공학박사 학위논문

Analysis of Technological Knowledge Flows in Business Model Innovation

비즈니스 모델 혁신의 기술 지식 흐름 분석

2017년 8월

서울대학교 대학원

산업공학과

안 윤 정

Abstract

Analysis of Technological Knowledge Flows in Business Model Innovation

Yoonjung An

Industrial Engineering

The Graduate School

Seoul National University

With the extensive applications of Information and Communication Technologies (ICTs) in development of new business models, the business model concept is becoming an important research topic in the field of innovation and technology management. The Internet has driven innovation and changes across the business landscape and opened the era of e-commerce. Since many of, or sometimes even entire, commercial activities thus can be conducted on the Internet platform, ICTs are becoming a key enabler in the creation of new business models. Companies are keen to leverage ICTs to develop specific methods underlying business models and seek protection for these new inventions under patent laws. These inventions, called business model patents or business method (BM) patents, refer to various commercial techniques that are usually based on digital or software-based technologies. The value of BM patents lies in the fact that these patents contain essential technological knowledge regarding business model innovation and also can facilitate business model development and innovation through knowledge flows. Despite the importance of technological knowledge flows in business model innovation, linkages between business model innovation/evolution,

technology and knowledge flows are a rather rarely explored subject. The present literature does not pay much attention to the technological basis of business model innovation, flows of knowledge in business technologies or quantitative analysis of business model innovation. To fill this research gap, the overall objective of this work is to explore technological knowledge flows in business model innovation based on patent analysis.

This study consists of three research themes. The first research theme is to develop a structured approach to measure technological knowledge flows in business model innovation. The proposed approach integrates two complementary methods, the patent citation analysis and text mining technique. The empirical study applies the proposed approach to measure knowledge flows through BM patents and reveals that BM patents actively participate in stimulating knowledge flows in business model innovation.

The second research theme is to identify patterns of technological knowledge flows in business model innovation. This study applies a dynamic approach to capture time-varying processes of knowledge flows in BM patents. A Hidden Markov Model, patent citation analysis and clustering technique are used to identify major temporal patterns of knowledge flows in BM patents.

The third research theme discusses positions or roles of BM patents regarding knowledge flows in business model innovation. This study propose a systematic framework directed at investigating different roles of BM patents that facilitate knowledge flows for innovations in social commerce. The framework mainly uses several citation-based indicators to identify core BMs and specifies their roles according to knowledge flow patterns.

This study extends overall understanding of the technological aspect of business model innovation by linking the concept of business model innovation with technological development and knowledge flows and providing systematic ways to utilize patent citation analysis and other effective techniques.

Keywords: Business model innovation, Knowledge flows, Business technology, Business method patents, Patent citation analysis
Student Number: 2013-30315

Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1. Background and motivation

With the emergence of the Internet and its massive adoption for e-commerce, companies have disrupted the traditional way of doing business, creating entirely new business models. The Internet has driven innovation and changes across the business landscape and opened the era of e-commerce. It has transformed the way companies communicate and share information with customers and deliver values to them (Boulton, et al., 2000; Damanpour and Damanpour, 2001). Since the customers' demand for online channels keeps increasing, electronic business is now an imperative, rather than an alternative. Whether started on the Web or not, many companies have come up with diverse e-commerce business models and continue to innovate their models (Damanpour and Damanpour, 2001; Weill and Woerner, 2013).

Innovation or extension of e-commerce business models today often involves technological innovation since e-commerce is based on the convergence of several major information technologies, such as computer networking and telecommunications, multimedia, information retrieval systems and electronic data interchange (EDI), and business practices (Vladimir, 1996). Also, digital technologies which enable virtualization, peer-to-peer networks, cloud computing, Internet of services and so on, are becoming a key driver in the creation of new business models (Baden-Fuller and Haefliger, 2013; Bharadwaj et al., 2013; Pagani, 2013). The biggest challenge most companies face today might be how they leverage these digital technologies to improve the way of doing business via the Internet.

The blend of technologies and business practices has resulted in the proliferation of business technologies which can be partly represented by patents. Business model patents, also known as business method (BM) patents, refer to commercial techniques that are usually based on digital or software-based technologies, such as such as computers, the internet and mobile devices (Morris, 2014). They are a special class of patents that allow companies not only to assert ownership over technologies but also to protect applications and ways of using these technologies (Morris, 2014). In recent years, the growth of the business technologies, especially in the e-commerce business industry, has been phenomenal (Wu, 2005).

The value of BM patents lies in that they can play an important role in facilitating business model development and innovation. Due to their informational content, patents are regarded as the most disembodied transfer medium of technological knowledge (Autant-Bernard et al, 2013). Such knowledge flows through patents can facilitate research and invention (Hu and Jaffe, 2003). Likewise, flows of knowledge in BM patents enable firms to obtain new software technology and business knowledge with less effort and raise the productivity of inventive activities, which can result in more efficient business model innovation. Previous studies have found that BM patents have a considerable amount of citations made to and received by other patents (Allison and Tiller, 2003; No et al., 2015; Wagner, 2008). This demonstrates that a large amount of knowledge has been exchanged through BM patents.

Also, the overall flows of knowledge in business model patents are expected to increase as these patents are constantly evolving and growing in numbers. BM patents can thus facilitate an innovation process in business model development by stimulating knowledge flows.

There are three main research streams for business model innovation, namely, corporate strategy, innovation and technology management and entrepreneurship (Wirtz et al., 2016). Research in the innovation and technology management field attempts to relate technology with business models. There are two complementary perspectives that characterize the research in this field: the first is that business models are the ways to commercialize innovative ideas and technologies; the second is that the business model represents a new subject of innovation, which complements the traditional subjects of process, product, and organizational innovation and involves new forms of cooperation and collaboration (Zott et al., 2011). In the innovation and technology management field, however, the business model is mainly seen as a mechanism that connects a firm's technology to customer needs and/or to other firm resources (Zott et al., 2011).

The linkages between business model innovation/evolution, technology and knowledge flows are a rather rarely explored subject. Most of studies that discuss the technological aspect of business model innovation mainly focus on patentability of innovative business methods involving novel applications of ICT. There are a limited number of recent studies that investigate flows of knowledge in business technologies (Chang et al., 2009; No et al., 2015). The present literature does not pay much attention to the technological basis of business model innovation, flows of knowledge in business technologies or quantitative analysis of business model innovation. Given that, we believe that it is time to improve our knowledge about business model innovation, applying the concept of technological development and patent analysis.

1.2. Research objectives

In the knowledge-based and digitized economy, technology is exerting an increasing influence on the way in which a business model can be created and adapted (Baden-Fuller and Haefliger, 2013). This is partly evidenced by the proliferation of business model patents, also known as business method (BM) patents, which cover some combinations of software and business methodology, focusing on methods, systems, or processes for conducting various aspects of e-commerce (Bagley, 2000; Wu, 2005). BM patents, as a crucial source of technological knowledge underlying business models, can greatly facilitate business model development and innovation through stimulating exchange of knowledge. Therefore, the overall objective of the present work is to explore technological knowledge flows in business model innovation based on patent analysis. Understanding business model innovation or evolution from a technological perspective will help firms to make appropriate strategic management decisions and capture opportunities for value creation through business model innovation. It will also provide important implications for formulating policy on e-commerce and related technologies.

The main research questions of this work are:

1. How can flows of technological knowledge in BM patents be measured?
2. What patterns of technological knowledge flows through BM patents can be identified?
3. How can BM patents be positioned based on their knowledge flow patterns?

The key objective of the first research question is to develop an approach that can quantitatively measure technological knowledge flows in business model innovation. The second research question aims to examine various flow patterns in business model innovation from a long-term view. The third research question focuses on identifying core business methods and classifying their roles based on knowledge flow patterns.

1.3. Scope and framework

This study links three important concepts, namely business model innovation, technology and knowledge flows, as shown in Figure 1.1. To be more specific, the scope of this study includes analysis of knowledge flows through BM patents, which can lead to faster business model innovation.

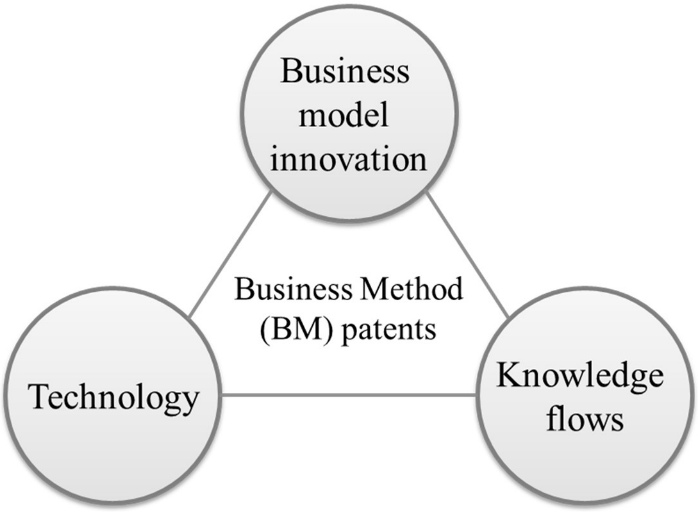


Figure 1.1 Scope of research

This study consists of three research themes that correspond to the above research questions. The scope of this study is depicted in Figure 1.2. Regarding the first research question, which is related to measurement of knowledge flows, this study attempts to develop a structured approach for measuring knowledge flows through BM patents. This approach is based on the integration of two complementary methods, patent citation analysis and text mining. The second question is related to identification of knowledge flow patterns. This study particularly focuses on identifying dynamic patterns of knowledge flows of BM patents. For methodology, a Hidden Markov Model is adopted and patent citation data is used as input data. Lastly, this study answers to the last research question by investigating roles of core BMs in social commerce from a knowledge flow perspective. Social commerce is specifically chosen for research area since it is now a dominant trend in ways of doing business via the Internet. Also, business models of social commerce are usually

IT-intensive. Core BMs are identified and their roles are defined based on citation indicators that measure different types of knowledge flows.

Research question	Theme	Topic
1. Measure	Development of a structured approach	A structured approach to explore knowledge flows through technology-based business methods
2. Pattern	Identification of dynamic patterns	Identifying dynamic knowledge flow patterns of business method patents with a Hidden Markov Model
3. Position	Investigation of knowledge transferors	Identifying core business methods in social commerce from a knowledge flow perspective

Figure 1.2 Research themes and topics

1.4. Thesis outline

The thesis consists of six chapters. The first chapter covers the background and motivation, research objectives, scope and framework and outline of the study. Chapter 2 lays the theoretical background for analyzing technological knowledge transfer in business model innovation. The main bodies of this thesis are organized according to the objectives presented in Section 1.2 (Figure 1.3). The basic research methodology used throughout the main bodies is patent citation analysis and the data source is 705 Class of the United States Patent and Trademark Office (USPTO) database. Chapter 3 develops a structured approach to measure knowledge flows through business method patents. The proposed approach integrates the patent citation analysis and text-mining technique and is applied to postage metering patents for the case study. Chapter 4 identifies major dynamic patterns of knowledge flows based on a Hidden Markov Model and clustering method, and conducts the case study using secure transaction patents. Chapter 5 investigates specific roles of core business methods in social commerce using several methods in addition to patent citation analysis, including a text-mining technique, cosine similarity measure and clustering method. Finally, Chapter 6 summarizes the conclusions and contributions of the study, and discusses its limitations and further research suggestions.

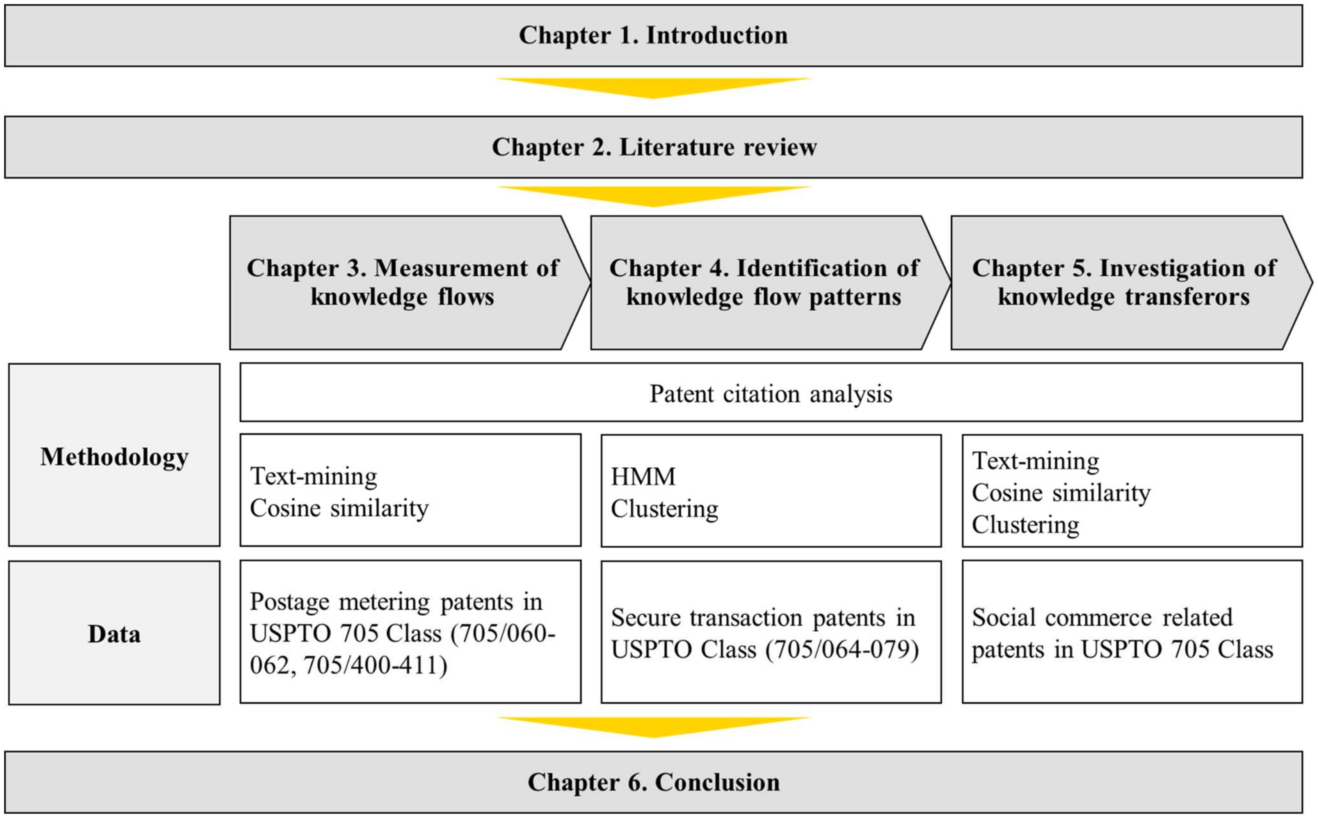


Figure 1.3 Thesis outline

Chapter 2. Literature Review

2.1. Business model innovation and technology management

Business model innovation is receiving increased attention in academia and industry alike. Business model innovation has established itself as a cornerstone of innovation – next to product, service, and process innovation (Wirtz et al., 2016). It has gained its importance in the recent past, especially since it can be an alternative or complement to product or process innovation which is often expensive and time-consuming (Amit and Zott, 2012). Companies can gain sustainable competitive advantage through successful implementation of business model innovation. Despite these high levels of interest and attention that have recently been paid to business model innovation, the extant literature draws a quite heterogeneous picture, which lacks conceptual clarity and consistency. Although business model innovation can be defined in numerous ways depending on the perspective or field of research, the comprehensive definition given by Wirtz (2016) can help understanding the concept: "Business model innovation describes the design process for giving birth to a fairly new business model on the market, which is accompanied by an adjustment of the value proposition and/or the value constellation and aims at generating or securing a sustainable competitive advantage."

