An integrated social network mining for product-based technology analysis of Apple

Juhwan Kim
Intellectual Property Strategy Agency, Seoul, Republic of Korea
Sunghae Jun
Department of Statistics, Cheongju University, Cheongju, Republic of Korea and Division of Industrial Management Engineering, Korea University, Seoul, Republic of Korea
Dong-Sik Jang
Division of Information Management Engineering, Korea University, Seoul, Republic of Korea, and
Sangsung Park
Graduate School of Management of Technology, Korea University, Seoul, Republic of Korea

Received 21 October 2016
Revised 28 March 2017
Accepted 9 April 2017

Abstract

Purpose – Patent contains vast information on developed technologies because of the patent system. So, it is important to analyze patent data for understanding technologies. Most previous studies on patent analysis were focused on the technology itself. Their research results lacked the consideration of products. But the patent analysis based on products is crucial for company because a company grows by sales of competitive products. The purpose of this paper is to propose a novel methodology of patent analysis for product-based technology. This study contributes to the product development strategy of a company.

Design/methodology/approach – The primary goal for developing technology is to release a new product. So it is important to analyze the technology based on the product. In this study, the authors analyze Apple's technologies based in iPod, iPhone, and iPad. In addition, the authors propose a new methodology to analyze product-based technology. The authors call this an integrated social network mining (ISNM). In the ISNM, the authors carry out a social network analysis (SNA) according to each product of Apple, and integrate all SNA results of iPod, iPhone, and iPad using the technological keywords.

Findings – In this case study, the authors analyze Apple's technologies according to Apple's innovative products, such as the iPod, iPhone, and iPad. From the ISNM results of Apple's technology, the authors can find which technological detail is more important in overall structure of Apple's technologies.

Practical implications – This study contributes to the management of technology including new product development, technological innovation, and research and development planning. To know the technological relationship between whole technologies based on products can be the source of intensification of technological competitiveness.

Originality/value – Most of studies on technology analysis were focused on patent technology itself. Though one of their research goals was to develop new product, they had their limits considering the products because they did not use the technology information in the technology analysis. The originality of this research is to use the product information in technology analysis using the proposed ISNM.

Keywords Social network analysis, Patent data, Apple products, Integrated social network mining, Technology analysis

Paper type Research paper

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2015R1D1A1A01059742). This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017R1D1A3B03031152).

Industrial Management & Data Systems
Vol. 117 No. 10, 2017
pp. 2417-2430

0263-5577
DOI 10.1108/IMDS-10-2016-0441

1. Introduction

A patent data analysis is used to analyze patent documents as part of a technology analysis in management of technology (MOT) (Roper *et al.*, 2011). A patent document contains rich results of researched and developed technology because the exclusive right of registered patent is protected by patent law (Hunt *et al.*, 2007; Roper *et al.*, 2011). Therefore, it is important to analyze the patent documents to understand the technology in a given domain, and many patent analysis studies for technology management have been introduced (Choi *et al.*, 2015; Jun *et al.*, 2012; Jun *et al.*, 2014; Noh *et al.*, 2015; Rodriguez *et al.*, 2015). Choi *et al.* (2015) analyzed patent data to make a predictive model of technology transfer using social network analysis (SNA), regression, and decision trees. Jun *et al.* (2012) proposed a methodology of patent analysis for vacant technology forecasting by combining clustering and matrix analysis. In addition, Jun *et al.* (2014) tried to solve the sparsity problem of patent document analysis by support vector clustering and Silhouette measure. A text-mining method of patent analysis for technological keyword selection and processing were studied by Noh *et al.* (2015). The techniques of natural language processing such as corpus and text collection were used for patent document analysis. Rodriguez *et al.* (2015) built a graph kernel-based measure to evaluate patent influence for patent citation analysis. Most studies of patent analysis were focused on the technology itself. They analyzed technological keywords or classification codes extracted from the patent documents. The results of most research works lacked the aspects of products. But the patent or technology analysis considering products is very important to a company because a company grows by the sales of innovative products. So in this paper, we propose a novel methodology of patent analysis for product-based technology. This study will implicate how our result of patent analysis could be applied to the product development strategy.

Patent analysis is a field of data science because patent data are also data. Data science includes data collection, data preparation, data analysis, visualization, and management (Stanton, 2013). Figure 1 shows the general data science process.

In the data science process, we begin the collection of data for the target domain. In general, the collected data contain many noises and are incomplete. So it is not suitable for data analysis methods such as statistics and machine-learning algorithms. Thus, we have to transform the collected data into structured data suitable for statistical analysis. In this step, we use the preprocessing based on text-mining techniques (Berry and Kogan, 2010; Han *et al.*, 2012; Kim and Jun, 2015; Tseng *et al.*, 2007). Kim and Jun (2015) transformed the patent documents issued by Apple into structured data using text-mining techniques. The structured data were a matrix composed of patent (row) and keyword (column). Tseng *et al.* (2007) studied on diverse techniques of text mining for patent analysis. They proposed a general methodology of patent text mining such as text segmentation and summarization,

