_i to the root of the concept hierarchy. The similarity of two concepts is $0 < \mathrm{sim}(c_1, c_2) \leq 1$, where 1 means that the two concepts are identical. The measure of semantic sentence similarity between two SAO structures, SAO_i and SAO_j, can be formulated by exploiting the matching average (Simpson and Dao 2005).

$$\mathrm{MatAvg}(\mathrm{SAO}_i, \mathrm{SAO}_j) = \frac{2 \times \mathrm{Match}(\mathrm{SAO}_i, \mathrm{SAO}_j)}{|\mathrm{SAO}_i| + |\mathrm{SAO}_j|}, \tag{2}$$

where $|\mathrm{SAO}_i|$ is the number of set concepts in the relevant SAO structure and $\mathrm{Match}(\mathrm{SAO}_i, \mathrm{SAO}_j)$ is the sum of similarity regarding the matching concepts between the sentences. The semantic similarity between two SAO structures is $0 < \mathrm{MatAvg}(\mathrm{SAO}_i, \mathrm{SAO}_j) \leq 1$, where 1 means that the SAO structures are identical. Using the WordNet semantic dictionary (Miller 1995) as a concept hierarchy, the similarity measurement between two words and the matching average for similarity measurement between two sentences, a .NET-based semantic sentence similarity measurement was implemented as a C# library (Simpson and Dao 2005). WordNet has already been used in a number of large-scale knowledge-processing applications, especially in patent analysis (Vasiliki et al. 2006; Yang et al. 2005; Verhaegen et al. 2011; Al-Shboul and Myaeng 2011). However, given that WordNet is a general-purpose language repository, it was extended for patent analysis by patent terminology and terms specific to the areas of technology examined in previous research.

To determine whether or not two SAO structures are identical, a threshold value t was used. Two SAO structures were considered semantically the same if the similarity score was larger than t. For example, when $t = 0.70$ and $\mathrm{MatAvg}(\mathrm{SAO}_i, \mathrm{SAO}_j) = 0.83$, then the two SAO structures can be considered identical:

$$\mathrm{SAO}_{ij} = \begin{cases} 1, & \text{if } \mathrm{MatAvg}(\mathrm{SAO}_i, \mathrm{SAO}_j) \geq t \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Finally, a measure of similarity between two documents, A and B, can be simply defined based on how many SAO structures the two documents share (Yoon and Kim 2011):

$$\mathrm{SIM}(\mathrm{A}, \mathrm{B}) = \frac{2 \times N_{\mathrm{SAO}}(\mathrm{A}, \mathrm{B})}{N_{\mathrm{SAO}}(\mathrm{A}) + N_{\mathrm{SAO}}(\mathrm{B})}, \tag{4}$$

where $N_{\mathrm{SAO}}(\mathrm{A})$ is the number of SAO structures in the patent document A, $N_{\mathrm{SAO}}(\mathrm{B})$ is the number of SAO structures in the product document B, and $N_{\mathrm{SAO}}(\mathrm{A}, \mathrm{B})$ is the number of identical SAO structures that the documents share. The similarity score of two documents is $0 \leq \mathrm{SIM}(\mathrm{A}, \mathrm{B}) \leq 1$, where 1 means that the two documents are identical.

Generation of product–patent infringement map

Two-mode network-infringement map

The two-mode matrix is yielded by measuring the technological similarity between documents, since the indexes of the rows and columns had different sets of entities (Borgatti 2009). Table 1 presents a sample of two-mode data where the rows correspond to patents while the columns correspond to products. The two-mode network-infringement map is developed as a bipartite graph based on the two-mode matrix created by using an appropriate cut-off value. This is useful for figuring out the technological similarity between a patent and product since the map is bipartite.

One-mode network-infringement map

To include the technological similarity between patents in the infringement map, the two-mode matrix was transformed into a one-mode matrix as in Table 2. The technological similarities between competitors' products are zero since the analyst is not interested in the technological similarities between competitor's products. Thus, the infringement map shows technological similarities between patents and products. The map is useful for figuring out similar patent groups or patent portfolios by clustering patents. The patent portfolio is an important tool for companies to manage intellectual property rights.

Table 1. Example of 2-mode data.

	Product 1	Product 2	…	…	Product *m*
Patent 1	0.028	0.029	…	…	0.110
Patent 2	0.023	0.024	…	…	0.113
…	…	…	…	…	…
…	…	…	…	…	…
Patent *m*	0.048	0.050	…	…	0.422

Table 2. Example of one-mode data.

	Patent 1	Patent 2	…	…	Patent *m*	Product 1	…	…	Product *n*
Patent 1	1	0.2727	…	…	0.3636	0.0275	…	…	0.1105
Patent 2	0.2727	1	…	…	0.2580	0.0289	…	…	0.1130
…	…	…	1	…	…	…	…	…	…
…	…	…	…	1	…	…	…	…	…
Patent *m*	0.3636	0.2580	…	…	1	0.0477	…	…	0.4221
Product 1	0.0275	0.0289	…	…	0.0477	1	0	0	0
…	…	…	…	…	…	0	1	0	0
…	…	…	…	…	…	0	0	1	0
Product *n*	0.1105	0.1130	…	…	0.4221	0	0	0	1

Identifying potential infringement between products and patents

Average Infringing Possible Product Index (AIPPI)

The Average Infringing Possible Product Index (AIPPI) is proposed to interpret the infringement map. The products that have a high possibility of patent infringement among competitors' products are identified by measuring AIPPI since they have high technological similarity with a company's own patents.

To measure AIPPI, this research utilises the concept of the actor degree centrality in network analysis. The simplest definition of actor centrality is that central actors must be the most active in the sense that they have the most ties to other actors in the network (Wasserman and Faust 1994). Thus, the higher is the degree of centrality of a product, the higher is the possibility of infringement. The product that has a high degree of centrality is linked to many patents that have high technological similarity. In other words, the product has a high potential to infringe on the company's own patents.