There are three main research streams for business model innovation, namely, corporate strategy, innovation and technology management and entrepreneurship (Wirtz et al., 2016). Figure 2.1, adapted from Wirtz et al. (2016), presents an overview of major publications in the different research streams over time. The viewpoint of corporate strategy considers business model innovation as either the reformulation of incumbent firms' corporate strategy or the novel creation of new market entrants' strategy. This connection to corporate strategy has emerged from the notion that a business model is the direct result of strategy. In the viewpoint of innovation and technology management, business model innovation is related to addressing operational aspects such as processes, linkages or structures. Studies taking the viewpoint of entrepreneurship focus on explaining how an existing or future company or business stream is to generate profit. However, this perspective has so far been lacking sufficient treatment when compared to the other two currents in the literature (Spieth, 2014; Wirtz et al., 2016).

Research in the innovation and technology management field attempts to relate technology with business models. There are two complementary perspectives that characterize the research in this field: the first is that a business model is the way to commercialize innovative ideas and technologies; the second is that the business model represents a new subject of innovation, which complements the traditional subjects of process, product, and organizational innovation and involves new forms of cooperation and collaboration (Zott et al., 2011). In the innovation and technology management field, however, the business model is mainly seen as a mechanism that connects a firm's technology to customer needs and/or to other firm resources (e.g., technologies) (Zott et al., 2011).

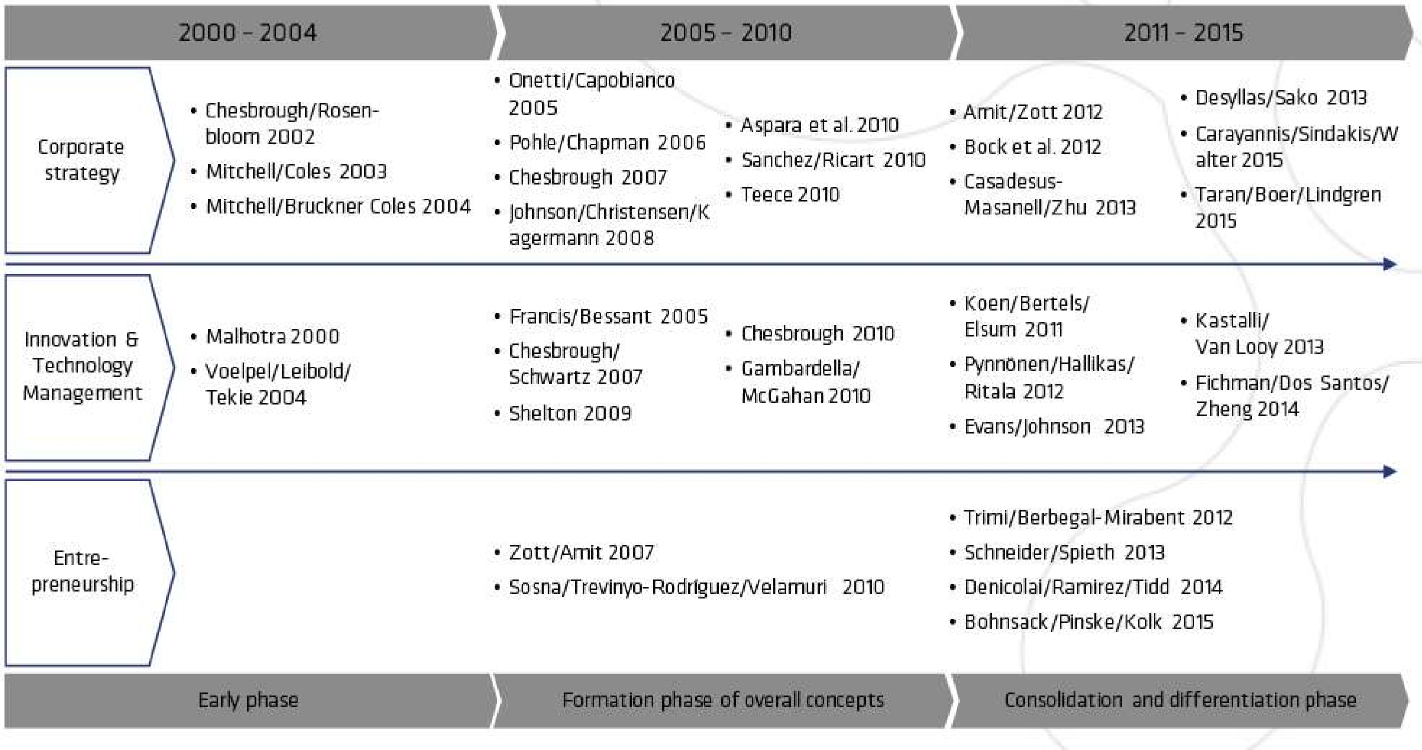


Figure 2.1 Main research streams for business model innovation and relevant publications

2.2. Knowledge flows

Knowledge flows are beneficial not only to individual economic entities that actively participate in exchanging knowledge but also to society as a whole. Economic entities can strengthen their innovative capabilities through obtaining new knowledge developed by other entities at costs generally lower than the original costs of developing it on their own (Verspagen and De Loo, 1999; Wang et al., 2012). They also can substantially increase the productivity of knowledge development and knowledge pool by exploiting other entities' knowledge as new ideas or data for research projects (Verspagen and De Loo, 1999). In addition, knowledge flows contribute to the increase of their market value, competitiveness and economic efficiency as knowledge is one of the core assets in the knowledge and technology driven economy (Liu et al., 2015; Plasmans and Lukach, 2010). All these benefits will, in turn, lead to faster technological advancement and economic growth of the society (Plasmans and Lukach, 2010; Verbeek et al., 2003).

There are several definitions for knowledge spillovers or knowledge flows, yet the fundamental concepts are the same – exchange or diffusion of ideas, knowledge, concepts, etc. between economic agents, which will help developing and extending their internal knowledge stock. Knowledge flows can take place through different means, such as purchase of capital goods with embodied technologies, publications, patents, conferences, networking, education and training and labor mobility (Dumont and Tsakanikas, 2001; Karlsson and Gråsjö, 2014).

The previous literature, especially in the field of economics, has used a number of distinct approaches to explore knowledge flows. Among multiple proxies, including R&D capital stocks/expenditures, international trade, human mobility, purchase of machinery, equipment and components and so on, patents have been exploited as a representative indicator of knowledge flows (Macdissi and Negassi, 2002). Patents are a valuable data source as they provide reliable and comprehensive technical information organized in a standard format. Patent citation information is particularly useful for measuring technology diffusion and knowledge flows. Citation frequencies can be a proxy for the amount of knowledge transferred from antecedents to descendants. Many studies have built "technology flow matrices" or statistical models, such as a logit, probit and Weibull model, based on patent data to quantify knowledge flows (Duguet and MacGarvie, 2005; Plasmans and Lukach, 2010; Verspagen and De Loo, 1999). Also, recent research has attempted to employ a more innovative approach, a text-mining technique, to identifying patterns of knowledge flows (No et al., 2015).

Unfortunately, research on knowledge flows is still limited in the aspects of methodology and research field. Although patent citations are used as a proxy for knowledge flows, it has some drawbacks. Patent citations may underestimate the actual extent of knowledge flows (Lukach and Plasmans, 2005) or may be biased by incorrect citing of sources (Duguet and MacGarvie, 2005). Also, most studies focus on static analysis of knowledge flows while knowledge flows are an intrinsically dynamic phenomenon. Although some models have been constructed based on the Bass diffusion model to generate dynamic diffusion patterns of knowledge (Kreng and Tsai, 2003; Tsai, 2008), they have a critical limit in applicability since the principle of the Bass model presupposes that diffusion patterns would follow an S-curve. Furthermore, the scope of research field revolves around high-tech industries including semiconductors and fuel cells, and not enough scholarly attention has been devoted to study knowledge flows with regard to business model innovation.

2.3. Business method (BM) patents and patent citation analysis

Business method (BM) patents are a valuable information source to enhance understanding of innovations in e-commerce. Although a new business model is itself unlikely to qualify for formal IP protection, specific BMs underlying it are fully patentable subject matters (Desyllas & Sako, 2013). BM patents generally cover some combinations of software and business methodology for conducting various aspects of e-commerce (Wu, 2005). These patents may also contain apparatus and article of manufacture claims implicating the computer environment in which the model operates (Bagley, 2000). In recent years, the volume of BM patents has grown at an ever-increasing rate (Chang, 2012) and consequently, research and analysis on BM patents is becoming crucial from both technological and business perspectives. Both academia and industry can analyze BM patents for various purposes, such as identifying promising BMs, analyzing and forecasting technological trends/developments, strategic technology and business planning, and identifying technological positions.

Investigating knowledge flows in BM patents is helpful for determining innovations that play a vital role in technological progress of e-commerce. Whether a certain technology has a critical impact on the evolution process depends on whether technological knowledge is useful for new developments and applications (Ho et al., 2014). Patent citation data is commonly used in empirical studies to measure knowledge flows and innovation performance (Guan and Chen, 2012). Patent documents contain a list of backward and forward citations which indicate technological antecedents and decedents of the particular innovation (Ko et al., 2014). Backward citations are citations made to prior patents and represent technological knowledge acquired by the inventor; forward citations are citations received by other patents and can be interpreted as the diffusion of knowledge encapsulated in a certain patent (Plasmans & Lukach, 2010). Also, forward citations are frequently used as a proxy for patent value or importance (Duguet & MacGarvie, 2005).

Although there are a large volume of research on knowledge flows using patents, it is difficult to find research discussing about knowledge flows of BM patents. Chang et al. (2009) analyzed technology diffusion within BM patents based on a citation network and clustered patents based on their technology diffusion patterns. No et al. (2015) identified BMs that facilitate knowledge flows with an approach integrating patent citation analysis and text-mining.

Chapter 3. Measurement of Knowledge Flows

3.1. Introduction

Business environment has been changing rapidly since the 1990s and companies have to constantly develop new business methods or business models to keep up with such changes and survive in the market. The most noticeable change among all is the substantial application of ICT in the business environment; ICT has become a new enabler for business communication and processing commercial transactions (Chang, 2012; Mounteney, 2002; National Science Board (US), 2002; Osterwalder, 2004; Pateli, and Giaglis, 2005; Wu, 2005). With ICT as a critical part of new BMs, BMs are now one of the patentable subject matters that gain special attention, and a growing number of companies try to seek patent protection for the new BMs (National Science Board (US), 2002). As a result of that, the number of BM patents grew rapidly over a very short period (Rausch, 2003; Wu, 2005).

Despite their importance in the business environment, it is difficult to find previous studies discussing about technology-based BMs with respect to their relationships with technologies, technology-based BMs of other kinds or knowledge flow. Recently, some researchers made attempts to study such topics. Kim et al. (2011) identified between technology-based services, which are represented by the BMs, and ICTs. Some studies aimed to explore technology diffusion of BMs (Chang, 2012) and identify internal technological relationships among the BM patents and patterns of business model evolution

(Lee et al., 2013). Other studies merely focused on explaining the theoretical background of impact of technology on BMs/business model innovation (Mounteney, 2002; National Science Board (US), 2002; Osterwalder, 2004; Wu, 2005).

There are different terminologies used for BMs that include ICTs, but there is no specific terminology or definition that is widely agreed upon among researchers or practitioners. Hunt (2001) called such BMs "computer-implemented business methods." Wu (2005) used two different terms which are "software-embodied business methods" and "internet business methods." Those BMs are also simply called "business methods" (Wagner, 2008) and "business methods based on Internet technologies" (Chang, 2012). In this paper, they will be called "technology-based business methods" and narrowly defined as "type of business methods limited to patentable subject matter classified in USPTO Class 705 which only includes business methods based on technologies."

The role of knowledge exchange is especially important in a knowledge and technology driven economy because it allows better penetration and diffusion of innovation and stimulates cooperation in R&D (Hu and Jaffe, 2003; Lukatch and Plasmans, 2002). There have been extensive studies emphasizing the importance of knowledge flow/spillover. Glaeser et al. (1991) suggest that knowledge flow/ spillover is directly linked to three factors of economic growth, which are specialization, competition and diversity, and they are characterized by a higher intensity of intra-industry knowledge spillover, inter-firm innovation flows and inter-industry knowledge exchange, respectively. Also, Huggins and Johnston (2010) argue that knowledge exchange through networking with various partners in different domains can open opportunities for novel combination and recombination of ideas or best-of-breed solutions that originate from different resource bases and knowledge bases. Such knowledge networks are thus an important aspect of the innovation process (Huggins and Johnston, 2010; Meagher and Rogers, 2004; Sammarra and Biggiero, 2008).

Although it is not difficult to conceptualize a phenomenon of knowledge flow, it is very difficult to measure the degree of knowledge flow (Lukatch and Plasmans, 2002). The two main methods are direct and indirect (Crespi et al., 2008). The main direct method is to use information in patent citations. The indirect method of measuring knowledge flows typically regresses total factor productivity (TFP) growth on factors thought to be potentially causing information flows, such as the presence of multinational enterprises (MNEs) or international-trade status (Crespi et al., 2008). They both have advantages and disadvantages, but we decide to use patent citation as a measure of knowledge flow due to the following reasons: patent citation is a certified evidence of previous knowledge used by the inventor (Nelson, 2009), data can be obtained easily and International Patent Classification (IPC) corresponds with the purpose of our study.

Since BMs have been disclosed to the public in the form of a patent, meaning they are more exposed to knowledge flows, it is worth studying important implications of the BM patents in terms of knowledge flow. The BM patents enable effective measurement of knowledge flow of BMs, with citation and other information. Knowledge flow is stimulated by the BM patents through active cited (backward citation) and citing (forward citation) patents. There exist some previous studies showing the empirical evidence that the BM patents not only cite a significant number of previous patents but are also cited by a substantial number of subsequent patents (Allison and Tiller, 2003; Wagner, 2008).

Both cited and citing patents represent knowledge flow in a similar manner, but the underlying economic rationales of these two processes differ (Duguet and MacGarvie, 2005; Plasmans and Lukach, 2010). Cited patents (backward citations) have been used to measure technological knowledge acquired by the patenting entity and thus regarded as knowledge utilization; on the other hand, citing patents (forward citations) have been interpreted as a measure of the knowledge diffusing outward from the patenting entity and thus regarded as knowledge dissemination (Duguet and MacGarvie, 2005; Plasmans and Lukach, 2010). The more frequently a patent is cited by patenting entities, the greater the related technology may have influenced, implying that the technological knowledge is more widely disseminated. Since the BM patents have a substantial number of both cited and citing patents, it can be interpreted that they play an important role in utilizing and disseminating knowledge.

Although there are numerous studies using patent citation information as a proxy for knowledge flow between technologies or actors, and patent citations encapsulate important information about knowledge flow, there are still some drawbacks to use citation information. Patent citations, which are linked to the patenting procedure itself, capture only the knowledge flows, thus underestimating the actual extent of knowledge flows (Lukach and Plasmans, 2005). Also, they could be biased by incorrect citing of sources; thus supplementary investigation is required to allow citation information to be confidently applied (Duguet and MacGarvie, 2005).

In order to overcome the drawback of citation based approach, text mining, using textual data to discover useful pattern, can be applied along with citation analysis. Co-word analysis is mainly utilized to explore the concept network in different fields since the nature of words, on which co-word analysis is based, can act as the important carrier of knowledge (Van Raan and Tijssen, 1993). Words and co-occurrences of words cover a much broader domain than citations (Leydesdorff, 1987).Words occur not only as indicators of links among documents but also internally within documents. Thus, the text data can be used to measure a degree/amount of knowledge transferred by measuring text similarities between patents while patent citation is used to measure a path of knowledge flows.