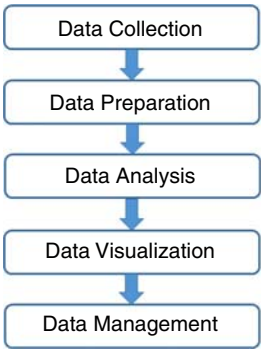

Figure 1.
General data science process

stemming, keyword extraction, and so on. Next, we analyze the structured patent data using statistics or machine-learning algorithms. In addition, we perform data visualization based on diverse tools such as SNA graphs (Butts, 2008; Scott, 2012; Sternitzke *et al.*, 2008). Scott (2012) explained SNA is to analyze the relationship between nodes (objects) in a social network. Butts (2008) made a package called "sna" for SNA using R data language. Furthermore, Sternitzke *et al.* (2008) visualized patent statistics using SNA. SNA is an efficient approach to patent data science as well as general data (Jun, 2012; Jun and Lee, 2014; Nettleton and Salas, 2016). Jun (2012) proposed a method based on SNA for forecasting central technology in given technology domain, and Jun and Lee (2014) showed the "small world effect" of SNA exists in technology areas. Also, Nettleton and Salas (2016) applied the SNA visualization to online data analysis, and got the meaningful performance. Finally, we use the results of patent analysis and visualization for data management. In this study, we use the data science process and propose a novel methodology for patent data analysis and visualization for understanding product-based technology. The proposed methodology is an integrated social network mining (ISNM), which combines diverse SNA models in order to improve the performance of the patent data science for technology analysis. The network models, including SNA, are necessary methods for discovering hidden knowledge from patent data (Jun and Park, 2016; Park *et al.*, 2016; Nettleton and Salas, 2016; Rodriguez *et al.*, 2015). Jun and Park (2016) compared the technologies of BMW with Hyundai to examine their technological competitions using SNA visualization, and Park *et al.* (2016) studied on the technological evolution in the technological field of three-dimensional printing using SNA. In our study, we select Apple's technology as our target domain because Apple is a leading company that has produced many innovative products, including the iPod, iPhone, and iPad. The case studies on Apple have been conducted in order to better understand the company's technological innovation (Jun and Park, 2013; Kim and Jun, 2015). In this paper, we apply the proposed ISNM to analyze Apple's technologies and products, and provide analytical results to determine the company's research and development (R&D) strategy. The remainder of this paper is organized as follows. Section 2 describes the patent analysis and SNA structure. Then, we show the proposed methodology and its application to Apple in Section 3. In Section 4, we discuss the interesting and contribution of this paper. Lastly, Section 5 concludes our research.

2. Patent data analysis using a SNA

A patent analysis is required in an interdisciplinary approach because diverse fields such as statistics, computer science, industrial engineering, or MOT use patent analyses (Hunt *et al.*, 2007; Roper *et al.*, 2011). We can use keywords or international patent classification (IPC) codes as input data for a patent analysis. In particular, the patent analysis result can be used effectively in MOT areas such as R&D planning or technological innovation and impacts. In this study, we conduct a patent data analysis in order to investigate the Apple's technology. In addition, we use the SNA to construct the proposed ISNM. SNA is to analyze the social structure between social objects (nodes) such as friend or family using graph theory based on computer data structure (Butts, 2008; Jun, 2012; Scott, 2012). Butts developed a package based on R data language for performing SNA. Instead of social objects (people), we use technologies or products for the nodes of SNA. We consider keywords and IPC codes for the SNA nodes of technologies based on products. In addition, we apply the SNA model to construct a new methodology for a technology analysis of Apple in our research.

3. ISNM for Apple's technology analysis

In this section, we describe the proposed methodology and perform a case study of Apple's technology using the proposed ISNM method. This is based on patent data science. Our patent data science also follows the process of general data science. Figure 2 shows the proposed process of patent data science used in our study.

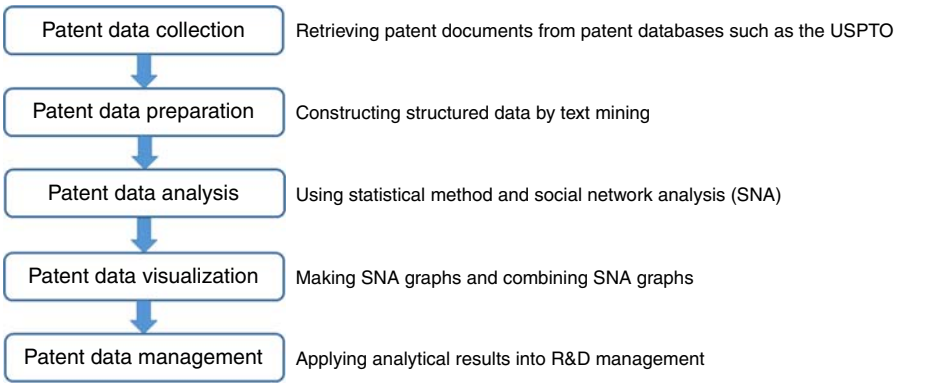

Figure 2.
Process for patent data science

After selecting the target technology, we collect related patent documents by retrieving them from the patent databases, such as the United States Patent and Trademark Office (USPTO) (USPTO, 2016). The collected patent documents are transformed into structured data using the text-mining techniques in the patent data preparation step. Here, we use R project system and its offered package for text mining (Feinerer *et al.*, 2008; Feinerer and Hornik, 2016; R Development Core Team, 2016). In the patent data analysis step, we use a statistical analysis for the patent technology analysis. Here, we consider the SNA for our patent data visualization. Finally, we apply the analysis and visualization results to patent data management, such as Apple's R&D planning. We study a methodology in order to understand the technological innovation process of Apple. In the innovative technology domain, Apple has evolved over many years. In addition, the company has released diverse innovative products in the market, including the iPod, iPhone, and iPad. Apple is a front-runner in the smart device industry. As a result, many researchers have studied the technological evolution of Apple's innovation. In order to analyze Apple's technologies, we propose using the ISNM. First, we classify entire patents according to three Apple products. A patent can overlap the technologies of the iPod, iPhone, and iPad. We also know that a patent technology can be used to develop various products. For example, Patent 1 might be needed to develop the technologies for the iPod and iPad. However, Patent 2 might only be used for the iPhone technology. In our research, we perform a technology analysis of Apple according to the technologies of these three products. Here, we considered an SNA for Apple's technology analysis. From the analytical results, we find the technological relationship between sub-technologies based on keywords or IPC codes. Before analyzing Apple's patent documents, we build the structured data, because the patent document data are not suitable for statistical analyses. To solve this problem, we transform the retrieved patent documents into structured data, as shown in Figure 3.