The degree centrality in the bipartite graph was utilised in this research, and a two-mode network was generated that had two entities–patents and products. The degree centrality index in the bipartite graph is different from the degree centrality typically used in a one-mode network, which is defined as the number of ties incident upon node i because the nodes were partitioned in exactly two mutually exclusive sets such that there were no ties wholly within either set in the bipartite graph (Borgatti and Everett 1997). The degree centrality index in the bipartite graph is defined as follows:

$$d_i = \frac{d_i}{n_{\text{prod}}}, \quad \text{for } i \in V_{\text{pat}}. \tag{5}$$

$$\text{AIPPI} = d_j = \frac{d_j}{n_{\text{pat}}}, \quad \text{for } j \in V_{\text{prod}}, \tag{6}$$

where V_{pat} is a set of patent nodes and V_{prod} is a set of product nodes. Also, d_i is the number of ties incident with patent i and product j in the opposing set. For a node in V_{pat}, the maximum number of ties it could have is n_{prod}, which is the number of nodes in V_{prod} (Borgatti and Everett 1997). Thus, the average infringing possible product (AIPP) is identified by interpreting d_j and the product with high d_j had high potential to infringe on the company's own patents. Using this index, a company can search for targets for aggressive patent strategy. In other words, the index can be used as a scanning index to locate the competitors' potentially infringing products.

Critical Infringing Possible Patent Index (CIPPI)

The Critical Infringing Possible Patent Index (CIPPI) is proposed to identify patents that are critical to competitors' products among a company's own patents. Identifying a critical patent among a company's own patents is important because most of the time a core technology or patent is powerful in industry and has a high potential to be infringed by many competitors' products. The CIPPI is suggested by calculating the maximum technological similarity to each competitor's product from the patent–product technological similarity matrix. Using this index, companies can search for the critical weapon of an aggressive patent strategy.

Subgrouping

Subgrouping methods such as cliques (Luce and Perry 1949), n-cliques (Luce 1950; Alba 1973), n-clans, n-clubs (Mokken 1979), k-plexes (Seidman and Foster 1978), lambda sets (Borgatti, Everett and Shirey 1990), and ls-sets (Seidman 1983) are not well suited for analysing a bipartite graph. In fact, bipartite graphs contain no cliques, as strictly defined by Luce and Perry. In contrast, bipartite graphs contain too many two-cliques and two-clans. The FACTIONS routine in UCINET (Borgatti, Everett and Freeman 2002) takes the bipartite graph as input and uses a combinatorial optimisation algorithm called Tabu Search (Glover 1989) to assign nodes to as many clusters as hypothesised by the user so as to maximise a fit criterion.

The two-mode data set should be transformed into a one-mode data set format for subgrouping using FACTIONS. The one-mode network is easier for observing the technological similarity between patents since similarity between patents is not considered in the two-mode network. Several groups consisting of patents and competitors' products are extracted by a grouping process based on technological similarity. The group characteristic is identified by investigating the IPC or keywords that the patents included. An area of common interest in technology can be extracted between the competitor's product and one's own patents in the same group, and the IPC provides for a hierarchical system of language-independent symbols for the classification of patents and utility models according to the different areas of technology to which they pertain. Keywords of the group, which were extracted by text mining, are suggested as detail contents in the area of technology.

Static analysis and dynamic analysis

Static analysis and dynamic analysis are conducted in this research. The AIPP is suggested with the AIPPI and a critical infringing possible patent (CIPP) is suggested with the CIPPI in both static analysis and dynamic analysis. The possibility of infringement on the patent and product is identified at the time of static analysis. The relationship between patents and products is visually represented with a connected arc on the map. AIPP and CIPP are suggested as a table with a value of degree centrality in the two-mode network and a maximum technology similarity. Moreover, an average technological similarity (ATS), which is an additional index, is presented for each competitor's product from the patent–product technological similarity matrix. AIPPI is the value that is extracted from degree centrality based on technological similarity, while ATS is the average value of technological similarity to each competitor's product. Thus, ATS offers additional information when patent and product nodes are not connected to each other; in other words, when the value of AIPPI is zero.

Dynamic analysis is conducted to understand the possibility of the change of patent infringement. The timing of patent appearances can be visually observed on the map as time goes by. Also, changes in infringement possibility based on technological similarities between patents and products can be observed with the table, which has the information on AIPP and CIPP. In conclusion, the possibility of patent infringement at present is confirmed by overall static analysis. Then details such as timing of patent appearances and change of infringement possibility based on technological similarity are confirmed by dynamic analysis.

Illustration

Data collection and treatment

The method proposed in this research was applied to the light emitting diode (LED) lamp technology and products because LED lamp technology is promising for the future and some patent infringement litigations are already ongoing.

In previous research on patent infringement with the SAO approach, Bergmann et al. (2007, 2008) conducted a case study using DNA chip technology. DNA chip technology has two suitable processes for gene expression analysis, the photolithographic procedure and the contact-tip-deposition-printing procedure. This study focused on the methodologies in this field of technology and tried to identify the possibility of infringement. Park, Yoon and Kim (2012) conducted a case study with technology for prostate cancer treatment in the pharmaceutical industry. Patents on pharmaceuticals are largely developed in regard to formulation, composition, and dosage of pharmaceuticals. It is not suitable for investigating infringement between products and patents, though these two technology fields have products like DNA chip fabrication equipment and medicine. However, products in LED technology field are diverse and it is possible to identify infringement between patents and products since LED lamps are tangible and there are patents in regard to devices as well as methods.

Five hundred and twenty patents, most of which were held by 10 patent holders, were collected from among 1892 patents related to LED lamps in USPTO. Patents were screened by a domain expert and design patents were not considered because they did not include textual technological information. It was supposed that the company of interest was OSRAM Sylvania Inc., which is one of the affiliates of OSRAM and the fourth corporation on our list of companies holding the most patents. OSRAM has 23 patents on LED lamps and validates our method of generating the patent infringement map since it has a patent-litigation case in reality.