As an attempt to provide a deeper understanding of technology-based BMs with regard to knowledge flow, the paper proposes a framework for exploring knowledge flows driven by technology-based BMs from their utilized technologies to disseminated technologies, by investigating both cited and citing patents. The proposed approach integrates patent citation analysis and text mining to explore the knowledge flow through technology-based BMs. First, knowledge flow path is traced using citation links between technology-based BM patents and their cited patents, which represent utilized knowledge sources, and between technology-based BM patents and their citing patents which represent disseminated knowledge. Then, co-word analysis as a text mining is integrated with the citation analysis to verify the degree of knowledge transferred between BM patents and cited/citing patents by measuring the text similarity between BM patents and their utilized/disseminated knowledge source. The integrated approach will lead to a better measurement of knowledge flow in terms of the degree of knowledge flow.

3.2. Proposed approach: integrating patent citation analysis and text mining

3.2.1. Overall research process

In addition to the fact that patent information is better protected from data disruption than other database, citation information provides citation links which can be used to analyze technological diffusion, valuation or impact, among various patents.

In many studies (Choi and Park, 2009; Choi and Park, 2009; Von Wartburg et al., 2005; Wang et al., 2013), patent citation information has been frequently used to construct the knowledge flow matrices for measuring knowledge flows. In this paper, patent citation and text data are integrated to classify technology-based BM patents as knowledge flow drivers, and to measure the degree of knowledge flow driven by technology-based BMs.

The proposed approach consists of four steps as shown in Figure 3.1. First, technology-based BM patents, which are base patents in the research, and their cited/citing patents are collected from the USPTO database. In the research, the base patents are the subject of analysis, which act as a mediator of knowledge flow through cited and citing patents. For the base patents, all the patents belonging to research-related patent class (USPTO classes, in our case) are collected. Second, keywords are extracted from the abstracts of base patents and citing/cited patents to measure the degree of knowledge flow. The lists of descriptors are standardized to delete a variant of the same word. For the third step, after constructing the keyword matrix of base patents, and of cited/citing patents, textual similarities between the base BM patents and the cited/citing patents which are clustered as patent classes (i.e. USPTO classes or subclasses) are computed to measure what degree of knowledge is actually exchanged between them. Cosine similarity is used as a similarity measure since it is easy to interpret and simple to compute for long and sparse vectors (Baeza-Yates and Frakes, 1992; Han et al., 2011; Salton and McGill, 1983). Lastly, patterns of knowledge flows driven by technology-based BMs are identified in terms of patent classes. In other words, the patent classes are categorized into different groups based on the degree of knowledge the base BM patents utilized or disseminated.

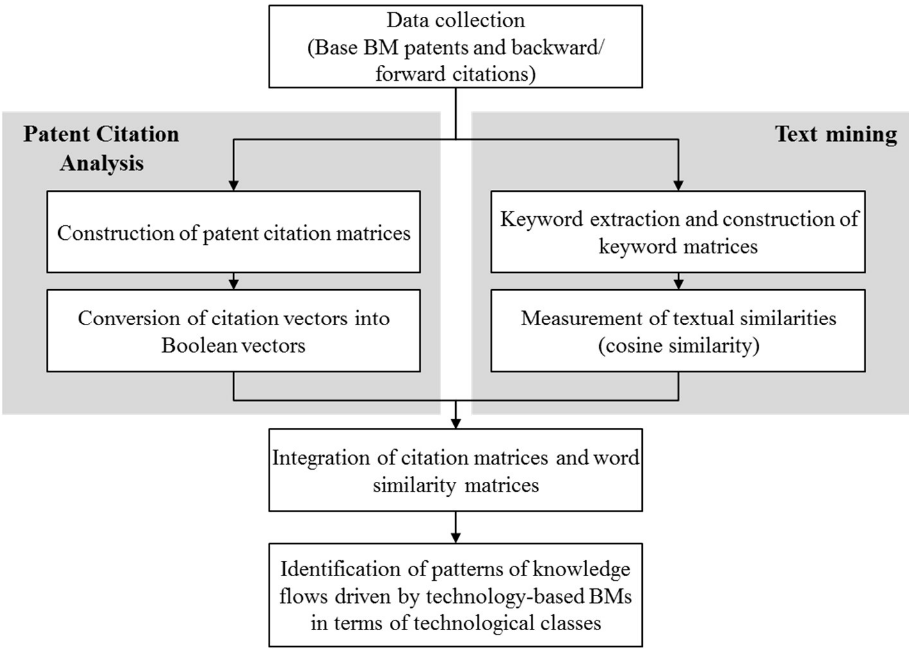


Figure 3.1 Overall process for exploring knowledge flows driven by technology-based BMs

3.2.2. Integrated approach by combining patent citation analysis and text mining

The idea of studying the text information combining with bibliometric methods and vice versa is not new to the literature. There are several studies combining bibliometric information with textual content to obtain improved performance in clustering (Janssens, 2007), classification (Calado et al., 2006) and bibliometric mapping (Janssens et al., 2006). However, most studies focus on combining co-citation with word analysis in the context of evaluative bibliometrics in order to improve efficiency of co-citation clustering and bibliometric mapping (Glenisson et al., 2005). Since text-based approach usually is based on rather rich vocabularies and peculiarities of natural language, the relationship between documents is somewhat fuzzy and not always reliable. On the other hand, if strict citation-based criteria are applied, that is, if non-periodical references and occasional coupling links are removed, the resulting citations-by-document matrix becomes extremely sparse. Combining two techniques helps to improve the reliability of relationship and the clustering algorithm as well (Janssens et al., 2008).

The text information and citation information are combined and configured, as shown in Figure 3.2, to measure the similarity among base BM patents and their cited/citing patents in order to explore the knowledge flows through technology-based business methods.

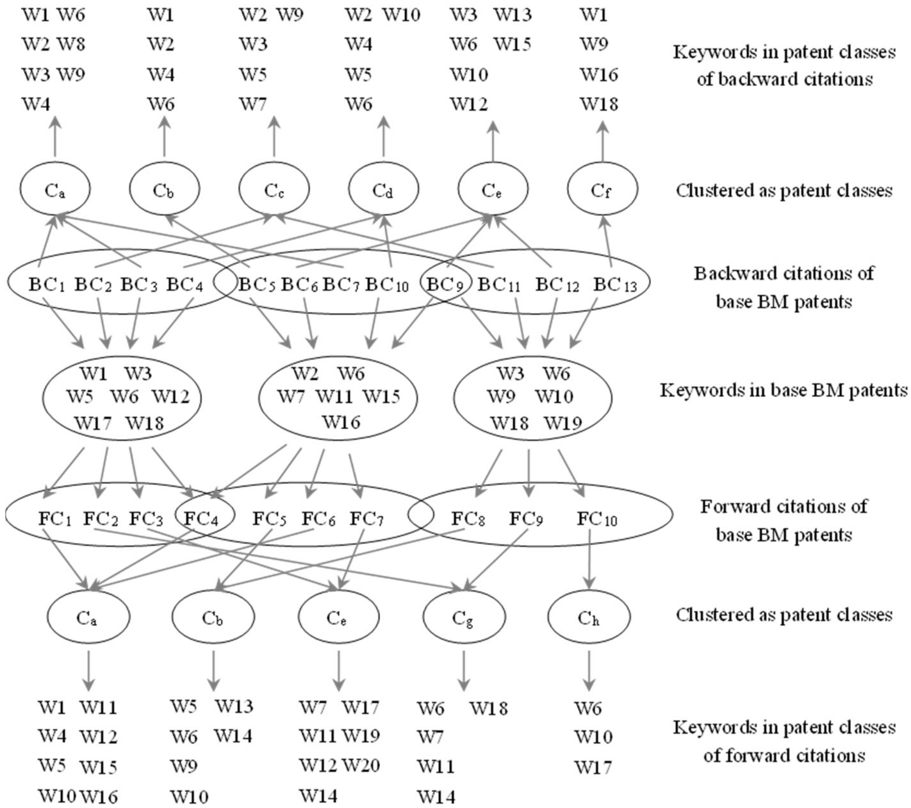


Figure 3.2 Citation analysis and text mining in integrated approach

3.2.3. Similarity measure

The issue of document similarity attained more attention with increasing interest in information retrieval. It is a measure used to compare two objects and to determine if they are related to the same topics (Zhang and Rasmussen, 2001). Similarity measures yield an indication of the relevance of an object, a document, to a given standard, the query. According to McGill et al. (1979), there are more than 60 different similarity measures; dice coefficient, cosine coefficient, overlap coefficient (Salton and McGill, 1983), and the spreading activation similarity measure. Each similarity measure has its strengths and weaknesses in practice.

The distance-based similarity measure and the angle-based cosine measure are the most popular measures. However, the distance-based similarity measure takes only the impact of the distance into account. Thus, documents with the same distance to the reference point shall have the same similarity regardless of the direction of the document (Zhang and Korfhage, 1999). The cosine measure can effectively identify documents in a vector document space that have the same indexing term distribution within each document (Trajtenberg et al., 1997). That is, the same proportion of weights is given for any pair of indexing terms between two documents if they have the same indexing terms. The most widely used measure is still the cosine similarity in the vector space model and the cosine similarity is easy to interpret and simple to compute for long and sparse vectors (Baeza-Yates and Frakes, 1992; Salton and McGill, 1983; Priego, 2003).

In the paper, keyword-based similarity values are measured, from the 'Patents by keywords' matrix, using the concept of the cosine measure of 'Salton &McGill' (1983) which is defined as the cosine of the angle enclosed between two vectors x and y. The cosine measure of 'Salton & McGill' has an advantage over the Pearson correlation in that the similarity is insensitive to the number of zeros as the cosine is not based on the mean of the distribution (Priego, 2003).

In this research, the similarity between a base BM patent and backward/forward citation, at patent-class level, is defined as

$$Sim(P,C) = \frac{\sum_{K=1}^{n} P_k C_k}{\sqrt{\sum_{K=1}^{n} {P_k}^2}\sqrt{\sum_{K=1}^{n} {C_k}^2}}$$

where P_k is the frequency of keyword k in base patent P and C_K is the frequency of keyword k in backward/forward citations at patent-class level. The boolean case of $P_k = 0$ when there is no keyword existing in corresponding patent, or $P_k = 1$ when there is keyword. The boolean case of $C_k = 0$ when there is no keyword existing in corresponding citation patent, or $C_k = 1$ when there is keyword. n is the total number of patent-class for base patents and citation patents. Since the cosine formula normalizes for the length of the word-profiles of both object (base patent) and query (citation patents), objects with long word-profiles can be penalized for their 'representational richness' if this does not correspond to the richness in the query's representation (Jones and Furnas, 1987).

The keyword-based matrices for base patents and for cited/citing patents at patent-class level are constructed, then the keyword-based similarity matrices for "base patents by cited/citing patents" are constructed based on the value of the keyword-based similarity. As seen in Figure 3.3, there are similarity values if there is a direct citation relationship and 0 otherwise, by using boolean vectors of 'base patent by cited/citing patents' as a weight.

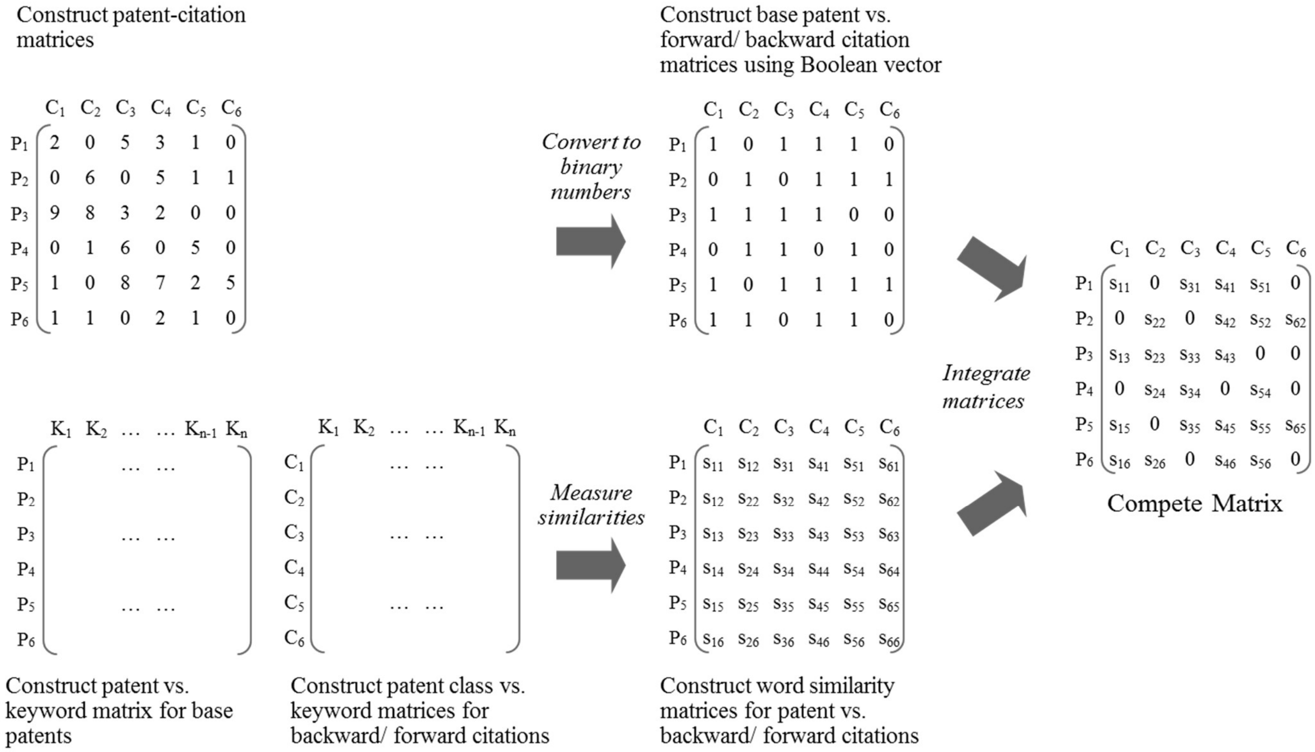


Figure 3.3 Integration of patent citation and text mining

3.3. Case study: postage metering system

3.3.1. Data collection

As explained in the research framework, all the data is retrieved from the USPTO database. The postage metering system related BM patents are selected as the base BM patents. Under the USPTO classification scheme, Class 705 is further divided into subclasses based on subjects, such as operations research, POS terminal or electronic cash register, electronic shopping and finance. Among the various subjects, postage metering system (060–062, 400–411) is selected as the base BM patents for our empirical study. The postage metering system BM patents belong to two different subclasses: cryptography and cost/price. Subclasses 060–062, belonging to cryptography, are defined as the subject matter wherein a charge for mailing an article is determined, markings representing this charge are affixed to the article, and respective modifications to an account balance are made. Subclasses 400–411, belonging to cost/price, are defined as the subject matter wherein the data processing or calculating computer comprises means for determining and printing cost required for mailing an article. In terms of a time span for building the dataset, the patents issued from 2004 to 2009 are specifically chosen because there needs to be a certain period of time for the base patents to have a sufficient number of both cited and citing patents. The number of base patents is 149 and that of cited and citing patents is 3,494 and 643, respectively. The detailed information about the base patents and citations are listed in Tables 3.1 and 3.2.