The patent-keyword matrix consists of patents (rows) and keywords (columns), with each element representing frequency of each keyword. In addition, the patent-IPC code matrix replaces "keyword" with "IPC code." The patent-keyword or IPC code matrices comprise structured data that is well suited to statistical analyses. Our structured matrix data are constructed for three products (i.e. iPod, iPhone, and iPad). Figure 4 shows our approach to the technology analysis of Apple's products.

	Keywords or IPC codes
Patents	Frequency

Figure 3.
Patent – keyword or IPC code matrices

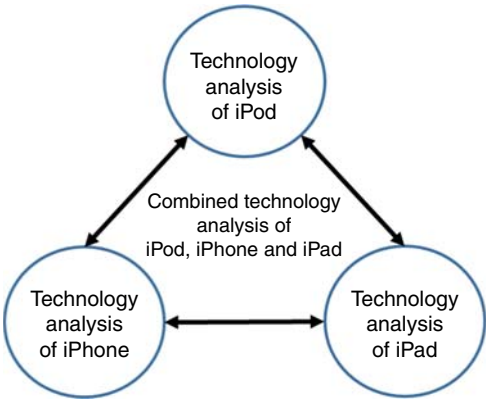

Figure 4.
Combined structure of technology analysis for Apple's products

In this study, we perform a patent analysis for Apple's iPod, iPhone, and iPad. In addition, we combine the three results in order to understand Apple's technological evolution using ISNM. To find Apple's technological evolution, we also identify the overall technology structure, based on the technology analysis results for the iPod, iPhone, and iPad. Using the results from Figure 4, we can determine Apple's R&D strategy and new product development. We retrieved all patents applied by Apple from 1978 to 2013 from various patent databases (USPTO, 2016; WIPSON, 2016). Finally, most patents applied by Apple were searched from the WIPS databases (WIPSON, 2016). In this paper, we collected all patents applied and registered by Apple. Thus, total 6,639 patent documents were retrieved, and we classified them into the three products, as shown in Figure 5.

The numbers of assigned patents for the technologies of the iPod, iPhone, and iPad are 3,465, 4,463, and 4,166, respectively. Here, we had help from the experts of Apple's technology in the Korea Intellectual Property Strategy Agency (www.kista.re.kr) when classifying the patents. We extracted the keywords from the retrieved patent documents for a technology analysis based on Apple keywords. Table I shows the top ten extracted keywords for Apple technologies and products.

Some keywords are included for all products, while others are unique to one product. Using the keywords in Table I, we constructed the integrated SNA result shown in Figure 6.

In our research, we made the SNA graphs according to Apple's products, and we linked up the three SNA graphs via common keyword of three products. In Figure 6, all SNA graphs have the "network" keyword in common. We find that the core keyword connecting the three products of Apple is "network." That is, the network technology is needed to develop all Apple products. We also know that the keywords "interface," "computer," and "control" are central technologies to the development of the iPod. In the iPhone product development, the technologies based on "wireless," "computer," and "network" are

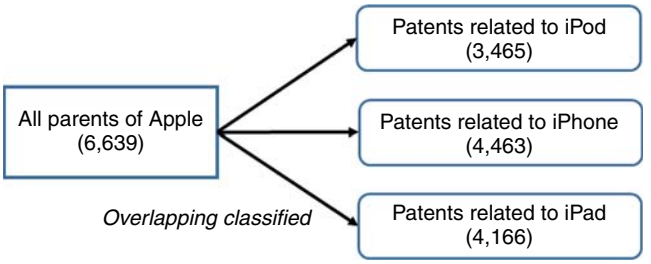

Figure 5.
Classified patents for Apple's technology

Table I.
Top 10 keywords for Apple's products

Rank	iPod	iPhone	iPad
1	computer	digital	signal
2	interface	computer	interface
3	control	mobile	cell
4	detection	network	protocol
5	network	wireless	encoding
6	error	protocol	frequency
7	audio	radio	message
8	cell	audio	mobile
9	frequency	transmission	network
10	operation	video	wireless