SAO structures for patents were extracted from the claim in 23 OSRAM patent documents and SAO structures for products were extracted from the description section of patent documents of rival corporations. SAO structures for the product extracted from the patents of every corporation were aggregated. In the process of aggregating patents, only patents that had a predefined IPC as related to LED lamps (e.g. F21: lighting, H01: basic electric elements) were selected. The SAO structure extraction process was conducted using Knowledgist 2.5, a textual language processing system. Then the semantic similarity between these extracted SAO structures was computed based on the similarity of the pairs of words (Simpson and Dao 2005). Since WordNet is a general-purpose language repository, some acronyms specific to the technology domain were not defined in WordNet and were added. These included CCT (correlated colour temperature), CRI (colour rendering index), DALI (digital addressable lighting interface), SDCM (standard deviation of colour matching) and ESD (electrostatic discharge). These domain-specific acronyms were based on the arranged terminologies from the LED technology companies' websites.

Two-mode product–patent infringement map

A product–patent infringement map is generated through the creation of a semantic patent-product matrix by measuring semantic sentence similarity. We generated a two-mode network infringement map and a one-mode network infringement map. The two-mode network infringement map was developed as a bipartite graph based on the (23 × 9) asymmetric two-mode matrix created, which represented 23 OSRAM patents and nine products of competitors. It showed how similar each of the 23 patents was to every competitor's product.

(1) Patent infringement static analysis

The two-mode product–patent infringement map was generated by static analysis, shown in Figure 2. The threshold was determined as 0.8 between SAO sentences and 0.3 between documents by conducting sensitive analysis. This map displays the possibility of patent infringement based on technological similarity between patents owned by OSRAM and competitors' products.

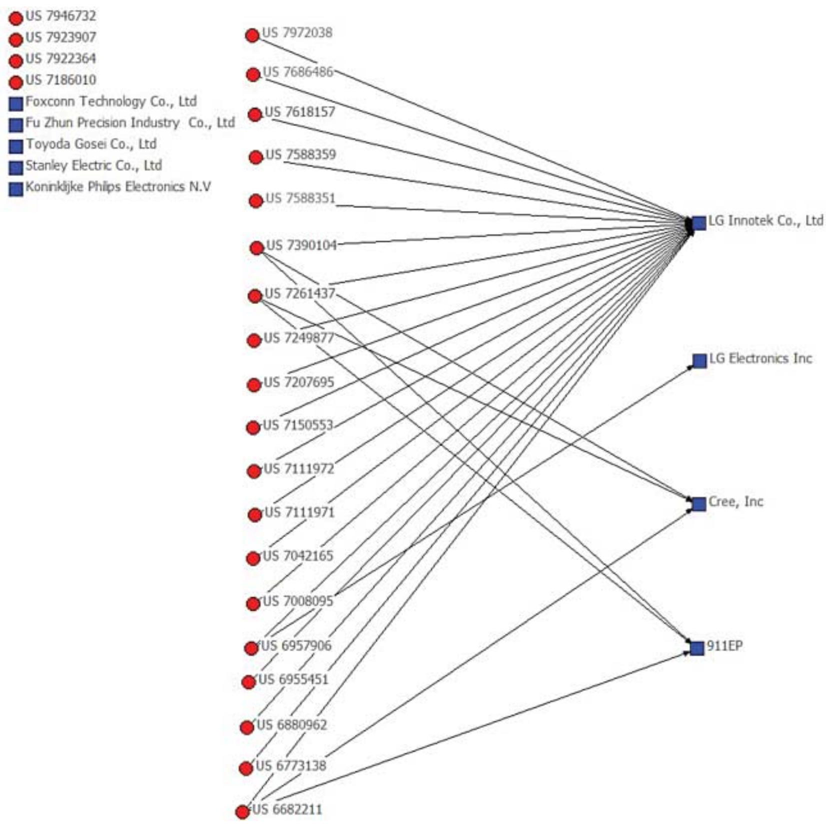

Figure 2. Two-mode network product-patent infringement map.

LG Innotek's product had the most links among the competitors' products. The products of Cree Inc., 911EP and LG Electronics also had links with each patent, as Figure 2 shows.

The AIPP, which was extracted based on the degree centrality in the network analysis, was proposed to identify the potential of competitors' products infringing on a company's own patents. Comparing the value of the ATS, some values of AIPPI are zero because the AIPPI is calculated based on degree centrality after generating a network using threshold 0.3 between documents. LG Innotek's product had the highest AIPPI among the competitor's products. The products of 911EP and Cree Inc. followed in sequence.

The CIPP was identified using the CIPPI, which shows the maximum technological similarity to each competitor's product from the patent–product technological similarity matrix. Table 3 represents the AIPPI, ATS and CIPPI between the product and patent. The AIPPs were the products of LG Innotek, critical patents that infringe on the competitor's product among the company's own patents based on the maximum technical similarity. For example, US7390104 was the CIPP for LG Innotek's product among OSRAM's 23 patents. Table 4 shows the top five AIPPs and a corresponding CIPP. The company scans the AIPP, and then finds out the CIPP based on technological similarity. For instance, since LG Innotek's product has the highest possibility of infringement for OSRAM's own patents on average, if patent-infringement litigation occurs between OSRAM and LG Innotek, patent US7390104 could be a critical patent among OSRAM's

Table 3. AIPP and CIPP between product and patent.