Table 3.1 Base patents (postage metering system BMs) and backward/ forward citations

Base/ citation	Number of patents	Descriptions
Base Patents	149	Issued from 2004 to 2009 - need a certain period of time for both backward and forward citations
Backward Citations	3494	Eliminated those have not been issued and those do not include abstract
Forward Citations	643	Eliminated those do not include abstract

Table 3.2 Subclasses of base patents

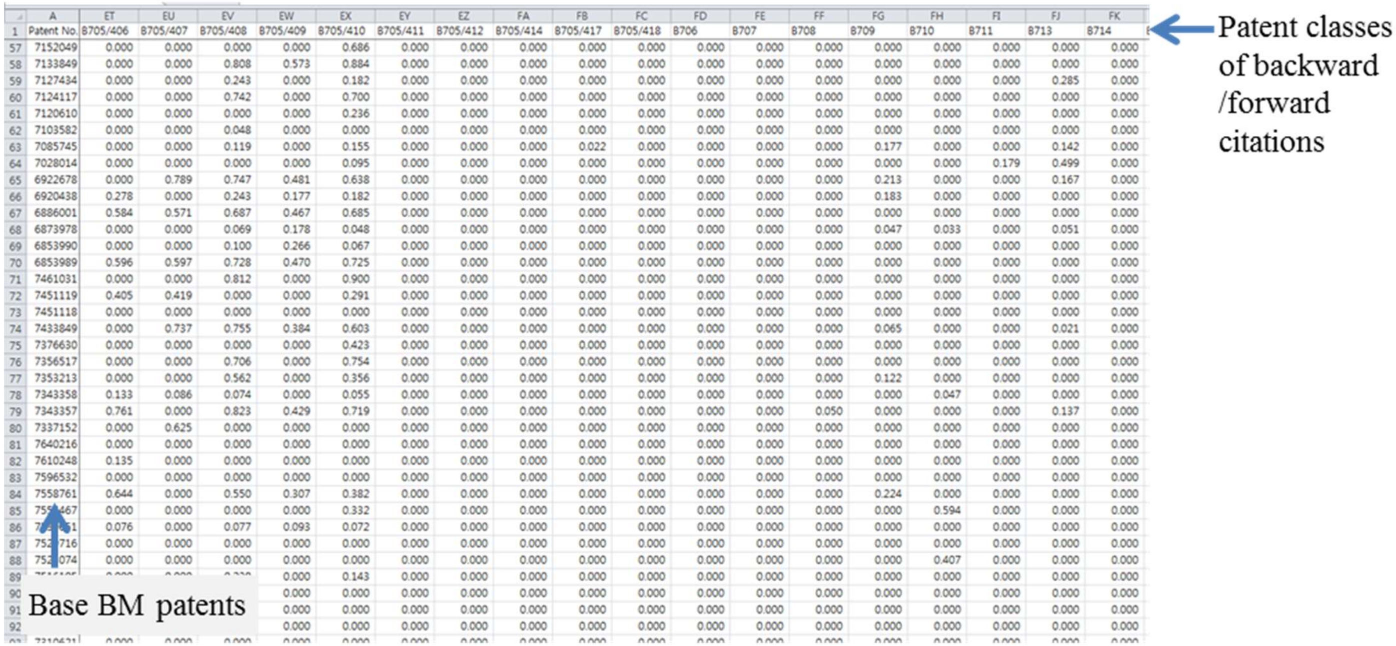
Subclass (14)	Title	No. of base patents
705/060	Postage metering system (Cryptography)	28
705/061	Reloading/recharging	3
705/062	Having printing detail (e.g., verification of mark)	16
705/401	Postage metering system (Cost/Price)	41
705/402	Special service or fee (e.g., discount, surcharge, adjustment, etc.)	11
705/403	Recharging	2
705/404	Record keeping	7
705/405	Data protection	1
705/406	With specific mail handling means	3
705/407	Including mailed item weight	8
705/408	Specific printing	20
705/409	Rate updating	1
705/410	Specialized function performed	7
705/411	Display controlling	1

3.3.2. Construction and integration of matrices

The dataset for base BM patents includes backward and citations and their text information as well. It is extended by the following operations:

1) Build the dataset for cited patents (backward citations) of base BM patents
2) Build the dataset for citing patents (forward citations) of base BM patents
3) Construct the Boolean matrices of "base patent vs. cited/ citing patents"
4) Build the dataset for keywords of base BM patents and cited/citing patents
5) Construct the keyword matrix of "base BM patents vs. keywords" and "patent class vs. keywords" for cited/citing patents
6) Measure similarities based on the keyword matrix of "base BM patents vs. keywords" and "patent class vs. keywords" for cited/citing patents
7) Integrate the Boolean matrices of "base BM patents vs. cited/citing patents" and keyword similarity matrices of "base BM patents vs. cited/citing patents" (seen in Figure 3.4).

The keywords are extracted from the abstracts, which contain the essential information of the patents, using the text mining package, "TextAnalyst 2.1." A total of 989 keywords are extracted; however, it should be noted that the software does not understand the context or meaning of words so that there may be some irrelevant or insignificant words included in the result. Those insignificant words are then manually screened out and a total of 490 keywords are selected as a final set. After mining keywords, the resulting lists of descriptors were standardized to eliminate different spellings and variants of the same terms. In contrast to the conventional co-word analysis which generates a symmetrical matrix with an empty diagonal, matrix of patents vs. keywords are asymmetrical.



	A	ET	EU	EV	EW	EX	EY	EZ	FA	FB	FC	FD	FE	FF	FG	FH	FI	FJ	FK
1	Patent No.	B705/406	B705/407	B705/408	B705/409	B705/410	B705/411	B705/412	B705/414	B705/417	B705/418	B706	B707	B708	B709	B710	B711	B713	B714
57	7152049	0.000	0.000	0.000	0.000	0.686	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
58	7133849	0.000	0.000	0.808	0.573	0.884	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
59	7127434	0.000	0.000	0.243	0.000	0.182	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.285	0.000
60	7124117	0.000	0.000	0.742	0.000	0.700	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
61	7120610	0.000	0.000	0.000	0.000	0.236	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
62	7103582	0.000	0.000	0.048	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
63	7085745	0.000	0.000	0.119	0.000	0.155	0.000	0.000	0.000	0.022	0.000	0.000	0.000	0.000	0.177	0.000	0.000	0.142	0.000
64	7028014	0.000	0.000	0.000	0.000	0.095	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.179	0.499	0.000
65	6922678	0.000	0.789	0.747	0.481	0.638	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.213	0.000	0.000	0.167	0.000
66	6920438	0.278	0.000	0.243	0.177	0.182	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.183	0.000	0.000	0.000	0.000
67	6886001	0.584	0.571	0.687	0.467	0.685	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
68	6873978	0.000	0.000	0.069	0.178	0.048	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.047	0.033	0.000	0.051	0.000
69	6853990	0.000	0.000	0.100	0.266	0.067	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
70	6853989	0.596	0.597	0.728	0.470	0.725	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
71	7461031	0.000	0.000	0.812	0.000	0.900	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
72	7451119	0.405	0.419	0.000	0.000	0.291	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
73	7451118	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
74	7433849	0.000	0.737	0.755	0.384	0.603	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.065	0.000	0.000	0.021	0.000
75	7376630	0.000	0.000	0.000	0.000	0.423	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
76	7356517	0.000	0.000	0.706	0.000	0.754	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
77	7353213	0.000	0.000	0.562	0.000	0.356	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.122	0.000	0.000	0.000	0.000
78	7343358	0.133	0.086	0.074	0.000	0.055	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.047	0.000	0.000	0.000
79	7343357	0.761	0.000	0.823	0.429	0.719	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.050	0.000	0.000	0.000	0.137	0.000
80	7337152	0.000	0.625	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
81	7640216	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
82	7610248	0.135	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
83	7596532	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
84	7558761	0.644	0.000	0.550	0.307	0.382	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.224	0.000	0.000	0.000	0.000
85	[illegible]	0.000	0.000	0.000	0.000	0.332	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.594	0.000	0.000	0.000
86	[illegible]	0.076	0.000	0.077	0.093	0.072	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
87	[illegible]	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
88	[illegible]	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.407	0.000	0.000	0.000
89	[illegible]	[illegible]	[illegible]	[illegible]	0.000	0.143	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
90					0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
91					0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
92					0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Figure 3.4 Integrated matrix of “Boolean citation matrix” and “Word-similarity matrix”

3.3.3. Patterns of knowledge flow

The technological classes are positioned on the two dimensional map based on the amount of knowledge exchanged with the base BM patents, as shown in Figure 3.5. Taking the mean value as a criterion, each class can be classified as either high (above mean) or low (below mean) in the knowledge utilization (KU) dimension or knowledge dissemination (KD) dimension. The classes are then categorized into three groups, depending on the amount of knowledge flow: high KU–high KD, high KU–low KD and low KU–high KD. The detailed characteristics of knowledge flow patterns are described in the following Table 3.3.

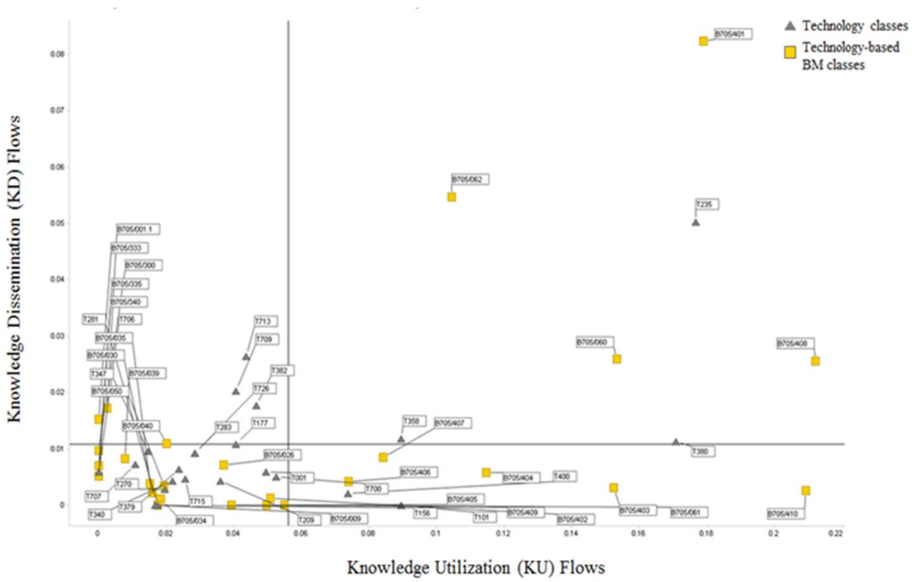


Figure 3.5 Positioning of technological classes with relation to knowledge flow driven by technology-based BMs

Table 3.3 Patterns of knowledge flows through technology-based BMs in terms of technological classes

Knowledge flow pattern	Description	Affiliated technological classes		
		Category	Characteristics	Examples
High KU – High KD	High degree of both knowledge inflow and outflow through technology-based BMs	Technology	ICTs mainly related to subject (cryptography & cost/price) specific technologies	Register, cryptography, etc.
		Technology-based BM	Postage metering system BMs that mainly include subject (cryptography & cost/price) specific technologies	Postage meter system (cryptography & cost/price), etc.
High KU – Low KD	High degree of knowledge inflow and low degree of knowledge outflow through technology-based BMs	Technology	General technologies related to printing	Printing, typewriting machines
		Technology-based BM	Postage metering system BMs that include supporting technologies	Specialized function performed, recharging, record keeping, etc.

Low KU – High KD	Low degree of knowledge inflow and high degree of knowledge outflow through technology-based BMs	Technology	ICTs related to data/image processing	Electrical computers and digital processing systems, image analysis
		Technology-based BM	BMs that include business system infra technologies	Special goods or handling procedure, automated electrical financial or business practice or management arrangement

3.3.3.1. High KU-High KD

The technological classes in this group are characterized by a high degree of both knowledge inflow and outflow through the technology-based BMs. They are mainly ICTs and the postage metering system BMs related to subject (cryptography & cost/price) specific technologies: for example, register cryptography, postage meter system (cryptography & cost/price), etc.

3.3.3.2. High KU-Low KD

The technological classes in this group are characterized by a high degree of knowledge inflow and a low degree of knowledge outflow through the technology-based BMs. The affiliated classes include general technologies related to printing and the postage metering system BMs with supporting technologies; for example, printing, typewriting machines, specialized function performed, recharging, record keeping, etc.

3.3.3.3. Low KU-High KD

The technological classes in this group are characterized by a low degree of knowledge inflow and a high degree of knowledge outflow through the technology-based BMs. The affiliated classes include ICTs related to data/image processing and the BMs with business system infra technologies; for example, electrical computers and digital processing systems, image analysis, special goods or handling procedure, automated electrical financial or business practice.

3.3.4. Classification of knowledge flow drivers

The base BM patents as knowledge flow drivers are classified based on the amount of knowledge exchanged between the base BM patents and backward/forward citations, as shown in Figure 3.6 and Figure 3.7. Taking the mean value as a criterion, the base BM patents at patent class level are then classified into three groups, depending on the amount of knowledge flow in and out: knowledge utilizing & disseminating group, knowledge utilizing group and knowledge disseminating group. The Tables 3.4 and 3.5 present the groups of knowledge flow drivers in technology classes and in technology classes & technology-based BM classes, respectively.

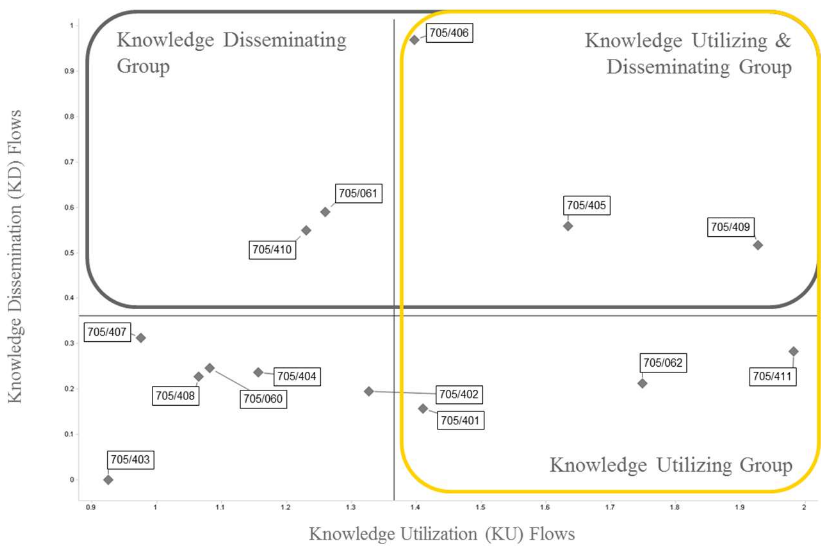


Figure 3.6 Positioning of base patent classes driving knowledge flows in

technology classes

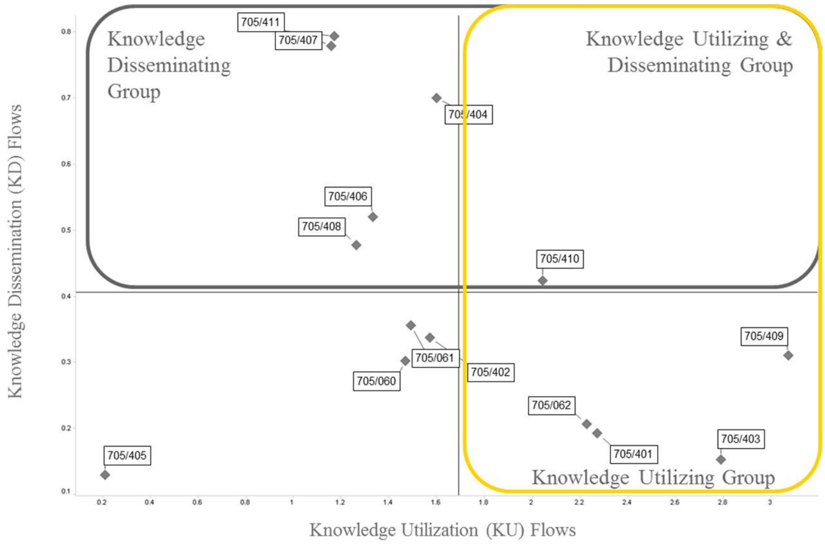


Figure 3.7 Positioning of base patent classes driving knowledge flows in technology-based BM classes

Table 3.4 Group of knowledge flow drivers: technology classes

Group	Subclass	Title
Knowledge Utilizing & Disseminating Group	705/409	Rate updating
	705/405	Data protection
	705/406	With specific mail handling means
Knowledge Utilizing Group	705/411	Display controlling
	705/062	Having printing detail (e.g., verification of mark)
	705/401	Postage metering system (Cost/Price)
Knowledge Disseminating Group	705/061	Reloading/recharging
	705/410	Specialized function performed

Table 3.5 Group of knowledge flow drivers: technology classes & technology-based BM classes

Group	Subclass	Title
Knowledge Utilizing & Disseminating Group	705/410	Specialized function performed
Knowledge Utilizing Group	705/409	Rate updating
	705/403	Recharging
	705/401	Postage metering system (Cost/Price)
	705/062	Having printing detail (e.g., verification of mark)
Knowledge Disseminating Group	705/411	Display controlling
	705/407	Including mailed item weight
	705/404	Record keeping
	705/406	With specific mail handling means
	705/408	Specific printing

3.3.4.1. Knowledge utilizing group

Knowledge utilizing group increases utility of the existing patents by acquiring a significant amount of knowledge from existing patents. Simply put, this group act as knowledge application. The group includes printing detail (e.g., verification of mark) and postage metering system (Cost/Price) as technology classes and technology-based BM classes

3.3.4.2. Knowledge disseminating group

Knowledge disseminating group spreads its valuable knowledge that will be widely used by the following patents and acts as knowledge provision. This group includes reloading/recharging and specialized function performed as technology classes, and record keeping, with specific mail handling means, including mailed item weight as technology-based BM classes.