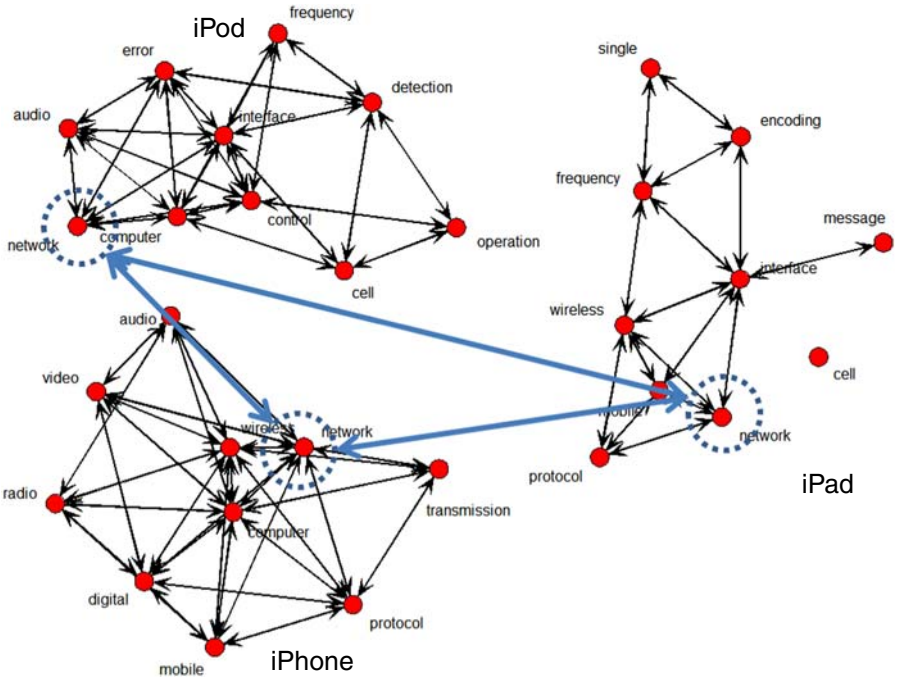

Figure 6.
Integrated SNA graph combined by technology analysis results for the iPod, iPhone, and iPad

important, because they have high degrees (in-degree and out-degree). In addition, the keywords of "interface" and "wireless" are major technology issues for the iPad. In this study, we show another approach to Apple's technology analysis using IPC codes. Table II shows the top 20 IPC codes for the three products.

We know that the technologies based on G06F and H04L are commonly required for all three Apple products. Table III represents the technological definition of each IPC code from the World Intellectual Property Organization (www.wipo.org).

In this paper, we use the above IPC codes to analyze the product-based technologies of Apple. To find the detailed relations between the technologies, we performed the ISNM using the IPC code data. First, we obtained the SNA graphs for each of the iPod, iPhone, and iPad using the top 20 IPC codes. Figure 7 shows the SNA graph for the iPod.

Rank	iPod	iPhone	iPad
1	G06F	G06F	G06F
2	H04L	H04L	H04L
3	G09G	H04B	H04N
4	H04B	H04N	G06K
5	G06T	G06K	H04B
6	H04N	G09G	G09G
7	G06K	G06T	G06T
8	H01Q	H04M	H04J
9	H04M	H04Q	G10L
10	H03K	H04J	H01Q
11	G11C	G10L	H04Q
12	H04J	H04W	H04M
13	H04R	H01Q	H04W
14	G01R	H03K	H03K
15	H03M	G11C	G11C
16	H03F	H04R	H04R
17	H01R	H03M	G01R
18	G02F	H01R	H03M
19	H04K	G01R	G03B
20	H03G	G03B	H03F

Table II.
Top 20 IPC codes of three Apple products

IPC code	Technological definition
G01R	Measuring electric and magnetic variables
G02F	Optical operation and property
G03B	Taking, projecting, or viewing photographs
G06F	Electric digital data processing
G06K	Recognition and presentation of data
G06T	Image data processing and generation
G09G	Arrangements and circuits for control of indicating devices
G10L	Speech analysis, synthesis, recognition, or processing
G11C	Static stores
H01Q	Aerials
H01R	Electrically conductive connections
H03F	Amplifiers
H03G	Control of amplification
H03K	Pulse technique
H03M	Coding, decoding or code conversion
H04B	Transmission
H04J	Multiplex communication
H04K	Secret communication, jamming of communication
H04L	Transmission of digital information, telegraphic communication
H04M	Telephonic communication
H04N	Pictorial communication, television
H04Q	Selecting
H04R	Loudspeakers, microphones, gramophone pick-ups, deaf-aid sets
H04W	Wireless communication networks

Table III.
Technological definition of top 20 IPC codes

The technologies based on G06F, H04L, G09G, G06T, and H04B are the core technologies used to develop the iPod, because they are linked to many other nodes (IPC codes). So, sustainable development of them is required for the technological innovation of the iPod product. In addition, this figure provides diverse information about the technologies related to iPod.

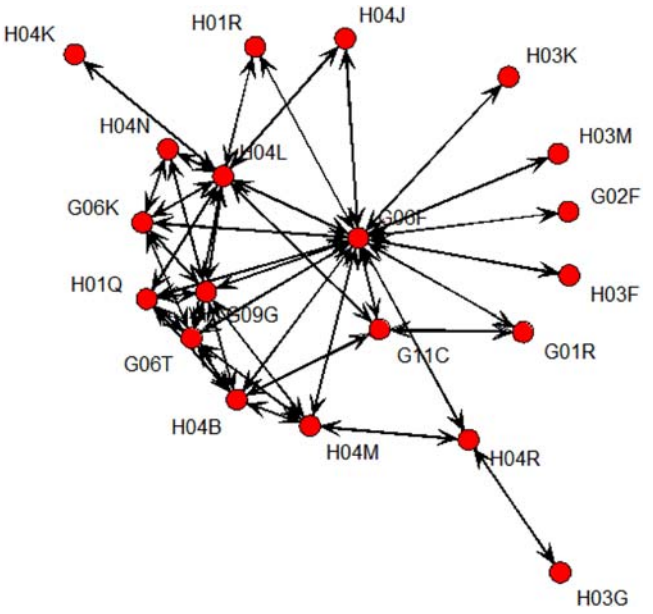

Figure 7.
SNA graph for the iPod using the top 20 IPC codes

For example, the technology of H03G is only associated with the technology of H04R, and affects the technologies of H04M or G11C via H04R. All connections of the SNA graph in Figure 7 are interpreted in this example. Figure 8 shows the SNA graph for the iPhone using the top 20 IPC codes.