	Foxconn Technology Co. Ltd	Fu Zhun Precision Industry Co. Ltd	LG Innotek Co. Ltd	LG Electronics Inc	Toyoda Gosei Co. Ltd	Stanley Electric Co. Ltd	Koninklijke Philips Electronics N.V	Cree Inc	911EP
US 7972038	0.028	0.029	0.367	0.214	0.078	0.110	0.082	0.140	0.110
US 7946732	0.023	0.024	0.178	0.000	0.065	0.093	0.093	0.088	0.113
US 7923907	0.025	0.027	0.213	0.077	0.072	0.102	0.103	0.185	0.123
US 7922364	0.033	0.035	0.296	0.242	0.094	0.132	0.116	0.157	0.151
US 7686486	0.023	0.024	0.711	0.167	0.065	0.093	0.093	0.198	0.226
US 7618157	0.014	0.015	0.541	0.250	0.040	0.057	0.063	0.126	0.142
US 7588359	0.023	0.024	0.711	0.167	0.065	0.093	0.104	0.209	0.226
US 7588351	0.012	0.012	0.514	0.286	0.033	0.048	0.043	0.116	0.120
US 7390104	0.035	0.037	0.714	0.229	0.100	0.140	0.156	0.321	0.330
US 7261437	0.050	0.052	0.493	0.083	0.139	0.193	0.205	0.330	0.368
US 7249877	0.025	0.027	0.340	0.077	0.072	0.102	0.108	0.196	0.235
US 7207695	0.033	0.035	0.630	0.121	0.094	0.132	0.116	0.178	0.194
US 7186010	0.022	0.023	0.273	0.130	0.062	0.088	0.062	0.110	0.114
US 7150553	0.028	0.029	0.490	0.214	0.078	0.110	0.118	0.258	0.254
US 7111972	0.018	0.019	0.439	0.200	0.053	0.075	0.084	0.180	0.185
US 7111971	0.028	0.029	0.571	0.143	0.078	0.110	0.123	0.215	0.199
US 7042165	0.018	0.019	0.488	0.000	0.053	0.075	0.084	0.180	0.173
US 7008095	0.014	0.015	0.541	0.250	0.040	0.057	0.063	0.126	0.130
US 6957906	0.008	0.009	0.438	0.364	0.023	0.033	0.038	0.071	0.085
US 6955451	0.029	0.030	0.560	0.138	0.081	0.115	0.128	0.214	0.242
US 6880962	0.015	0.016	0.579	0.235	0.043	0.061	0.069	0.149	0.153
US 6773138	0.047	0.049	0.333	0.133	0.130	0.181	0.147	0.246	0.212
US 6682211	0.048	0.050	0.537	0.130	0.133	0.185	0.206	0.382	0.422
AIPPI	0	0	0.826	0.043	0	0	0	0.130	0.130
ATS	0.026	0.027	0.476	0.167	0.073	0.104	0.105	0.190	0.196
CIPPI	0.050	0.052	0.714	0.364	0.139	0.193	0.206	0.382	0.422

Table 4. Top 5 AIPP and a corresponding CIPP in patent infringement static analysis.

Rank	AIPP	AIPPI	ATS	CIPP	CIPPI
1	LG Innotek Co. Ltd	0.826	0.476	US 7390104	0.714
2	911 EP	0.130	0.196	US 6682211	0.422
3	Cree Inc	0.130	0.190	US 6682211	0.382
4	LG Electronics Inc	0.043	0.167	US 6957906	0.364
5	Koninklijke Philips Electronics N.V	0	0.105	US 6682211	0.206

own patents. In other words, OSRAM would be able to identify LG Innotek's product as a target for patent-infringement litigation and US7390104 could be used as a critical weapon in aggressive patent infringement litigation strategy. The AIPPI, ATS and CIPPI are only numerical values for relative comparison and are not absolute numerical values to represent the possibility of infringement. Therefore, the analyst needs to review the patents and competitor's product in detail when making a patent strategy. Nonetheless, this method can be a useful supporting tool for decision making regarding patent strategy in that it suggests AIPP and CIPP, which help the

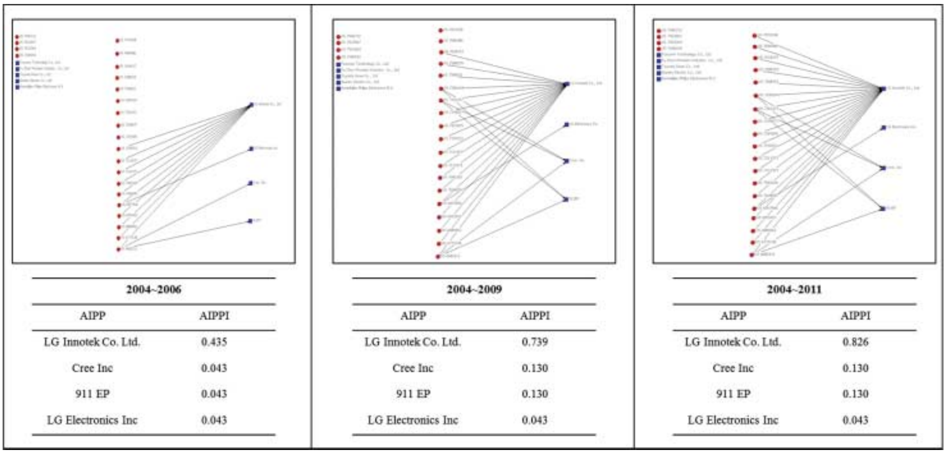

Figure 3. Cumulative dynamic analysis of two-mode network product-patent infringement map.

user to set the scope before searching for details on all of the patents and products in a specific technology or product field.

As reviewing patent documents, OSRAM holds patents regarding LED lamp structure, arrangement, assembly and light source. LG Innotek possesses patents related to LED lamp module, assembly and light sources and LG Electronics has LED light source patents. Cree and 911EP hold patents regarding lamp package and lamp assembly. Product nodes of these firms are linked to patent nodes of OSRAM in the product–patent map because the theme of technology is partially similar in general. However, Foxconn Technology and Fu Zhun Precision Industry possess patents regarding LED lamp cooling apparatus and heat dissipation. Toyoda Gosei and Stanely Electric hold patents regarding vehicle lamp and light guide. Koninklijke Philips Electronics owns patents related to high performance of LED lamp system and colour of light.