3.3.4.3. Knowledge utilizing/ disseminating group

Knowledge utilizing/ disseminating group facilitates knowledge flow by both acquiring and disseminating a significant amount of knowledge. The affiliated classes include data protection, with specific mail handling means, rate updating as technology classes and specialized function performed as technology-based BM class.

3.4. Implication and conclusion

The study proposed an elaborated approach that explores knowledge flows through technology-based BMs. The proposed approach is integrating patent citation analysis and text mining technique, so that the range and degree of knowledge flows are measured together. In addition, possible problems that may arise when only text mining or citation analysis is used alone are reduced by combining citation analysis and text mining when verifying the degree of knowledge flows between technology-driven BMs and their cited/citing patents. The degree of knowledge measured by the proposed approach can induce the patterns of knowledge flow in detail. The technological classes are categorized into three different groups based on the amount of knowledge they provided to or received from the technology-based BM patents: high knowledge utilization–high knowledge dissemination, high knowledge utilization–low knowledge dissemination, and low knowledge utilization–high knowledge dissemination.

As technology-based BMs have become a patentable subject matter, they played a critical role in knowledge flow. Since ICTs are the integral part of the technology-based BMs, it is often overlooked that most of technological knowledge flow occurs between the ICT classes and the technology-based BMs. It is, however, found that not only the ICT classes but also other general technologies exchange a substantial amount of knowledge with the technology-based BMs. Some technology-based BMs utilize the knowledge from general technologies more than from technology-based BMs as well as disseminate its knowledge to technologies more than to technology-based BMs. In contrast, some technology-based BMs utilize the knowledge from technologies, but its main area of knowledge dissemination is with technology-based BMs. It is also found that there are active knowledge flows between technology-based BMs of different classes.

The proposed approach showed its strengths for handling the unstructured documents in exploring knowledge flows. Moreover, this paper contributed to an improved understanding of the value and function of technology-based BMs and BM patents from the knowledge flow perspective. It also identified the significant sectors in which knowledge is actively exchanged through the technology-based BMs. It is suggested to focus on the development of these sectors to stimulate coevolution of the technology-based

BMs. Since many small and medium companies struggle with lack of necessary skills and knowledge (Grandon and Pearson, 2004), the proposed approach and the classification of knowledge flow drivers can help decision makers to get a comprehensive view of technology-based BMs. Pathways of knowledge dissemination will give an advantage to obtain the benefits of R&D without having to pay its full cost.

However, the proposed approach should be carefully applied to practice. Setting the appropriate level of analysis is one of the important issues. Also, it has to be accounted that the co-word analysis used as text mining is performed on the keywords, and the text is not analyzed directly. This can lead to a problem in measuring textual similarity. For example, the similarity between two texts can be very high while the semantic contents in the texts are actually quite different. This problem can be mitigated by applying supplementary techniques, such as subject–action–object (SAO) analysis, or semantic similarity measures, such as Latent Semantic Analysis (LSA), Hyperspace Analogue to Language (HAL), Generalized Latent Semantic Analysis (GLSA), Cross-Language Explicit Semantic Analysis (CL-ESA) and Pointwise Mutual Information - Information Retrieval (PMI-IR).

Our case study has some limitations and requires further studies. The scope of BM patent data used in the case study is limited to postage metering system. Data from a wider range should be used to understand the overall system of technology-based BMs. Also, the positioning map only shows a snapshot of the phenomenon. Creating maps for different time periods will present the evolving stages of knowledge flows brought about by the technology-based BMs.

Chapter 4. Identification of Knowledge Flow Patterns

4.1. Introduction

A business model – a way to create, deliver and capture values – is not just a fundamental element of doing business anymore but rather becoming an important locus of innovation (Amit and Zott, 2012; Chesbrough, 2011; Desyllas and Sako, 2013). Conventional R&D innovations had long been regarded as a dominant growth strategy. However, these innovations often require a massive amount of investment, both in terms of time and money (Amit and Zott, 2012). As an alternative or complement to the conventional innovation, more companies are shifting their focus to business model innovation (Desyllas and Sako, 2013).

Business model innovation today is often implemented by information and communication technology (ICT) as a whole range of economic activity can be conducted through the Internet. While a new business model is itself unlikely to qualify for formal intellectual property (IP) protection, specific business methods underlying it can be protectable under the Patent Act (Desyllas and Sako, 2013; Wagner, 2008). These patents are generally called business method (BM) patents and cover some combinations of software and business methodology, including methods, systems, or processes for conducting various aspects of e-commerce (Bagley, 2000; Wu, 2005).

Therefore, valuable technological knowledge regarding business model innovation, especially online business model innovation, can be found in BM patents.

While BM patents, in general, accelerate technological progress and motivate business model innovation by stimulating exchange of knowledge, different BM patents play different roles in the growing process. BM patents have transferred a considerable amount of knowledge, as evidenced by a large number of citations made to and received by other patents (Allison and Tiller, 2003; No et al., 2015; Wagner, 2008). This implies that BM patents as a whole enable firms to obtain new technological and business knowledge with less effort and raise the productivity of inventive activities, which can result in more efficient business model innovation. A closer look, however, reveals that individual BM patents contribute to business model innovation in different ways (No et al., 2015). For example, some BM patents stimulate business model innovation by disseminating its knowledge, whereas some BM patents stimulate innovation by mediating knowledge interactions. Although there has been an attempt to identify such critical BM patents and their roles based on knowledge flows, the time-varying characteristics of a knowledge flow process have not been discussed. Whether a certain technology is a driver of evolution process depends on whether technological knowledge is useful for the new development and applications over a substantial period of time (Ho et al., 2014). Consequently, we need to observe knowledge flow patterns of BM patents with the long-term perspective to understand how different BM patents play their roles in business model innovation.

We adopt a Hidden Markov Model (HMM) as a method to identifying such dynamic patterns of knowledge flows driven by BM patents. An HMM is "a statistical tool for modeling generative sequences that can be characterized by an underlying process generating an observable sequence (Blunsom, 2004)." We use backward and forward citations, which are common proxy measures for knowledge inflows and outflows, as input datasets of an HMM. Backward citations are citations to prior patents made by a particular patent and have been used to measure inflows of technological knowledge; forward citations are citations to a particular patent made by subsequent patents and have been used to measure outflows of technological knowledge (Duguet and MacGarvie, 2005; Plasmans and Lukach, 2010). Although backward and forward citations contribute to knowledge flows in a similar manner (Plasmans and Lukach, 2010), their patterns and implications can be quite distinct from each other. Thus, they are used as separate inputs for an HMM. Once individual knowledge flow patterns are generated, those with similar characteristics are grouped together with a clustering method to identify major patterns. In the case study, we illustrate the application of the proposed approach on BM patents in 16 technological subclasses related to secure transactions.

An HMM is specifically employed in our study for the following reasons. First, an HMM is a statistical model in which the system being modeled is assumed to be a stochastic process. In general, the types of modelling methods can be dichotomized into the class of deterministic models, and the class of statistical models (Rabiner, 1989). Although both deterministic and stochastic models have had good success in various research fields, stochastic models can be more informative and flexible as it accounts for the uncertainty due to varying behavioral characteristics. In the field of technology and innovation management, stochastic processes have been applied to patent citations to observe technological trends or predict future technological impacts as they can capture the randomly varying characteristics of patent citations (Lee et al., 2017; Lee et al., 2016). That is, an HMM is capable of characterizing various temporal patterns of knowledge flows, rather than generating one or a specified number of representative patterns, such as an S-curve, based on certain assumptions. Second, an HMM is a dynamic model that accounts for time-dependent changes in the state of the system. Some statistical models, such as a finite mixture model, are similar to an HMM in that it provides a natural representation of heterogeneity in a finite number of latent classes. However, static methods cannot represent temporal patterns of knowledge flows through BM patents. Third reason is that the output of an HMM is represented by a sequence of a fixed number of states, instead of continuous values. This characteristic is the main advantage over classical dynamic approaches for stochastic modelling (i.e., autoregressive methods) in this type of research as it allows us to compare various long-term patterns more easily from a marco view and provide the practical implications of each state. Lastly, an HMM facilitates a clustering process for identifying major patterns. Analyzing technological trajectories with patent information often involves multivariate times series, which consist of a number of patents, patent citations, co-classifications and so on (Lee et al., 2016). Clustering multivariate trajectories is not an easy task because there is no proper way of defining the distance between arbitrary multivariate time series (Ghassempour et al., 2014). An HMM helps to overcome this issue by associating each multivariate trajectory with a sequence of states. In sum, an HMM is a very flexible tool that has no general theoretical limit in regard to statistical pattern classification (Bilmes, 2006) and thus expected to be more widely used in the field of technology and innovation management.

As BM patents play an increasingly important part in business model innovation, tracing their knowledge flow patterns can provide valuable insights in formulating more effective strategies or policies. With the proposed approach, firms can capture potentially influential or promising BM patents and identify their technological position in business model innovation. This helps to forecast the future direction of technology for business model innovation more accurately and thus informed innovation decisions can be made.

4.2. Hidden Markov Models

In real world processes, signals are often generated as observable outputs of sources. The goal of pattern recognition, therefore, is to characterize such real-world signals in terms of signal models (Fink, 2014; Rabiner, 1989). In the field of speech recognition research, HMMs have established the dominating processing paradigm, effectively superseding all competing approaches (Fink, 2014). HMMs have been actively used in other recognition fields including handwriting recognition and activity recognition and their application has been further expanded to weather prediction, financial time series analysis, robotics, computervision, computational biology and so on (Lee and Cho, 2011).). The great popularity of this modelling technique can be attributed to the fact that there is “no general theoretical limit to HMMs given enough hidden states, rich enough observation distributions, sufficient training data, adequate computation, and appropriate training algorithms (Bilmes, 2006).”

Recently, some attempts have been made to employ HMMs in the technology and innovation management studies. These studies have been successfully applied HMMs on patent data in analyzing growth patterns of technologies and innovations and technology life cycle. Lee et al. (2011) and Lee et al. (2012) identified the representative growth trends in the information and communication technology sector and energy sector based on the approach that incorporates HMMs, a clustering technique and a single patent indicator (i.e., variable). A similar approach was proposed to estimate stages of technology life cycles at the individual patent level and empirical analysis was conducted on laser technology in lithography (Lee et al, 2017). Lee et al. (2016) also analyzed life cycle patterns of molecular amplification diagnosis technology with multiple patent indicators, exploiting the advantage of HMMs in clustering multivariate trajectories. These studies have found HMMs a very useful tool to reflect uncertain and dynamic nature of a technology’s progression.

As can be noticed from its name, an HMM is developed based on the Markov model that is used to model generative sequences with a finite set of states and probabilities for transitioning from one state to another. These two models share the underlying assumption that the current state is solely dependent on the previous state. The major difference between these models is how they treat the notions of observation and state. While the sequence of observations (i.e., outputs) is simply the sequence of states visited in a traditional Markov model, the observations are the values taken by a number of variables and their probability distributions are a function of the "hidden" state in an HMM (Ghassempour et al., 2014). In short, an HMM can be defined as "a statistical tool for modeling generative sequences that can be characterized by an underlying process generating an observable sequence (Blunsom, 2004)." By separating visible observations from hidden states, HMMs can generate much more complex dynamic patterns than traditional Markov models, while maintaining computational efficiency (Ghassempour et al., 2014). Since the models are very rich in mathematical structure, they can be used in an extensive range of applications (Rabiner, 1989). They, when applied properly, also work well in practice and enable us to realize important practical systems, such as prediction systems, recognition systems and identification systems, in an efficient manner (Rabiner, 1989).

The use of an HMM in real-world applications involves one of three basic problems. The first problem is about choosing the model which best matches the observations, to which the solution is the forward-backward procedure (Rabiner, 1989). The second problem is related to decoding which is the process of estimating the state sequence that is most likely to have produced an observation sequence (Blunsom, 2004; Rabiner, 1989). One solution to this problem is to use the Viterbi algorithm. It is similar to the forward algorithm but the difference is that the transition probabilities are maximized at each step (Blunsom, 2004; Viterbi, 1967). The last problem involves adjusting the model parameters to maximize the probability of the observation sequence given the model. Since there is no known way to solve this problem analytically, an iterative procedure, such as Baum-Welch (or equivalently the expectation-maximization (EM)) method or gradient techniques must be used (Rabiner, 1989).

4.3. Proposed approach

4.3.1. Data

This study focuses on technologies for business model innovation and accordingly collects the relative patents from the largest patent database in the world, the United States Patent and Trademark Office (USPTO). US Patent Classification (USPC) system is a hierarchical way of assigning the technological class to which every patent belongs. In the USPTO database, BM patents belong to Class 705 that is entitled "Data Processing: financial, business practice, management, or cost/price determination" and is defined as follows:
"This is the generic class for apparatus and corresponding methods for performing data processing operations, in which there is a significant change in the data or for performing calculation operations wherein the apparatus or method is uniquely designed for or utilized in the practice, administration, or management of an enterprise, or in the processing of financial data.
This class also provides for apparatus and corresponding methods for performing data processing or calculating operations in which a charge for goods or services is determined.

This class additionally provides for subject matter described in the two paragraphs above in combination with cryptographic apparatus or method (USPTO, 2016)."

4.3.2. Research process

4.3.2.1. Select patent citations as a proxy for knowledge flows

This research considers two different types of citations, namely backward and forward citations since they both contribute to knowledge flows. The patterns and implications of backward and forward citations, however, may be quite distinct from each other as the underlying economic rationale of these two processes differ (Duguet and MacGarvie, 2005) and thus they are used as separate inputs for an HMM. Detailed descriptions of the measures are provided below:

- Backward citations (knowledge inflows): Backward citations are citations to prior patents made by a particular patent and regarded as the technological knowledge acquired by the patenting entity (Duguet and MacGarvie, 2005; Plasmans and Lukach, 2010). Backward citations, therefore, represent knowledge utilization which is closely related with a firm's productivity in R&D and innovation activities (Duguet and MacGarvie, 2005; Plasmans and Lukach, 2010). Many previous works have used backward citations as a proxy for knowledge spillovers across geographical, social or technological spaces (Jaffe et al., 2000; Narin and Olivastro, 1998).

The extent of knowledge acquired or absorbed by the entity can be measured by the frequency, that is, the total count, of backward citations (Ko et al, 2014; Wang et al., 2012).