Similar to the SNA graph of iPod, the technology of G06F is the core technology for the iPhone. Also, the technologies of H04N, G09G, G06K, and G06T are the candidates for the core technology. The technologies of G11C and G01R are developed together, and influence to the core technology of G06F. The next figure shows the SNA graph for the iPad.

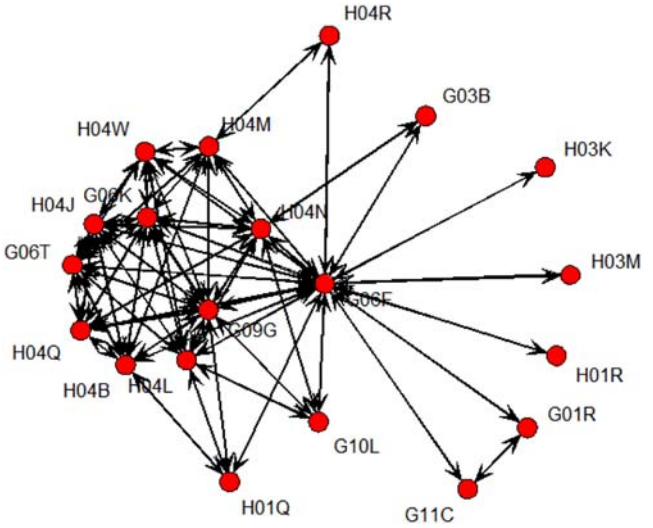

Figure 8.
SNA graph of iPhone using top 20 IPC codes

As in the case of the iPod and iPhone, G06F is central to the technology of the iPad. The technology of G06F is defined as "electric digital data processing" in Table III. So we conclude this technology is the most basic for the Apple's products of iPod, iPhone, and iPad. The technologies based on G09G and H04J are also necessary for the iPad. Besides, the technologies based on H04L, G06K, and H04B are the candidates for developing the iPad. The following table shows the SNA degrees of the top 20 IPC codes in Apple's innovative products (Table IV).

The greater the degree, the more important the IPC code is. G06F has the biggest value of degree in all products. This is the same result in Figures 7-9. To find the technological

	iPod		iPhone		iPad	
Rank	IPC	Degree	IPC	Degree	IPC	Degree
1	G06F	34	G06F	38	G06F	38
2	H04L	20	G09G	24	H04L	20
3	G09G	16	H04N	20	G09G	20
4	G06T	14	G06K	20	H04J	18
5	H04B	12	H04L	18	H04N	16
6	G06K	10	G06T	18	G06K	16
7	H01Q	10	H04J	18	H04B	16
8	H04M	10	H04B	16	G06T	14
9	H04N	8	H04M	16	H01Q	14
10	G11C	8	H04W	16	H04Q	10
11	H04R	6	H04Q	14	G10L	8
12	H04J	4	G10L	8	H04M	8
13	G01R	4	H01Q	8	H04W	8
14	H01R	4	G11C	4	G11C	4
15	H03K	2	H04R	4	H04R	4
16	H03M	2	G01R	4	G01R	4
17	H03F	2	G03B	4	G03B	4
18	G02F	2	H03K	2	H03K	2
19	H04K	2	H03M	2	H03M	2
20	H03G	2	H01R	2	H03F	2

Table IV.
Degrees of top 20 IPC codes of iPod, iPhone, and iPad

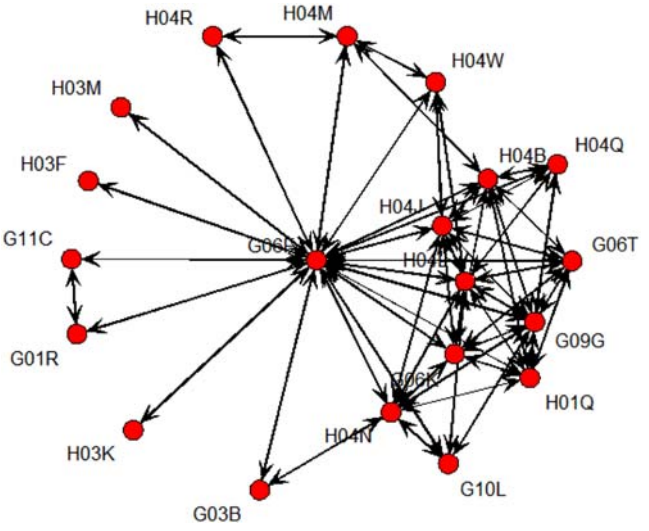

Figure 9.
SNA graph of iPad using top 20 IPC codes

associations among Apple's three products, we construct the SNA graphs using the top 10 IPC codes, based on the degrees. Figure 10 shows the SNA graph of the iPod.

We know that G09G and G06F are located at the center. In addition, the G06T technology plays a key role in developing the iPod technology. In contrast, the H03F technology is of little importance. So we conclude the technologies of G09G, G06F, and G06T are core technologies for developing the iPod product. Figure 11 shows the SNA graph of the iPhone using the top ten IPC codes.

In the case of iPhone, the IPC codes of G06F, G06K, G06T, and G09G are all core technologies. In comparison with the iPod, the iPhone needs more sub-technologies for its development.