(2) Patent infringement dynamic analysis

After conducting static analysis, dynamic analysis can obtain additional information from the infringement map, such as the appearance of new AIPP, change of AIPP rank and change of CIPP according to the time period. The appearance of product launch timing is not considered in this paper since information on products is considered as an aggregate of information from the description section of patent documents. Three cumulative product–patent infringement maps were generated by dynamic analysis, which was conducted every three years from 2004 to 2011 and based on patent issue year in the case of this paper. The map included cumulative patent data so that the information on the appearance of new patents and new links was visually represented according to time period, shown in Figure 3. Furthermore, AIPP with AIPPI is suggested for each time period. In this case, four AIPPs appeared on three maps: LG Innotek, Cree Inc., 911 EP and LG Electronics Inc. The rank of AIPPI for each AIPP did not change, but the value increased as time passed. For example, the AIPPI with LG Innotek increased from 0.435 to 0.826. Table 5 presents the change of CIPP for each time period. It suggests the change of CIPP among patents of OSRAM according to every competitor's product. The CIPP for LG Innotek was US6880962 with 0.579 of CIPPI in the first time period, whereas it changed to US7390104 with 0.714 of CIPPI in the second time period as new patents appeared. However, the CIPP for Cree Inc. was US6682211 and there was no change. As interpreting the dynamic analysis, newly appearing OSRAM patents

Table 5. Cumulative dynamic analysis of CIPP.

	Time period					
	2004–2006		2004–2009		2004–2011	
AIFP	CIPP	CIPPI	CIPP	CIPPI	CIPP	CIPPI
Foxconn Technology Co. Ltd	US 6682211	0.048	US 7261437	0.050	US 7261437	0.050
Fu Zhun Precision Industry Co. Ltd	US 6682211	0.050	US 7261437	0.052	US 7261437	0.052
LG Innotek Co. Ltd	US 6880962	0.579	US 7390104	0.714	US 7390104	0.714
LG Electronics Inc	US 6957906	0.364	US 6957906	0.364	US 6957906	0.364
Toyoda Gosei Co. Ltd	US 6682211	0.133	US 7261437	0.139	US 7261437	0.139
Stanley Electric Co. Ltd	US 6682211	0.185	US 7261437	0.193	US 7261437	0.193
Koninklijke Philips Electronics N.V	US 6682211	0.206	US 6682211	0.206	US 6682211	0.206
Cree Inc	US 6682211	0.382	US 6682211	0.382	US 6682211	0.382
911EP	US 6682211	0.422	US 6682211	0.422	US 6682211	0.422

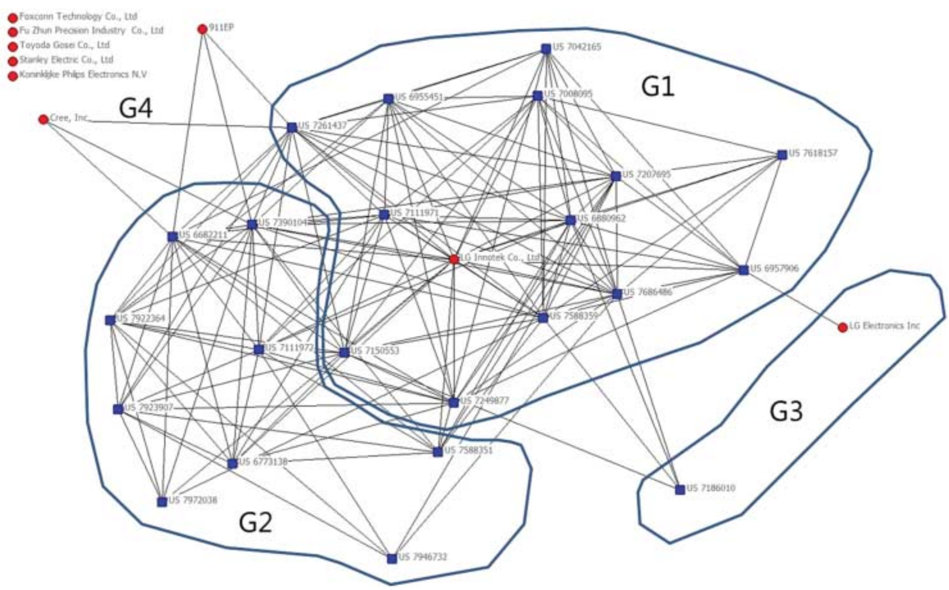

Figure 4. Subgroups from transformed one-mode network infringement map.

were technologically similar to LG Innotek's products because the CIPPI changed as time passed. However, OSRAM has not been developing technology similar to Cree Inc.'s product, since the CIPPI has not changed over time. Thus, OSRAM should consider an investigation of LG Innotek's product rather than that of Cree Inc.

As reviewing contents of patents in dynamic analysis, OSRAM largely holds patents regarding LED lamp structure in all periods while patents regarding light source are only in the 2004–2006 period. US6880962 which is a patent related to light source was the CIPP of LG Innotek in the 2004–2006 period. However, the CIPP of LG Innotek in the next two periods is changed to US7390104 which is patent regarding portable LED lamp structure. The CIPP of LG Electronics,

Table 6. Density table of subgroups.

	Group 1	Group 2	Group 3	Group 4
Group 1	0.80	0.35	0.09	0.03
Group 2	0.35	0.75	0.00	0.09
Group 3	0.09	0.00	0.00	0.00
Group 4	0.03	0.09	0.00	0.00

Table 7. Subgroups in one-mode network infringement map.

Group	Member	Main IPC	
G1	US 7686486, US 7618157, US 7588359, US 7261437, US 7249877, US 7207695, US 7150553, US 7111971, US 7042165, US 7008095, US 6957906, US 6955451, US 6880962, LG Innotek Co., Ltd	F21V-014/00, F21V-021/00, H01L-033/00, H05B-037/02	F21V-029/00, F21V-007/04, F21S-008/10,
G2	US 7972038, US 7946732, US 7923907, US 7922364, US 7588351, US 7390104, US 7111972, US 6773138, US 6682211	F21V-003/00, F21L-004/02, H01R-033/00	
G3	US 7186010, Fu Zhun Precision Industry Co., Ltd, LG Electronics Inc, Koninklijke Philips Electronics N.V	B60Q-001/00	
G4	Foxconn Technology Co., Ltd, Toyoda Gosei Co., Ltd, Stanley Electric Co., Ltd, Cree, Inc 911EP	—	

which usually possesses patents related to light source, matched US6957906 which is a patent regarding light source in all time periods without change.