- Forward citations (knowledge outflows): Forward citations are citations a particular patent receives from subsequent patents and thus can be interpreted as an indication that this patent has contributed to the dissemination of knowledge (Duguet and MacGarvie, 2005; Plasmans and Lukach, 2010). It has long been evidenced that forward citations have a positive correlation with the patent's quality and economic value (Hall et al., 2005; Yang et al., 2015). In other words, highly cited patents are considered more valuable and important. The challenging methodological problem of forward citation analysis lies in the long accumulation process which may cause truncation issues (Nemet, 2012). This problem restricts its application in capturing the dynamic process of knowledge flows lead by very recent innovations.

4.3.2.2. Measure time series citation data for BM subclasses

The unit of analysis is a subclass under the BM patent class, i.e., Class 705 of USPTO. The BM patent class consists of business method technologies of a wide variety of subjects, such as health care management, insurance and accounting, whereas each subclass under the BM patent class includes more homogenous technologies. Therefore, a subclass can be considered a specific type of technology. Knowledge flow patterns for specific technologies can be identified and compared via analyzing citations of different subclasses.

For each subclass, the bivariate (backward and forward citations) time series data of length T is collected and organized in a general form $O_{1:T} = (O_1^1, O_1^2,\ O_2^1, O_2^2,\ \dots\ ,\ O_T^1, O_T^2)$. The length T varies depending on the selection of a time interval. Since every patent contains the dates of registration and citation, the time interval can range from a day to a few years. If the time interval is too short or too long, in other words, if the length T is too long or too short, it may be difficult to interpret the result or draw meaningful conclusions. An appropriate time interval for our case study is a year.

4.3.2.3. Construct a HMM and generate sequences of knowledge flow states

A set of backward and forward citation trajectories can be mapped into an HMM given a number of states, that is a probability density over the space of trajectories, $P(T \mid \lambda)$. P takes a certain functional form and λ is a set of parameters, A, B and π. These elements of an HMM are explained as follows (Rabiner, 1989):

1) A finite number of states (states of knowledge flows)

$$S = \{s_1, s_2, \dots, s_N\}. \tag{1}$$

2) A discrete set of possible symbol observations (backward and forward citations)

$$V = \{v_1, v_2, \dots, v_M\}. \tag{2}$$

3) State transition probability distribution

$$A = \{a_{ij}\},\ a_{ij} = P(q_{t+1}=s_j \mid q_t=s_i);i,j = 1,\ \ldots,\ N \tag{3}$$

where qt denotes the state at time t.

4) Observation symbol probability distribution in state

$$P(O_t=v_k \mid q_t=s_j);\ j = 1,\ \ldots,\ N;\ k = 1,\ \ldots,\ M \tag{4}$$

where Ot, denotes the observation at time t.

5) Initial state distribution

$$\pi = \{\pi_i\},\ \pi_i= P(q_1=s_i);i=1,\ \ldots,\ N. \tag{5}$$

These parameters are estimated using the Baum-Welch method, which is a special case of the EM method. The EM method iterates between two steps: (1) the expectation step (E-step), and (2) the maximization step (M-step). In the E-step, current parameters of the model is assumed and the expected values of necessary statistics are computed. In the M-step, these statistics are used to re-estimate the model parameters so as to maximize the expected likelihood of the parameters. The two steps are repeated until a certain criterion is fulfilled (Li and Biswas, 1999). The detailed information of the EM method is given by Bilmes (1998) and the references therein.

Unfortunately, parameter estimation methods do not estimate the optimal number of hidden states (the order), so that an alternative criterion, such as the Akaike information criterion (AIC) or Bayesian information criterion (BIC), should be used to determine this parameter. Here, we use the Bayesian information criterion (BIC) (Schwarz, 1978), which is defined as following:

$$\mathrm{BIC}(J) \quad = \quad -2\log L(J) + \mathrm{n_p} \log(N) \tag{6}$$

where $L(J)$ is the likelihood for the model order J, n_p is the dimensionality of the model parameter space and N is the number of cases in data. The model order J with the lowest BIC is selected as the optimal number of hidden states (Lee et al, 2012).

Finally, the most likely sequence of states, $\widehat{s_{1:T}} = (\widehat{s_1}, \widehat{s_2}, \dots, \widehat{s_T})$, for each BM patent subclass is generated through decoding. There are largely two different types of decoding: local and global decoding. While the former refers to the derivation of the most likely state at date t, the latter looks for the most probable sequence of states (Mergner, 2009). Since local decoding partially ignores the transition probability matrix, using local decoding for each t separately cannot determine the most probable sequence (Mergner, 2009). Moreover, the computational cost would be very high. Therefore, we employ global decoding, the Viterbi algorithm, that is more effective and efficient in the context of Markov switching models. The Viterbi algorithm calculates the most probable path, keeps track of the states that make up this path, and backtracks from the final state to the first state to choose the most probable state at each instant (Lee et al., 2017).

Although BIC is employed to choose the appropriate number of states, this does not guarantee that the selected model is appropriate. Residuals are a common tool for assessing the fit of a model in many branches of statistical modelling. In HMM, the goodness-of-fit of a fitted model is generally assessed using one of two types of pseudo-residuals: the uniform pseudo-residual and normal pseudo-residual (Zucchini and MacDonald, 2009). We the normal pseudo-residual which can be defined as follows:

$$z_t = \Phi^{-1}(u_t) = \Phi^{-1}\left(F_{X_t}(x_t)\right) \tag{7}$$

where Φ is the distribution function of the standard normal distribution and X a random variable with distribution function F. Then $Z \equiv \Phi^{-1}(F(X))$ is distributed standard normal. If these normal pseudo-residuals are distributed standard normal, with the value of the residual equal to 0 when the observation coincides with the median, then the fitted model is valid (Zucchini and MacDonald, 2009). Model adequacy is checked by the qq-plot of the pseudo-residuals against the theoretical quantiles of the standard normal distribution.

4.3.2.4. Cluster knowledge flow state sequences to identify major patterns

Clustering analysis helps to identify major patterns of knowledge flow through BM patents by measuring distances between the state sequences of BM patent subclasses and partitions them in such a way that sequences in the same cluster are more similar to each other than to those in other clusters. To measure the distance between state sequences, we adopt the clustering attribute, $Duration_i$, proposed in the previous studies (Lee et al., 2016; Lee et al., 2017) and modify it to obtain a better clustering result. $Duration_i$ is the total time units spent in each state i, but this may not characterize the overall trend of a sequence. For example, the time units spent in each state can be the same for two sequences while trends in the sequences are very different from one another – an increasing trend in one sequence and a decreasing trend in the other. Therefore, we divide the total time length into several periods and measure the time units spent in each state during each period. The attributes can be represented as $Duration_i^j$, where i and j denotes the state and period respectively. When the number of states and periods are n and m, the number of attributes will be $n \times m$.

Hierarchical clustering methods are preferred over non-hierarchical methods in our study since we do not have prior domain knowledge to determine the best number of representative patterns and thus need to look at partitions at various levels of details. The process of hierarchical clustering can be presented in a dendrogram. It is more informative than the unstructured set of clusters returned by non-hierarchical methods. Although hierarchical methods do not require that the number of clusters be specified in advance (Érdi et al., 2013), determining the optimal number of clusters from a hierarchical structure or termination condition still involves a subjective decision to a greater or lesser extent (Lee et al., 2011).

There are two types of hierarchical methods based on how the hierarchical decomposition is formed. The agglomerative approach is the bottom-up approach in which each object starts in its own cluster and the clusters are iteratively merged into larger clusters until all the objects are in a single cluster or certain termination conditions are satisfied. In contrast, the divisive approach is the top-down approach that starts with all the objects in the same cluster and iteratively split the cluster into several smaller subclusters until each object is in its own cluster. A challenge with divisive methods is that they typically use heuristics partitioning due to high computational costs and this can lead to inaccurate results, especially when the number of objects is large (Han et al., 2011). Therefore, the agglomerative clustering is employed in this study.

4.4. Case study

4.4.1. Data

Among a variety of subclasses in Class 705, those related to secure transaction (64-79) are selected for our case study due to two main reasons. First, the patents in these subclasses enable secure electronic transactions which have contributed to the explosive growth of online business environments. Security technologies are the principal design factor for online business models and have been applied in a wide range of e-commerce areas. It is thus anticipated that these patents have been actively exchanging knowledge with other patents. Second, 16 subclasses are an adequate number to identify and compare different patterns of knowledge flows.

We collect a total of 1,265 patents registered from 1976 to 2015 within these subclasses and these patents had made altogether 41,217 citations of previous patents and cited by 57,873 subsequent patents. Table 4.1 presents the numbers of patents, backward citations and forward citations for each subclass. The number of citations only includes patent citations; it does not include citations made to or received by non-patent literature and special patents.

4.4.2. Analysis and results

Our data is composed of 16 knowledge flow trajectories corresponding to 16 subclasses, where each trajectory is a set of time series data for two variables, backward and forward citations, as shown in Table 4.2. The length of a trajectory varies across the subclasses due to the difference in age of the subclasses. The lengths of trajectories do not have be the same for modeling (Ghassempour et al, 2014). We assume that the all trajectories begin at the period 0 since the period 0 can be considered as the time when the first patent of each subclass was issued. This can be a valid assumption since we identify trajectories of knowledge flows for the subclasses from their emergence regardless of the time of their emergence, as in the life cycle model.

Table 4.1 Selected subclasses under secure transaction

Subclass	Title	Beginning year of patent registration	Number of patents	Number of backward citations	Number of forward citations
64	Secure transaction (e.g., EFT/POS)	1996	310	11,170	5,996
65	Including intelligent token (e.g., electronic purse)	1988	120	3,991	6,666
66	Intelligent token initializing or reloading	1984	30	390	2,039
67	Including authentication	1980	225	9,408	5,932
68	Balancing account	1993	23	701	1,951
69	Electronic cash detail (e.g., blinded, divisible, or detecting double spending)	1990	41	730	4,116
70	Home banking	1985	10	163	989
71	Including key management	1980	47	1,160	1,677
72	Verifying PIN	1976	42	438	1,709
73	Terminal detail (e.g., initializing)	1977	10	103	789
74	Anonymous user system	1995	60	2,253	1,818
75	Transaction verification	1981	160	4,746	6,073
76	Electronic credential	1980	91	3,165	4,964
77	Including remote charge determination or related payment system	1997	28	744	3,028
78	Including third party	1988	44	1,047	5,240
79	Including a payment switch or gateway	1998	24	1,008	4,885

Table 4.2 Time-series citation matrix

Sub- class	Type of citation	1 1976	2 1977	3 1978	… …	38 2013	39 2014	40 2015
64	Back	-	-	-	…	724	1,788	184
	For	-	-	-	…	1,016	1,152	768
65	Back	-	-	-	…	476	814	71
	For	-	-	-	…	756	733	393
…	…	…	…	…	…	…	…	…
72	Back	5	0	0	…	84	48	0
	For	0	1	4	…	227	260	158
…	…	…	…	…	…	…	…	…
78	Back	-	-	-	…	376	424	188
	For	-	-	-	…	0	0	0
79	Back	-	-	-	…	628	725	357
	For	-	-	-	…	31	6	99

Since the optimal number of hidden states is unknown, the dataset is modelled using HMMs of a different number of hidden states. In the beginning, BIC is computed with two hidden states. As the number of hidden states increases, the BIC value gradually decreased and reached its minimum at four hidden states. Therefore, the correct model structure is found to be a four-state model and the parameters estimated for this model are shown in Table 4.3. Because the variables, backward and forward citations, consist of count data with a large variance, the data is almost certainly over-dispersed (Nemet et al., 2012) and skewed, suggesting that a skew normal distribution would be a better fit than a normal or Poisson distribution. For that reason, the variables are modeled according to a multivariate skew normal distribution.

An increase in state numbers should not be directly translated into an increase in knowledge flows because these numbers are randomly assigned to simply distinguish different states. These state numbers are reassigned so that the state number increases as the mean values for emission probability distributions of backward and forward citations increase, as shown in the last two columns of Table 4.3. In other words, State 1 is characterized by the lowest level of knowledge flows in both directions whereas State 4 is characterized by the highest level. This would help to understand and interpret the results more easily. The mean values for backward and forward citation distributions in State 1 are 3.0 and 2.1, which means that a subclass exchanges only a trivial amount of knowledge when it is in State 1. The mean values for backward and forward citation distributions in State 2 are 4.8 and 30.5 and those in State 3 are 44.1 and 100.6. State 4 is associated with large mean values, 388.2 and 445.4, for the respective measures. The bivariate plots for the four states are presented in Figure 4.1.

The transition probabilities show that a subclass at any point in time is most likely to remain in the same state at the next point in time. As can be seen from the initial and transition probabilities, most of the BM subclasses begin at State 1 – the initial probability for State 1 is 0.937 – and tend to remain in that state with a probability of 0.850. The transition probability that a subclass would be in a stationary position (i.e., no transition to other states) is also higher than 0.8 when it is in other states, and the probability is the highest (0.914) in State 4.

Table 4.3 HMM parameters

	Initial probability	Transition probability				Mean of emission distribution	
		State 1	State 2	State 3	State 4	Back-ward	Forward
State 1	0.937	0.850	0.060	0.089	0.000	3.0	2.1
State 2	0.000	0.000	0.889	0.111	0.000	4.8	30.5
State 3	0.063	0.000	0.069	0.832	0.099	44.1	100.6
State 4	0.000	0.000	0.000	0.086	0.914	388.2	445.4

The goodness-of-fit of the fitted model is then assessed using the normal pseudo-residual. The qq-plot in Figure 4.2 clearly shows that the selected model provides an acceptable fit.