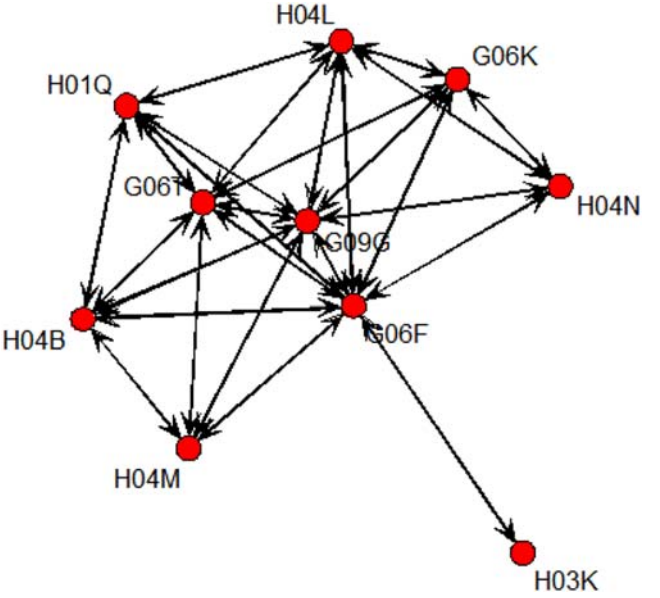

Figure 10.
SNA graph of iPod using top 10 IPC codes

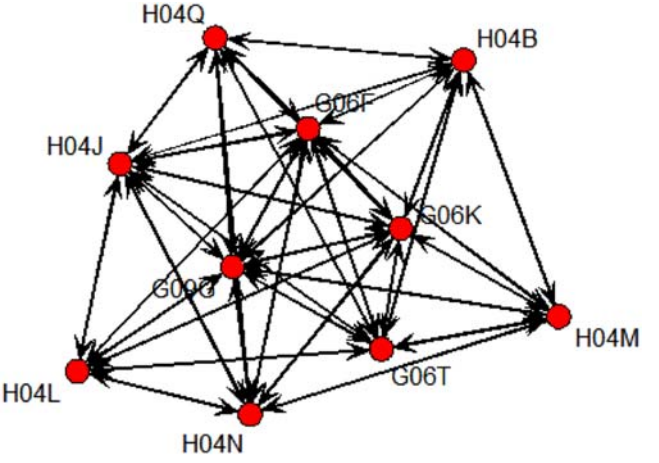

Figure 11.
SNA graph of iPhone using top 10 IPC codes

That is, we found more sub-technologies are needed than iPod for producing iPhone. Figure 12 shows the SNA graph of the iPad using the top ten IPC codes.

Unlike the cases of iPod and iPhone, most IPC codes of iPad are linked each other. That is, the development of iPad depends on a diverse technological fusion. Next, we combined the three SNA graphs in order to understand the technological relationship between Apple's products. Figure 13 represents the combined SNA graph for the iPod and iPhone.

The iPod and iPhone are connected by the technologies based on H04B and H04M. Thus, we expect the new Apple product to use the technologies of H04B and H04M. Figure 14 shows the combined SNA graph for the iPod and iPad.

The technologies based on H01Q, H04N, and H04L connect the iPod with iPad. Using these technologies, we can develop a new technology that combines the iPod and iPad. Figure 15 combines the SNA graphs of the iPhone and iPad.

We know that the connection between the iPhone and iPad depends on the technologies of H04J, H04L, and G06T. Finally, we combine the SNA graphs of all three products in Figure 16.

We find that the technology of G06K connects the three Apple products. Thus, to develop a new innovative product, Apple will need to plan an R&D strategy based on the G06K technology. We knew the technological definition of G06K is the technology of "recognition and presentation of data." Therefore, the integrated technologies related to data recognition and presentation are required for Apple's technological innovation.