One-mode product–patent infringement map

Figure 4 shows four subgroups that were extracted from the transformed one-mode network infringement map using the FACTIONS routine in UCINET with the final proportion correction at 0.802. The value is a good fit, as the final proportion correction can be considered an indicator of fit measurement. Group 1 (G1) consisted of the patents of LG Innotek's product, which had the highest AIPP. Group 2 (G2) consisted of only OSRAM's patents. Group 3 consisted of the patents of OSRAM and LG Electronics and the products of other competitors. Group 4 consisted of only the products of other competitors; however, this group is like the outsider of the network since the group has very low technological similarity to the OSRAM's patent. FACTIONS also indicated the density within groups that were identified. Table 6 shows the density of groups; G1 has highest density within the group and G2 follows in sequence.

In previous studies, Bergmann et al. (2007, 2008) and Park, Yoon and Kim (2012) suggested comparison between patents using multidimensional scaling based on semantic similarity, whereas this study suggests comparison between companies' patents and competitors' products using networks. They focused on evaluating the risk of patent infringement for defensive patent strategy, while this research focused on the exploitation of aggressive patent strategy by measuring the

Table 8. IPC and keyword in subgroup.

Group	IPC	Description	Keywords
G1	F21V-014/00	Changing the characteristics or distribution of the light emitted by adjustment of parts	electrical, head, reflector, connected linearly, exterior, engine, resistor, transverse, surface, arranged, longitudinal
	F21V-021/00	Supporting, suspending, or attaching arrangements for lighting devices	
	H01L-033/00	Semiconductor devices with at least one potential-jump barrier or surface barrier specially adapted for light emission; Processes or apparatus specially adapted for the manufacture or treatment thereof or of parts thereof	
	H05B-037/02	Circuit arrangements for electric light sources in general - Controlling	
G2	F21V-003/00	Globes; Bowls; Cover glasses	chips, positioned, optical, adjacent, contact, intermediate, latching, thermal, receptacle, adapter, connection, socket, tubular
	F21L-004/02	Lighting devices or systems thereof, being portable or specially adapted for transportation - Electric lighting devices with self-contained electric batteries or cells	
	H01R-033/00	Coupling devices specially adapted for supporting apparatus and having one part acting as a holder providing support and electrical connection via a counterpart which is structurally associated with the apparatus	
G1 & G2	F21V-029/00	Cooling or heating arrangements	heat, circuit, sink, assembly, support, board, extending, mounted, coupling, conductive, vehicle
	F21V-007/04	Reflectors for light sources - Optical design	
	F21S-008/10	Lighting devices intended for fixed installation - specially adapted for vehicles	
G3	B60Q-001/00	Arrangement of optical signalling or lighting devices, the mounting or supporting thereof or circuits therefor	frusto-conical, surface, cover, circumferential, fixed, flange, concave, double-sided, studs, aperture

possibility of infringement between patents and products. Thus, the proposed patent map suggests a company's patents that have a possibility of infringement with competitors' products.

Furthermore, a company's analogous patent groups can be considered that company's patent portfolio.

Table 7 summarises the characteristics of the subgroup in the one-mode network infringement map. Even though the patent includes many IPCs, we only considered the main IPC for extracting the core characteristic. All of the patents belong to the four categories of IPC, which were: F (mechanical engineering, lighting, heating, weapons, blasting), H (electricity), B (performing operations, transporting) and G (physics). Both G1 and G2 shared three common IPC sections and a respective group had several IPC sections.

Table 8 summarises the information on IPC sections and the corresponding keywords, which were the most frequent in the group of patents after eliminating noise such as definite articles (e.g. 'the'), conjunctions (e.g. 'and', 'but') and general keywords in the technology field (e.g. 'LED', 'lamp', 'light'). The overall technology theme was suggested by representing the IPC sections and specific keywords regarding the technology theme were also suggested by keywords. Most of the suggested keywords corresponded with the technology theme. For example, G1 had a technology theme on changing light emission, arrangement for light devices, semiconductor devices and circuit arrangement control. Suggested keywords corresponded with the technology theme and included 'reflector for light emission', 'resistor for semiconductor devices', and 'transverse, longitudinal for circuit arrangement control'. G1 was a significant subgroup for OSRAM in that LG Innotek, the most infringing possible product on average, was included with many patents. Thus, OSRAM should focus on common interests in the technology area and specific keywords in G1 when they make a decision on patent strategy. In other words, the method can be used as a tool to support decision making.

Conclusions

The results of analysis in this study highlight the possibility of patent infringement between OSRAM Sylvania and LG Innotek. In fact, patent litigation occurred between OSRAM Opto semiconductor and LG Innotek in 2011. OSRAM Opto and OSRAM Sylvania are closely related to each other and conduct joint research and development. In addition, OSRAM is vertically systemised with OSRAM Opto semiconductor for LEDs, OSRAM LED systems for modules, and OSRAM Sylvania for lamps. OSRAM Opto is in charge of packaging and epi chips, and OSRAM Sylvania is in charge of lamps. OSRAM Sylvania could be selected instead of OSRAM Opto semiconductor after searching the patents related to LED lamps since this research focuses on patent infringement with regard to the product. Thus, LG Innotek is considered a potential target for patent infringement when it comes to products in the future, as the analysis showed that OSRAM Opto, which is a closely connected firm, has already had a patent litigation with LG Innotek and LG Electronics.

This research suggested using an automated SAO-based semantic technological similarity measurement method to identify patent infringement in respect to the product by generating a product–patent infringement map. This method can be used as a tool to support decision making when a company's patent strategy is determined since the map suggests the average infringing possible product among competitor's products and the critical infringing possible patent among one's own patents. The results of dynamic analysis and subgrouping imply useful information for deploying patent strategy. Forecasting possible patent infringement is useful for reducing the risk from the perspective of reacting to patent litigation. Furthermore, a company can gain a competitive advantage in the market by deploying a strategic patent-litigation strategy.