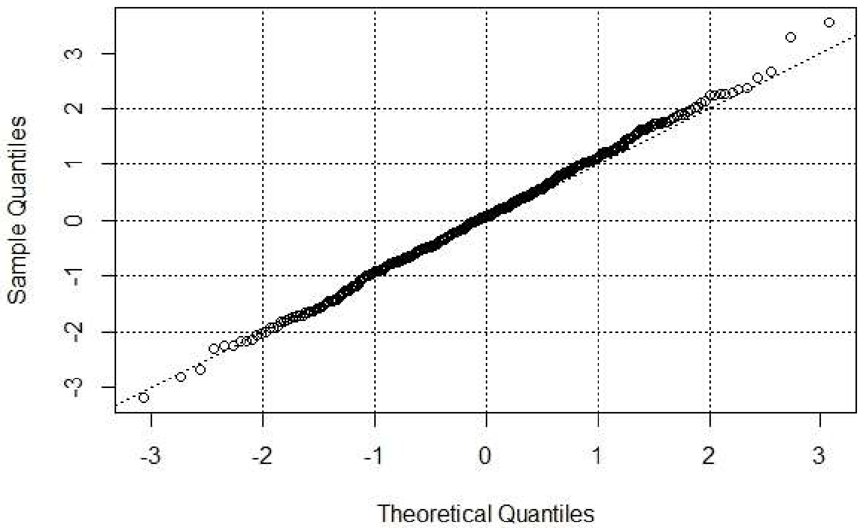


Figure 4.1 QQ-plot of pseudo residuals

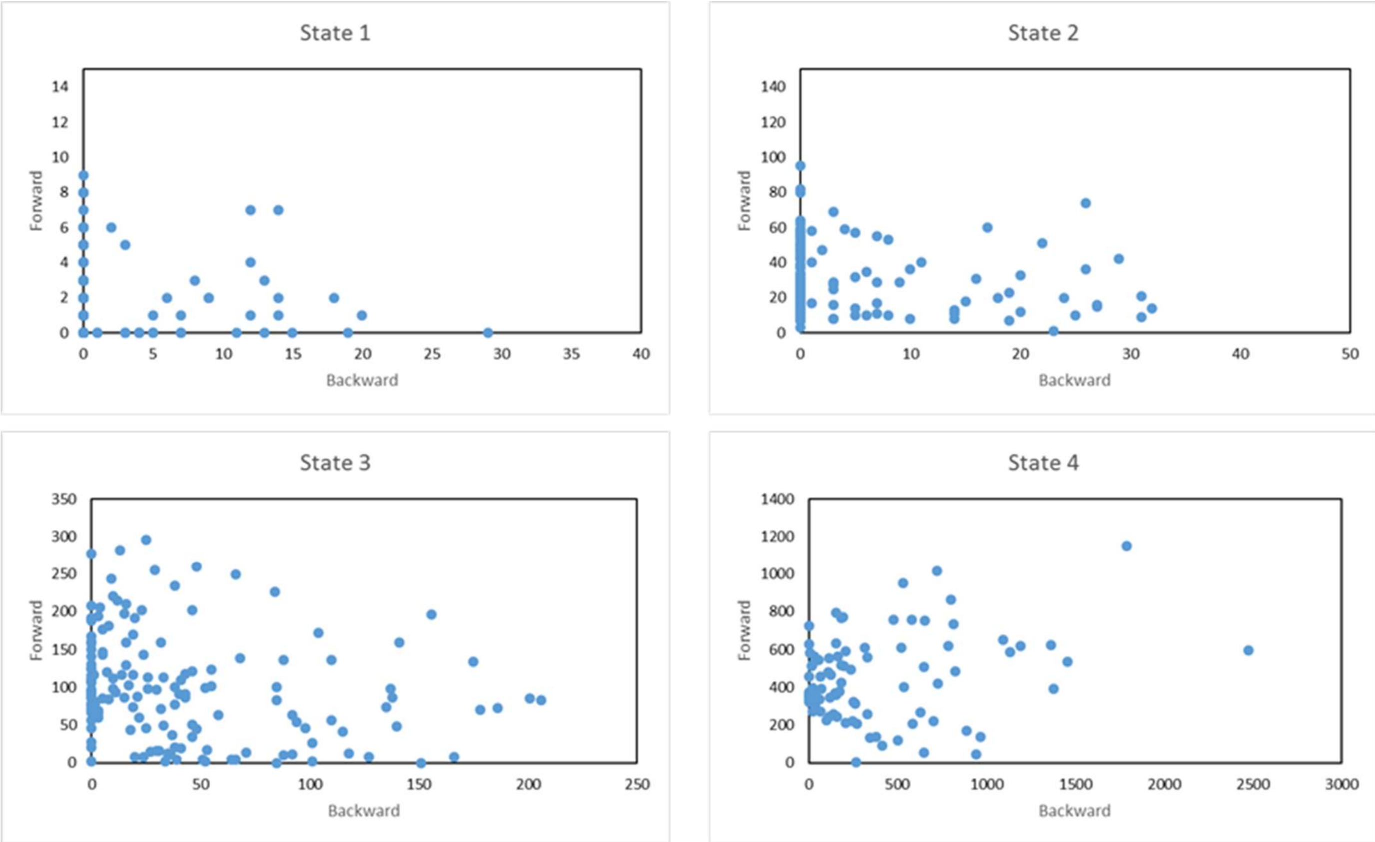


Figure 4.2 Bivariate plots for four states

Based on the four-state model, we derive temporal state changes for individual subclasses and the results are displayed in Figure 4.3. The horizontal axis represents the time period and the vertical axis represents the state. Although most of the subclasses exhibit an increasing tendency, the rate at which the knowledge flows increase and variability in knowledge flows may differ.

We then perform clustering analysis to identify major patterns of knowledge flows. We divide the total time length into four periods, each of which consisting of 10 time units, and measure the time units spent in each state during the period. Since there are four states, we have a total of 16 attributes. The agglomerative hierarchical clustering method is used to group the individual patterns into a hierarchy of clusters. By taking a close look at the dendrogram and the individual patterns, four clusters are selected as the representative knowledge flow patterns for the secure transaction BM subclasses (Figure 4.4).

Figure 4.3 Temporal state changes

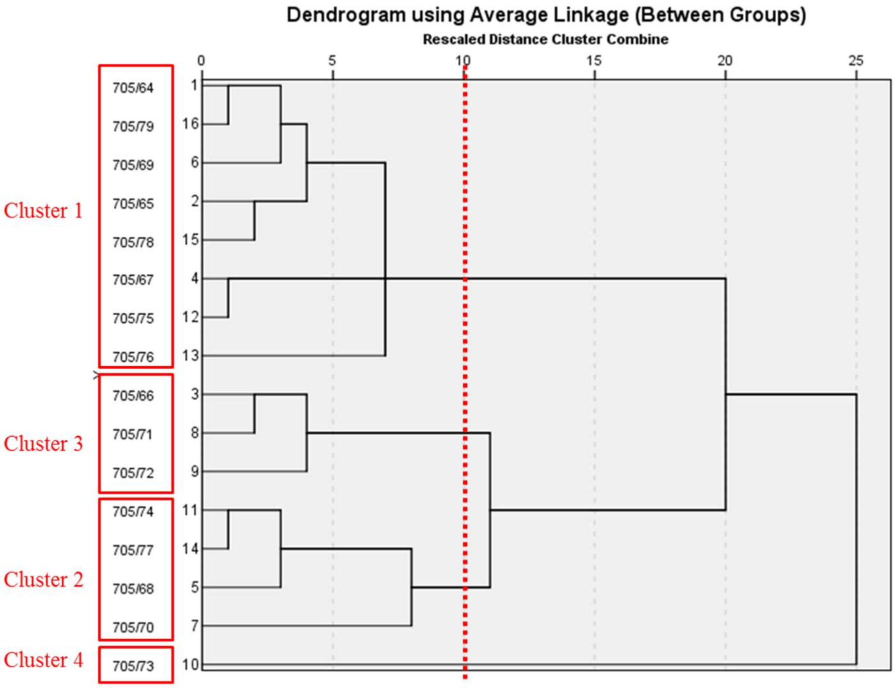


Figure 4.4 Dendrogram generated by AHC algorithm

4.4.3. Discussions

4.4.3.1. Major patterns of knowledge flows

Each state needs to be defined and explained in detail for better interpretation of the results. The amount of both knowledge inflows and outflows in State 1 is close to zero. It is a trivial amount in comparison with those in other states so that State 1 can be considered as a "closed" state. In other words, a BM subclass rarely exchanges knowledge when in this state. State 2 is a "minor dissemination" state which represents a small yet somewhat noticeable amount of knowledge outflows. When compared with State 1, knowledge outflows is larger but knowledge inflows is nearly the same. This implies that external knowledge is still underutilized and productivity in R&D and innovation activities regarding new business models is likely to be low in this state. It may be natural that a BM subclass tends to be in the closed state or minor dissemination state during their early phase. However, if a BM subclass stays in these states for a long period, it means that they have not exerted much impact on development of new business models and thus managers or policy makers need to take some actions to promote knowledge flows. State 3 is a "dissemination-intensive" state. Overall, a substantial amount of knowledge is exchanged in this state, yet there is a large gap between inflows and outflows – outflows is much larger than inflows. It is important to recognize the value of a BM subclass in this state as a knowledge source of business model innovation. At the same time, more efforts have to be placed on improving knowledge utilization. State 4 is a "stabilized exchange" state in which the amounts of knowledge inflows and outflows are significantly larger than other states and quite balanced. This is the most desirable state as BM subclasses in this state are able to not only play the role of knowledge source but also identify, assimilate and utilize new external knowledge. Therefore, these BM subclasses can be considered as a critical driver of business model innovation.

Four clusters of knowledge flow patterns, denoted by a state sequence, are examined and summarized in Table 4. Cluster 1 is the largest group composed of eight subclasses (64, 65, 67, 69, 75, 76, 78 and 79). These subclasses include cryptographic apparatus and methods designed for or utilized in electronic funds transfer (EFT) processing, point-of-sale (POS) terminal processing, transaction verification and so on. They are used for securing general transactions or financial data processing in a variety of internet-based businesses, ranging from online retail sectors to online financial services.

Table 4.4 Characteristics of clusters

Cluster	Size	Associated BM subclasses		Pattern
		Numbers	Titles	
1	8	64, 65, 67, 69, 75, 76, 78, 79	Secure transaction (e.g., EFT/POS), Including intelligent token (e.g., electronic purse), Including authentication, Electronic cash detail (e.g., blinded, divisible, or detecting double spending), Transaction verification, Electronic credential, Including third party, Including a payment switch or gateway	Thriving with low variability
2	4	68, 70, 74, 77	Balancing account, Home banking, Anonymous user system, Including remote charge determination or related payment system	Newly rising
3	3	66, 71, 72	Intelligent token initializing or reloading, Including key management, Verifying PIN	Gradually growing
4	1	73	Terminal detail (e.g., initializing)	Stagnant

In the typical pattern of Cluster 1, the knowledge flows rise rapidly at a certain point and then increase gradually over time. Cluster 1 can be considered as a stable and steadily growing medium for knowledge exchange. Most of the BM subclasses in this group show a jump from the closed state to the dissemination-intensive state in the range of the period 20 to 23. They then progress to the stabilized exchange state between the period 30 and 33. Since these BM subclasses have reached the maximum state, they are found to be capable of both utilizing and disseminating volumes of knowledge.

The knowledge flow pattern of Cluster 1 is closely associated with the trend of innovations in e-commerce sectors. For example, the exponential growth of e-commerce began in 1995 (period 20) with the first widespread use of the Web to advertise products and lasted until 2000 (period 25) (Laudon and Traver, 2013). After a few years of consolidation, e-commerce entered the new era with the appearance of social networking and user generated content sharing in 2006 (period 31) and with the introduction of smartphones in 2007 (period 32) (Laudon and Traver, 2013). In this respect, we can assume that the BM subclasses for secure transaction in Cluster 1 have proliferated and exchanged a considerable amount of knowledge, keeping pace with the major growth pattern of e-commerce technologies.

Both Cluster 2 and 3 as well show an increasing tendency of knowledge flows but have not yet completely reached State 4. Cluster 2 is formed of four BM subclasses (68, 70, 74 and 77) that are primarily related to security for back office activities, such reconciliation of cumulative transaction data, and personal financial activities, such as home banking. The BM subclasses in Cluster 2 are relatively recent inventions as they appear around the period 20. Cluster 2 rapidly goes up to the dissemination-intensive state at the period 30 and displays some ups and downs between the dissemination-intensive state and stabilized exchange state.

The BM patents in Cluster 3 are associated with specific cryptographic techniques or tools including a token, key and PIN. E-commerce offers efficient and convenient online services, but at the same time, it creates new risks and security threats. Cryptographic tools are therefore a crucial technical requirement for any online transaction activities. The patterns for the BM subclasses (66, 71 and 72) in Cluster 3 stays in the minor dissemination state for a long period and goes up to the next state at around the period 32, which is even later than Cluster 2. The trend shows that Cluster 2 and 3 are now becoming an important knowledge intermediary.

Cluster 4 includes only one BM subclass (73), which is terminal detail (e.g., initializing), and presents a knowledge flow pattern that is very different from the others. It remains stagnant in minor dissemination state for almost 20 periods. The reason might be high technological uncertainty, low applicability, low customer acceptance, high costs or any combination of these (Lee et al., 2017).

4.4.3.2. Methodological implications and extensions

One of our goals in study is to employ an HMM that is sound, easy to explain and effective in clustering technological trajectories. To our knowledge, this is the first attempt to use such tool in tracing long-term trends of knowledge flows through BM patents, which are an important driver of business model innovation. Application of HMMs is a quite recent topic in the technology and innovation management field. A simple HMM has been successfully employed in the technology and management studies because the aim of these studies is to observe and explore evolutionary trends and life cycle of technology from a macro-view rather than solving specific problems or making accurate predictions (Lee et al., 2016; Lee et al., 2017; Lee et al., 2011; Lee et al., 2012). However, we have to admit that our approach that is based on a simple HMM has some limitations: we implicitly assume geometrically distributed sojourn time and does not take account heterogeneity between observations. These limitations require further studies to employ the extensions of an HMM, such as a hidden-semi Markov model and mixed HMM. The extensions of an HMM have some advantages over a simple HMM. For example, a hidden-semi Markov model can model the runlength distributions instead of implicitly following the geometric distribution of a HMM. Also, a mixed HMM permits greater flexibility in modeling correlation structure by relaxing the assumption that the observations are independent (Altman, 2007). We believe this study can provide an impetus for future research on a dynamic process of knowledge flows, especially from a methodological perspective.

4.5. Conclusions

As business models are becoming important assets to gain a competitive edge, a growing number of companies are attempting to innovate their business models and protect related technologies as patents. The emergence of BM patents has promoted knowledge exchange, as evidenced by a substantial amount of backward and forward citations. Although BM patents are gaining momentum in exchanging knowledge, such phenomenon is often analyzed statically without considering a time dimension. The static analysis cannot capture dynamically changing processes of knowledge flows in BM patents and as a consequence cannot provide valuable insights for developing strategies or policies for business model innovation. Thus, we proposed an HMM to identify temporal patterns of knowledge flows in BM patents, taking two proxy measures, backward and forward citations, as input data.

Our analysis characterizes different knowledge flow patterns for the 16 subclasses of secure transaction BM patents. Since the overall BM patents are a relatively immature field which is evolving rapidly, all but one of subclasses related to secure transactions show growing patterns in knowledge flows. Although the growth rate and variability differ by clusters, we can conclude that the BM patents for secure transactions, in general, play increasingly important roles in advancement of business models. They not only enable successful functioning of e-commerce applications but also facilitate the transfer of knowledge between BM technologies, which can lead to creation of new innovations. Thus, appropriate policy and investment in this area will ensure continuous and faster innovation of e-commerce business models.

We believe this study contributes to the field of knowledge flows in several ways. First, we attempt to model dynamic patterns of knowledge flows without any arbitrary hypothesis or bias by employing an HMM, which is a very flexible tool for characterizing various temporal patterns. Second, we identify major knowledge flow patterns through clustering of individual patterns and examine their characteristics. Lastly, we focus on knowledge flows of BM patents, which can motivate business model innovation, rather than conventional R&D innovation.

Our study offers some important implications for firm managers and policy makers. Our approach is useful in capturing potentially productive and valuable BM patents that can deliver a better return on investment. For example, BM patents that show a rapidly increasing trend can be considered an emerging business technology. In addition, investigating knowledge flow patterns can help firms to identify their technological position in business model innovation and formulate appropriate strategies. BM patents that stay in the state of a low-level flow, such as the closed or minor dissemination state, for a long time indicate that both productivity in innovation activities of a firm and value of its BM patents are likely to be low. In such case, managers should encourage employees to engage in external professional networks or consider strategic R&D cooperation with other entities to foster knowledge flows. On the other hand, BM patents that are steadily growing or already in the state of a high-level flow, such as the stabilized exchange state, indicate that the firm has an ability to fully exploit external knowledge as new ideas for business model innovation and at the same time, its knowledge has substantial effects on development of new business models. To keep up this momentum, managers should make efforts to invest time and resources toward managing the external knowledge sources and developing BMs with more general applicability. Furthermore, it is desirable to have a balanced picture of knowledge utilization and knowledge dissemination (Plasmans and Lukach, 2010), so the type of knowledge flows need to be taken into account when formulating strategies. For example, a firm with BM patents in the dissemination-intensive state should focus on improving its absorptive capacity. In sum, the proposed approach support managers or policy makers to forecast the future direction of BM patents more accurately and accordingly they can make informed decisions on technology investment and business model management.