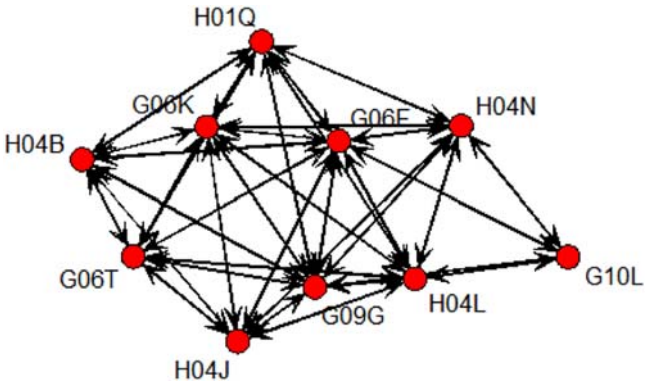

Figure 12.
SNA graph of iPad using top 10 IPC codes

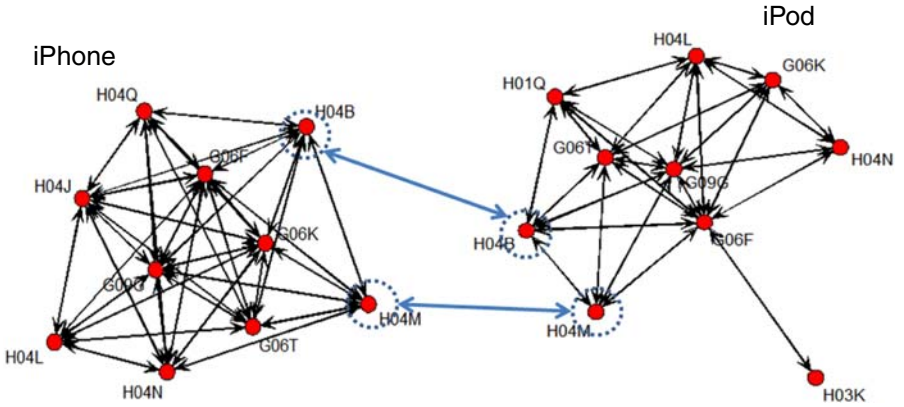

Figure 13.
Combined SNA graph for the iPod and iPhone

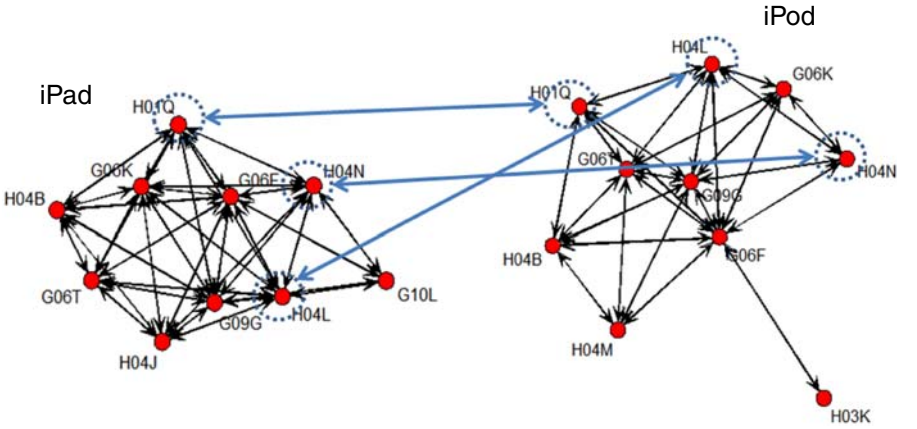

Figure 14.
Combined SNA graph with iPod and iPad

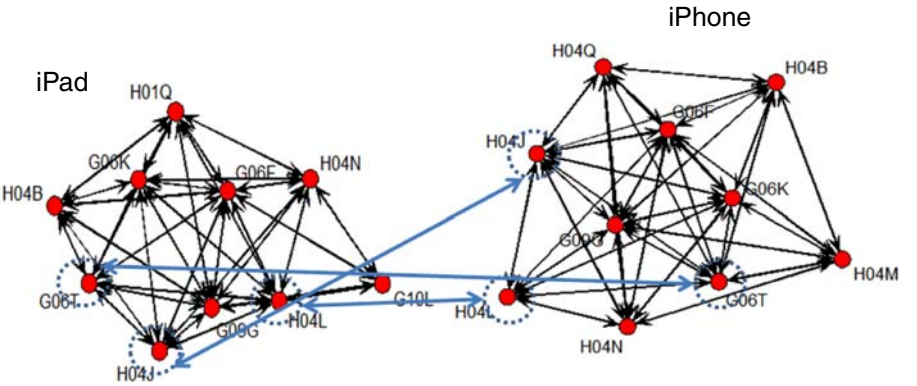

Figure 15.
Combined SNA graph of the iPhone and iPad

4. Discussion

The previous works of patent analysis took the focus to technology development in R&D planning. But the final goal of R&D planning is new product development innovatively to improve the competition of a company. To accomplish this goal, we carried out the study on a methodology of patent analysis for product-based technology. Using the results of our research, a company can build the management plan of technology before the R&D could perform. In this paper, we selected iPod, iPhone, and iPad for the case study of our research, because they were the representative product innovation and carved out new market for smart device. But except in them, Apple made other products such as iMac, MacBook, and Apple Watch. In addition, Apple has researched and developed the technologies for producing smart car. We performed the patent analysis to understand the Apple's technology based on products by two approaches. First, we analyzed the technological keywords from the patent documents of Apple. The second approach is to analyze the IPC codes included in the patent documents. According to the purpose of technology management, we can combine the results of two approaches for the detailed application.

5. Conclusions

In this study, we proposed a methodology for analyzing Apple's technology using ISNM. We selected three Apple products, namely the iPod, iPhone, and iPad, as Apple's innovative