This research contributes to strategic patent-management research in that a new approach is suggested that identifies the possibility of patent infringement between a patent and product. Nonetheless, there are limitations to the method in that each product is not considered a real business because the product is considered an aggregate of patents owned by competitors, and the description section of the patent substitutes for the product document. These limitations will be solved when well-organised product document is collected at the company level. If an official product database like the patent database were constructed, the method would be supplemented by this source of information.

Only technological similarity was considered when it came to mapping possible patent and product infringement, although patent-infringement litigation is closely related to a complex business situation. Furthermore, the legal characteristics of a patent also need to be considered. If a patent that is involved in patent litigation is an essential patent, antitrust laws apply to the case because the essential patent allows for the last mover in the industry to easily use the outcome of research and development. Thus, many factors need to be considered, as the patent itself has characteristics that are related to technology, legal aspects, and business. Nevertheless, this method can be used to set the overall scope based on the technological view before in-depth analysis by humans reflecting the legal and business views.

For future research, well-organised product documents should be identified to apply this method more accurately. In addition, other factors such as those related to legal factors and business are considered as well as technological similarity as used to identify patent infringement. Moreover, a more accurate quantitative or qualitative method of validation is necessary.

Acknowledgements

This work was partially supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2012R1A1A1011934).

Notes on contributors

Byungun Yoon is an Associate Professor in Department of Industrial & Systems Engineering of Dongguk University. His work experience includes an IT consultant in LG and a visiting scholar in the Centre for Technology Management (CTM) of the University of Cambridge. His theme of study has involved patent analysis, new technology development methodology and visualisation algorithms. His current interest is in enhancing technology roadmapping and product designing with data mining techniques. He has authored articles published in *Research Policy*, *R&D Management*, *Technological Forecasting and Social Change*, *Technology Analysis & Strategic Management*, among others.

Inchae Park is currently a PhD candidate in Department of Industrial & Systems Engineering of the Dongguk University. His theme of study has involved technology forecasting, data mining and patent analysis.

References

Alba, R.D. 1973. A graphic-theoretic definition of a sociometric clique. *Journal of Mathematical Sociology* 3: 113–26.

Al-Shboul, B., and S. Myaeng. 2011. Query phrase expansion using wikipedia in patent class search. In *Proceedings of the 7th Asia Conference on Information Retrieval Technology*, 115–26. Berlin: Springer.

Arundel, A. 2001. The relative effectiveness of patents and secrecy for appropriation. *Research Policy* 30: 611–24.

Berger, F., K. Blind, and A. Cuntz. 2012. Risk factors and mechanisms of technology and insignia copying – a first empirical approach. *Research Policy* 41: 376–90.

Bergmann, I., D. Butzke, L. Walter, J.P. Fuerste, M.G. Moehrle, and V.A. Erdmann. 2008. Evaluating the risk of patent infringement by means of semantic patent analysis: The case of DNA chips. *R&D Management* 38: 550–62.

Bergmann, I., M.G. Moehrle, L. Walter, D. Butzke, V.A. Erdmann, and J.P. Fuerste. 2007. The use of semantic maps for recognition of patent infringements: A case study in biotechnology. Special issue: Open innovation between and within organizations. *Zeitschrift fuer Betriebswirtschaft* 4: 69–86.

Borgatti, S.P. 2009. Two-mode concepts in social network analysis. In *Encyclopedia of complexity and system science*, ed. R.A. Meyers, 8279–91. New York: Springer.

Borgatti, S.P., and M.G. Everett. 1997. Network analysis of 2-mode data. *Social Networks* 19: 243–69.

Borgatti, S.P., M.G. Everett, and L.C. Freeman. 2002. *Ucinet for Windows: Software for social network analysis*. Cambridge, MA: Harvard Analytic Technologies.

Borgatti, S.P., M.G. Everett, and P.R. Shirey. 1990. LS sets, lambda sets and other cohesive subsets. *Social Networks* 12: 337–57.

Bryant, J.L. 1997. *Protecting your ideas: The inventor's guide to patents*. San Diego, CA: Academic Press.

Carree, M.A., L. Klomp, and A.R. Thurik. 2000. Productivity convergence in OECD manufacturing industries. *Economics Letters* 66: 337–45.

Cross, R., S. Borgatt, and A. Parker. 2001. Beyond answers: Dimensions of the advice network. *Social Networks* 23: 215–35.

Durham, A.L. 2004. *Patent law essentials: A concise guide*. Westport, CT: Praeger.

Fellbaum, C. 2005. WordNet and wordnets. In *Encyclopedia of language and linguistics*, 665–70. Oxford: Elsevier.

Franzosi, R. 1994. From words to numbers: A set theory framework for the collection, organization and analysis of narrative data. *Sociological Methodology* 24: 105–36.

Fujii, A., M. Iwayama, and N. Kando. 2007. Introduction to the special issue on patent processing. *Information Processing and Management* 43: 1149–53.

Gelsing, L. 1992. Innovation and the development of industrial networks. In *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*, ed. B. Lundvall, 116–28. London: Frances Pinter.

Gerken, J., M. Moehrle, and L. Walter. 2010. Patents as an information source for product forecasting: Insights from a longitudinal study in the automotive industry. Paper presented at the R&D Management Conference 'Information, Imagination and Intelligence in R&D Management', 29 June–2 July 2010, Manchester, UK.

Glover, F. 1989. Tabu search – Part 1. *ORSA Journal of Computing* 1: 190–206.

Golden, J.M. 2007. Patent trolls and patent remedies. *Texas Law Review* 85: 2111–61.

Hall, B.H., and R.H. Ziedonis. 2001. The patent paradox revisited: An empirical study of patenting in the US semiconductor industry, 1979–1995. *The RAND Journal of Economics* 32: 101–28.

Kim, Y., J. Suh, and S. Park. 2008. Visualization of patent analysis for emerging technology. *Expert Systems with Applications* 34: 1804–12.

Knoke, D., and J. Kuklinski. 1982. *Network analysis*. London: Sage.

Lee, C., B. Song, and Y. Park. 2013. How to assess patent infringement risks: A semantic patent claim analysis using dependency relationships. *Technology Analysis & Strategic Management* 25: 23–38.

Lemley, M.A., and C. Shapiro. 2007. Patent holdup and royalty stacking. *Texas Law Review* 85: 1991–2048.

Leoncini, R., M. Maggioni, and S. Montressor. 1996. Intersectorial innovation flows and national technological system network analysis for comparing Italy and Germany. *Research Policy* 25: 415–30.

Luce, D. 1950. Connectivity and generalized cliques in sociometric group structure. *Psychometrika* 15: 169–90.

Luce, D., and A. Perry. 1949. A method of matrix analysis of group structure. *Psychometrika* 14: 95–116.

Majewski, S.E., and D.V. Williamson. 2004. Incomplete contracting and the structure of R&D joint venture contracts. In *Intellectual property and entrepreneurship*, ed. G. Libecap, Vol. 15, 201–228. Oxford: Elsevier.

Marseden, P., and E. Laumann. 1984. Mathematical ideas in social structural analysis. *Journal of Mathematical Sociology* 10: 271–94.

Miller, G.A. 1995. WordNet: A lexical database for English. *Communications of the ACM* 38: 39–41.

Miller, G.A., C.D. Beckwith, D. Fellbaum, and G.K. Miller. 1990. WordNet: An online lexical database. *International Journal of Lexicograph* 3: 235–44.

Moehrle, M.G., L. Walter, A. Geritz, and S. Muller. 2005. Patent-based inventor profiles as a basis for human resource decisions in research and development. *R&D Management* 35: 513–24.

Mokken, R. 1979. Cliques, clubs and clans. *Quality and Quantity* 13: 161–73.

Moore, G.E. 1965. Cramming more components onto integrated circuits. *Electronics Magazine* 38: 114–17.

OECD. 2008. *The economic impact of counterfeiting and piracy*. Paris: Organisation for Economic Co-operation and Development.

Park, Y., and M. Kim. 1999. A taxonomy of industries based on knowledge flow structure. *Technology Analysis and Strategic Management* 11: 541–49.

Park, H., J. Yoon, and K. Kim. 2012. Identifying patent infringement using SAO based semantic technological similarities. *Scientometrics* 90: 515–29.

Pohlmann, T., and M. Opitz. 2013. Typology of the patent troll business. *R&D Management* 43: 103–20.

Reitzig, M., J. Henkel, and C.H. Heath. 2007. On sharks, trolls, and their patent prey unrealistic damage awards and firms' strategies of 'being infringed'. *Research Policy* 36: 134–54.

Resnik, P. 1999. Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language. *Journal of Artificial Intelligence Research* 11: 95–130.

Savransky, S. 2000. *Engineering of creativity: Introduction to TRIZ methodology of inventive problem solving*. Boca Raton, FL: CRC Press.

Schmiele, A. 2013. Intellectual property infringements due to R&D abroad? A comparative analysis between firms with international and domestic innovation activities. *Research Policy*, 42: 1482–95.

Seidman, S. 1983. Internal cohesion of LS sets in graphs. *Social Network* 5: 97–107.

Seidman, S., and B. Foster. 1978. A graph-theoretic generalization of the clique concept. *Journal of Mathematical Sociology* 6: 139–54.

Simpson, T. and T. Dao. 2005. Wordnet-based semantic similarity measurement. http://www.codeproject.com/KB/string/semanticsimilaritywordnet.aspx.

Soo, V.W., S.Y. Lin, S.Y. Yang, S.N. Lin, and S.L. Cheng. 2006. A cooperative multi-agent platform for invention based on patent document analysis and ontology. *Expert Systems with Applications* 31: 766–75.

Synder, D., and E. Kick. 1979. Structural position in the world system and economic growth 1955–1970: A multiple network analysis of transnational interactions. *American Journal of Sociology* 84: 1096–126.

Tsourikov, V.M., L.S. Batchilo, and I.V. Sovpel. 2000. Document semantic analysis/selection with knowledge creativity capability utilizing subject–action–object (SAO) structures. United States Patent No. 6167370.

Vasiliki, Z., Y. Kompatsiaris, V. Papastathis, H. Eleftherohorinou, L. Wanner, G. Piella, E. Pianta, L. Serafini, and M. Giereth. 2006. Specification of the semantics-based content- and meta-representation formalism for patent documentation. Report of Project titled in Advanced Patent Document Processing Techniques.

Verhaegen, P., J. Dhondt, J. Vertommen, S. Dewulf, and Duflou, J. R. 2011. Automatically characterizing products through product aspects. In *Global product development*, ed. A. Bernard, 595–605. Berlin: Springer.

Wallerstein, M.B., M.E. Mogee, and R.A. Schoen. 1993. *Global dimensions of intellectual property rights in science and technology*. Washington, DC: National Academies Press.

Wasserman, S., and K. Faust. 1994. *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.

Yang, S.Y., S.Y. Lin, S.N. Lin, S.L. Cheng, and V.W. Soo. 2005. An ontology-based multi-agent platform for patent knowledge management. *International Journal of Electronics Business Management* 3: 181–92.

Yoon, B. 2008. On the development of a technology intelligence tool for identifying technology opportunity. *Expert Systems with Applications* 35: 124–35.

Yoon, J., and K. Kim. 2011. Generation of patent maps using SAO-based semantic patent similarity. *Entrue Journal of Information Technology* 10: 19–27.

Yoon, J., and K. Kim. 2012. Detecting signals of new technological opportunities using semantic patent analysis and outlier detection. *Scientometrics* 90: 445–61.

Yoon, J., H. Park, and K. Kim. 2013. Identifying technological competition trends for R&D planning using dynamic patent maps: SAO-based content analysis. *Scientometrics* 94: 313–31.