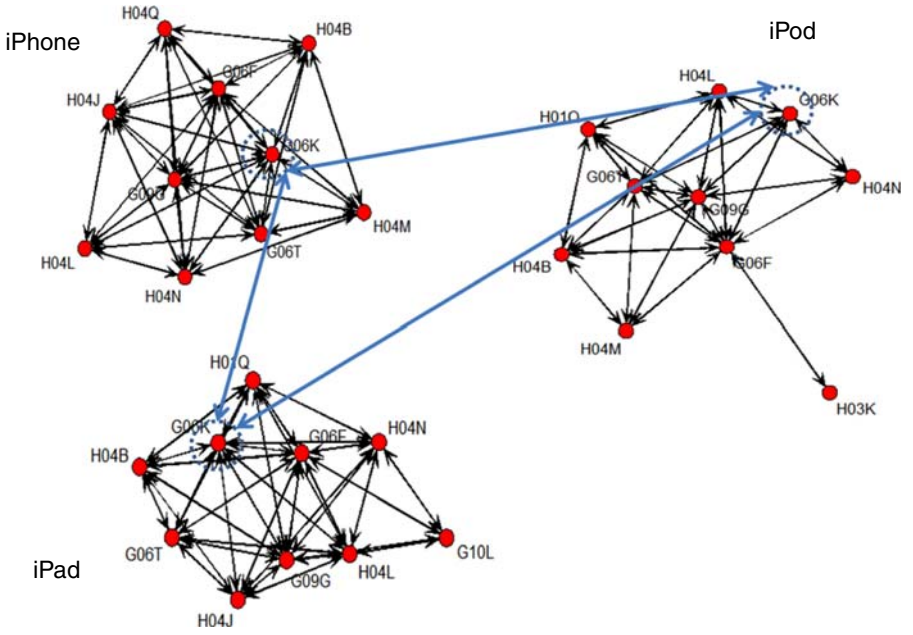

Figure 16.
Combined SNA graph for the iPad, iPhone, and iPad

technologies. We collected all patent documents applied for by Apple from various patent databases. With the help of Apple technology experts, we classified the patents according to the three products. Then, we transformed the patent documents into structured data, because the retrieved patent documents are not suitable for statistical analyses such as a SNA. In our research, we analyzed the structured patent data according to the Apple products using technological keywords and IPC codes. First, we performed the ISNM on the patent keyword data, obtaining the SNA graphs for the iPod, iPhone, and iPad. Next, we combined the three SNA graphs to determine the technological relationship between the three innovative products. Then, we used the IPC codes of Apple patents as input data to our ISNM. As in the case of the patent keyword analysis, we built the SNA graphs of each product, and then combined the SNA graphs to determine the technological connections between Apple patents. This research contributes to Apple's new product development and R&D planning. We focused on SNA visualization and the degree of Apple's technology analysis. However, there are many approaches to SNA patent technology analyses. In future work, we will consider other SNA methodologies for Apple's technology analysis. The future research direction is to combine the SNA visualization and statistical inference to illustrate the validity of technology analysis. Using this new process, we can expect new innovative products of Apple.

References

Berry, M.W. and Kogan, J. (2010), *Text Mining Applications and Theory*, Wiley, West Sussex.

Butts, C.T. (2008), "Social network analysis with SNA", *Journal of Statistical Software*, Vol. 24 No. 6, pp. 1-51.

Choi, J., Jang, D., Jun, S. and Park, S. (2015), "A predictive model of technology transfer using patent analysis", *Sustainability*, Vol. 7 No. 12, pp. 16175-16195.

Feinerer, I. and Hornik, K. (2016), "Package 'tm', Text Mining Package", CRAN of R project, Vienna.

Feinerer, I., Hornik, K. and Meyer, D. (2008), "Text mining infrastructure in R", *Journal of Statistical Software*, Vol. 25 No. 5, pp. 1-54.

Han, J., Kamber, M. and Pei, J. (2012), *Data Mining: Concepts and Techniques*, 3rd ed., Morgan Kaufmann, Waltham, MA.

Hunt, D., Nguyen, L. and Rodgers, M. (2007), *Patent Searching Tools & Techniques*, Wiley, Hoboken, NJ.

Jun, S. (2012), "Central technology forecasting using social network analysis", *Communications in Computer and Information Science*, Vol. 340, pp. 1-8.

Jun, S. and Lee, S. (2014), "A small world network for technological relationship in patent analysis", *Advances in Intelligent Systems and Computing*, Vol. 271, pp. 91-100.

Jun, S. and Park, S. (2013), "Examining technological innovation of Apple using patent analysis", *Industrial Management & Data Systems*, Vol. 113 No. 6, pp. 890-907.

Jun, S. and Park, S. (2016), "Examining technological competition between BMW and Hyundai in the Korean car market", *Technology Analysis & Strategic Management*, Vol. 28 No. 2, pp. 156-175.

Jun, S., Park, S. and Jang, D. (2012), "Technology forecasting using matrix map and patent clustering", *Industrial Management & Data Systems*, Vol. 112 No. 5, pp. 786-807.

Jun, S., Park, S. and Jang, D. (2014), "Document clustering method using dimension reduction and support vector clustering to overcome sparseness", *Expert Systems with Applications*, Vol. 41 No. 7, pp. 3204-3212.

Kim, J. and Jun, S. (2015), "Graphical causal inference and copula regression model for Apple keywords by text mining", *Advanced Engineering Informatics*, Vol. 29 No. 4, pp. 918-929.

Nettleton, D.F. and Salas, J. (2016), "A data-driven anonymization system for information rich online social network graphs", *Expert Systems with Applications*, Vol. 55, pp. 87-105.

Noh, H., Jo, Y. and Lee, S. (2015), "Keyword selection and processing strategy for applying text mining to patent analysis", *Expert Systems with Applications*, Vol. 42 No. 9, pp. 4348-4360.

Park, S., Kim, J., Jang, D., Lee, H. and Jun, S. (2016), "Methodology of technological evolution for three-dimensional printing", *Industrial Management & Data Systems*, Vol. 116 No. 1, pp. 122-146.

R Development Core Team (2016), "R: A language and environment for statistical computing", R Foundation for Statistical Computing, Vienna, ISBN 3-900051-07-0, available at: www.R-project.org (accessed May 1, 2016).

Rodriguez, A., Kim, B., Lee, J.M., Coh, B.Y. and Jeong, M.K. (2015), "Graph kernel-based measure for evaluating the influence of patents in a patent citation network", *Expert Systems with Applications*, Vol. 42 No. 3, pp. 1479-1486.

Roper, A.T., Cunningham, S.W., Porter, A.L., Mason, T.W., Rossini, F.A. and Banks, J. (2011), *Forecasting and Management of Technology*, John Wiley & Sons, Hoboken, NJ.

Scott, J.G. (2012), *Social Network Analysis*, 3rd ed., Sage, London.

Stanton, J. (2013), *An Introduction to Data Science*, Syracuse University, Syracuse, NY.

Sternitzke, C., Bartkowski, A. and Schramm, R. (2008), "Visualizing patent statistics by means of social network analysis tools", *World Patent Information*, Vol. 30 No. 2, pp. 115-131.

Tseng, Y., Lin, C. and Lin, Y. (2007), "Text mining techniques for patent analysis", *Information Processing and Management*, Vol. 43 No. 5, pp. 1216-1247.

USPTO (2016), "The United States Patent and Trademark Office", available at: www.uspto.gov (accessed April 15, 2016).

WIPSON (2016), "WIPS Corporation", available at: www.wipson.com; http://global.wipscorp.com (accessed April 10, 2017).

Corresponding author
Sangsung Park can be contacted at: hanyul@korea.ac.kr