Some tasks still remain for future research. The first task is to elaborate the proposed methodology. This study only considers the direction of knowledge flows – inflow and outflow – for developing proxy measures, but these measures can be further specified depending on the focus of research. One example can be defining the measures by the types of patent classes with which BM patents exchange knowledge. HMMs can also be performed several times with different combinations of measures to explore more meaningful implications. Moreover, the extensions of an HMM, such as a hidden-semi Markov model and mixed HMM, can be employed as they have some advantages over a simple HMM. For example, a hidden-semi Markov model can model the runlength distributions instead of implicitly following the geometric distribution of a HMM. Also, a mixed HMM permits greater flexibility in modeling correlation structure by relaxing the assumption that the observations are independent (Altman, 2007). The second task is expanding the dataset for a case study. The data used in this research is confined to 16 subclasses of BM patents. Incorporating a wider range of subclasses will enhance understanding of knowledge flows in BM patents.

Chapter 5. Investigation of Knowledge Transferors

5.1. Introduction

Social features or social media applications have become an important element of e-commerce as posts and comments shared by people around the globe are determining the success of products and services and shaping market trends (Anderson et al., 2011). With the explosive growth of social networks, traditional e-commerce, which is based on one-to-one interactions of the customer and seller, have been transformed into a more social and interactive form of e-commerce, referred to as social commerce (Stephen & Toubia, 2009; Wang & Zhang, 2012). There is no standard definition for social commerce, yet it is generally understood as a combination of e-commerce and social activities. By incorporating social features that support customers' participation in the developing, selling, buying and marketing of products and services, social commerce can offer novel business models which are beneficial to both sellers and buyers (Curty and Zhang, 2013; Huang and Benyoucef, 2013). In social commerce, sellers can convert customers into brand advocates, and at the same time, buyers can share shopping experiences with others and make better-informed purchasing decisions (Ng, 2013). Clearly, social commerce is a paradigm shift in ways of doing business and opens up new business opportunities (Busalim, 2016).

In response to this paradigm shift, firms are keen to find ways to embed social features in their business models through novel applications of information and communication technology (ICT). Business processes, practices, and operations are increasingly implemented by ICT and even an entire business model can be embedded in digital code (Ovans, 2000). This is especially true for social commerce that is heavily mediated by technological capabilities and advancements, such as Web 2.0, Cloud Computing and Service Oriented Architecture (SOA) (Wang & Zhang, 2012). Thus, technology is largely responsible for development of new business models in social commerce by promoting social interactions, supporting integration between social and commercial activities, and innovating functionality. Based on ICT, a growing number of companies have developed new business methods (BMs) that leverage social relationships into commercial gain and protected these BMs by means of a patent.

Since BMs have become a key enabler of new business models in social commerce, in-depth analysis of these BMs is crucial for understanding the overall trend and further capturing new opportunities in social commerce. Specifically, we need to identify BMs that have a significant impact on technological change and innovation through observing knowledge interactions of BMs. The importance of a certain technology, especially in the aspect of growing process, often depends on the extent to which it contributes to knowledge flows and diffusion (Ho et al., 2014). Therefore, BMs that facilitate knowledge flows are potentially valuable and productive as they can play a critical role in the whole evolution process.

In this regard, this paper aims to propose a framework for identifying core BMs in social commerce from a knowledge flow perspective. Our framework uses patents as a data source since BMs that are commercially valuable in practice can be found in patents. Moreover, patents encapsulate important technological knowledge and patent citations report the potential knowledge flows between the citing and the cited patents (Duguet and MacGarvie, 2005). One issue is that patent classification schema do not provide classes or subclasses specifically assigned to business method patents for social commerce. That is, these patents are organized merely following a standard classification scheme which does not consider features of social commerce. Hence, we collect BM patents related to social commerce and then reclassify them into clusters based on textual similarities. Since textual data in each patent contains the most comprehensive information, patents with high textual similarities can be regarded as similar technologies. Next, we evaluate each BM cluster using citation-based indicators. Patent citation data is widely used in empirical studies to measure knowledge flows (Guan and Chen, 2012). Finally, we determine core BMs and investigate their differentiated roles in knowledge flows.

5.2. Social commerce

The definition for social commerce has not yet been clarified. Social commerce can be defined from different perspectives as it involves multiple disciplines, including marketing, computer science, sociology and psychology (Huang and Benyoucef, 2013). The definition or usage of the term can also vary as social commerce evolves. In the early research, social commerce referred both networks of sellers and networks of buyers. That is, social commerce was considered as a subset or evolution of traditional e-commerce (Stephen and Toubia, 2009; Liang and Turban, 2011). Some researchers have described social commerce combining e-commerce, social media and Web 2.0 technologies. Liang and Turban (2011) define social commerce as "the delivery of e-commerce activities, services and transactions throughout social media environment, mostly on social networks and by employing Web 2.0 software."

Huang and Benyoucef (2013) give a similar definition: "an Internet-based commercial application, leveraging social media and Web 2.0 technologies which support social interaction and user generated content in order to assist consumers in their decision making and acquisition of products and services within online marketplaces and communities." Some of definitions are associated with two major configurations of social commerce Web sites: one is commercial features added to social networking Web sites to allow for advertisements and transaction; the other one social networking capabilities added to traditional e-commerce Web sites to take advantage of the power of social networking (Liang and Turban, 2011). Stephen and Toubia (2010) and Wang and Zhang (2012) describe that social commerce involves buying and selling of products and services through social media, which corresponds to the first configuration. Some researchers, on the other hand, stress the second configuration and consider social commerce as social features applied to e-commerce or online marketplaces (Huang and Benyoucef, 2013). Although social commerce can be developed in different ways, they all share the similar underlying concept which is a combination of commercial and social activities

(Liang and Turban, 2011). The most comprehensive and configuration-neutral definition (Yadav et al., 2013) refers to social commerce as “exchange-related activities that occur in, or are influenced by, an individual's social network in computer-mediated social environments, where the activities correspond to the need recognition, pre-purchase, purchase, and post-purchase stages of a focal exchange.”

Social commerce can be differentiated from e-commerce in terms of business goals, customer connection and system interaction (Huang and Benyoucef, 2013; Busalim, 2016). While the business goal of e-commerce is mainly maximizing efficiency in transaction and commercial activities, such as sophisticated searches and one-click buying, that of social commerce is focused more on social activities, such networking, collaborating and information sharing (Huang and Benyoucef, 2013). With regard to customer connection, e-commerce customers usually interact with websites individually and independently from other customers whereas social commerce customers are enabled to have community-based interactions on social network-based platforms (Huang and Benyoucef, 2015; Stephen and Toubia, 2009). Regarding system interaction, e-commerce provides a one-way browsing where information from customers is rarely sent back to businesses or other customers. Social commerce, on the other hand, provides more social and interactive environment which allows customers to share their information with businesses or other customers (Huang and Benyoucef, 2013). Differences between e-commerce and social commerce can be further analysed in many different aspects, including value creation, business models, technology, design and so on ((Baghdadi, 2016).

When compared with traditional e-commerce, social commerce can bring more benefits to both customers and sellers by utilizing social features that support interaction, communications and collaborations to assist in buying, selling and marketing of products and services (Ng, 2013). Social commerce sites allow consumers to engage, communicate and share information with other people regarding their purchasing experiences to support customers' problem solving and decision-making (Ng, 2013). That is, customers can build mutually beneficial relationships. The social activities of customers can also create word of mouth or viral marketing effects which can be more powerful and cost-effective than traditional advertising (Han and Kim, 2016; Huang and Benyoucef, 2015). In addition, firms can better understand their customers by getting real-time customer feedbacks, build strong customer-to-seller relationships, and eventually, convert customers into brand advocates (Han and Kim, 2016; Huang and Benyoucef, 2013).

5.3. Research framework

5.3.1. Overall research framework

Figure 5.1 shows the overall framework of our research. The proposed approach consists of three phases: (1) Data collection, (2) Classification of BMs in social commerce and (3) Identification of core BMs. The first phase collects BM patents for social commerce and extract two types of information, keywords and citations, from the patents using a text-mining technique. In the second phase, textual similarities between all pairs of BM patents are measured using the cosine similarity. Based on the textual similarities, the BM patents are grouped into clusters, each of which represents a particular business technology in social commerce. The last phase evaluates BM clusters using citation-based indicators. The core BM clusters are then identified with the indicator values and classified based on their roles in knowledge flows. The detailed process is described in the following section.

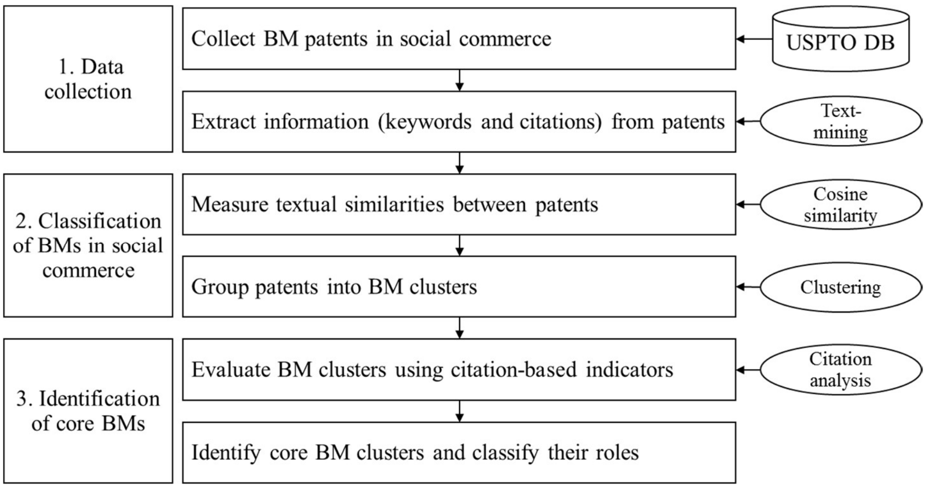


Figure 5.1 Overall research framework

5.3.2. Detailed process

5.3.2.1. Data collection

This study collects patents from The United States Patent and Trademark Office (USPTO) database. The USPTO database is the largest patent system in which patents from nearly all technological fields, including business technologies in e-commerce, have been accumulated for a long period of time (Ma and Lee, 2008; Wu, 2005). In the USPTO database, business method patents belong to 705 class entitled "Data Processing: financial, business practice, management, or cost/price determination." However, the USPTO classification system does not provide subclasses specifically assigned to social commerce BMs – that is, the classification system does not distinguish social commerce BMs from other e-commerce BMs. Due to this issue, it is not easy to collect BM patents that have major attributes of social commerce. Hence, we limit data collections to BMs based on social networks or communities to ensure that the collected patents are related to both social and commercial activities. These social commerce related patents are initially collected using search terms, such as "social commerce," "social network," "community-based" and so on. Irrelevant ones are then filtered out through a screening process.

There are largely two types of information need to be extracted from the patents: keywords and citations. The keywords are extracted from the patent abstracts which contain essential information about the patents. We use a popular text-mining method that is based on TF-IDF (term frequency-inverse document frequency) criteria. Since the TF-IDF method takes into account not only commonality but also uniqueness of terms (Lin et al., 2017), this can help eliminating terms that are too generally and widely used. Nevertheless, there may remain some terms that are not much informative or meaningful, so these terms need to be removed manually. The keywords will be used for grouping of patents with similar contents or subjects.

We use two different citations, namely backward and forward citations, in this analysis, and the corresponding data is obtained from the "Reference cited" section and "Referenced by" section of the patents. The citations will be used to evaluate BMs and identify core BMs among them.

5.3.2.2. Classification of BMs in social commerce

The collected patents need to be grouped into clusters with similar contents so that each cluster represents a particular business technology in social commerce. As explained in the previous section, 705 class does not provide subclasses specifically assigned to patents for social commerce. That is, grouping the patents based on the United States Patent Classification (USPC) system does not effectively identify BMs in social commerce, capturing their unique features.

To address this issue, we classify the patents into clusters based on textual similarities. Since textual data in patents contains the most comprehensive information that best characterize patents, patents with high textual similarities can be regarded as similar technologies. With the keywords extracted in the previous phase, the patents are represented as keyword vectors where each vector component indicates the frequency of the corresponding keyword. Textual similarities between patents are then measured using the cosine similarity which can effectively identify documents that have the same indexing term distribution within the each document. The cosine measure is also easy to interpret and simple to compute for long and sparse vectors (No et al., 2015). Once textual similarities between all patent pairs are measured, the values are presented in a patent similarity matrix. Next, we apply a matrix-based clustering technique to the patent similarity matrix for clustering. Matrix-based clustering, generally called DSM (Dependency Structure Matrix) clustering, finds subsets of DSM elements that are mutually exclusive or minimally related. After generating the clusters, we examine keywords that are closely related to social commerce activities to characterize and understand each cluster.

5.3.2.3. Identification of core BMs

In the last phase, we use patent citation indicators to identify core BMs from a knowledge flow perspective. Our framework distinguishes citations into four different types since each type of citation represents a distinct flow pattern of knowledge (Karvonen and Kässi, 2013). Figure 5.2 depicts these four types of citations: (1) backward citation within the BM patent class, (2) forward citation within the BM patent class, (3) backward citation beyond the BM patent class, and (4) forward citation beyond the BM patent class. Backward and forward citations here indicate knowledge absorption and diffusion. Patent classifications are used to delineate technological domains, so that any cited/citing patents in the BM class are considered internal knowledge of the BM domain those in different classifications are considered knowledge external to the BM domain. In this regard, we use four indicators, each of which corresponds to each type of citation, following the concept proposed by Ko et al. (2014).The formulas and descriptions for the indicators are summarized in Table 5.1. BMs are evaluated using these indicators and then core BMs are identified and classified based on their roles in knowledge flows (Table 5.2). The detailed descriptions of roles of core BMs are as follows:

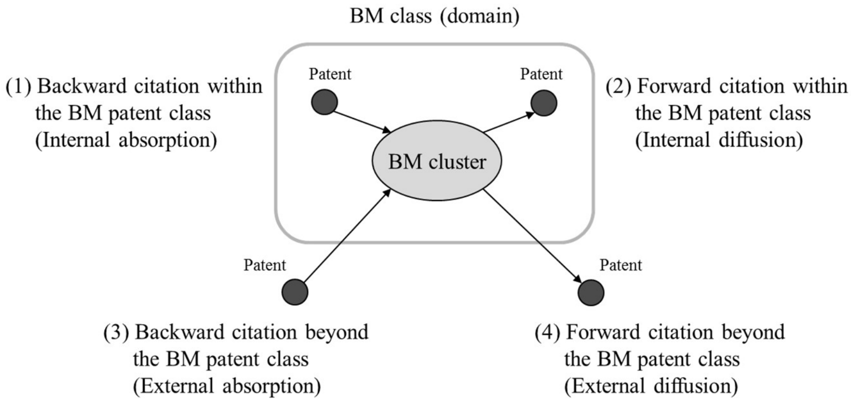


Figure 5.2 Types of citations and associated knowledge flow patterns

Table 5.1 Citation-based indicators

Indicator	Formula	Description
Internal absorption	$IA_i = \frac{\sum_{td(j)=BM} KF_{j,i}}{\sum_i \sum_j KF_{i,j}}$	The amount of knowledge a BM cluster absorbs within the its domain
Internal diffusion	$ID_i = \frac{\sum_{td(j)=BM} KF_{i,j}}{\sum_i \sum_j KF_{i,j}}$	The amount of knowledge a BM cluster diffuses within the its domain
External absorption	$EA_i = \frac{\sum_{c(i)\neq BM} KF_{j,i}}{\sum_i \sum_j KF_{i,j}}$	The amount of knowledge a BM cluster absorbs from other technological domains
External diffusion	$ED_i = \frac{\sum_{c(i)\neq BM} KF_{i,j}}{\sum_i \sum_j KF_{i,j}}$	The amount of knowledge a BM cluster diffuses to other technological domains

where $KF_{j,i}$: knowledge flow from patent j to BM cluster i (patent j is cited by a patent in BM cluster i), $td(j)$: technological domain (patent class) to which patent j belongs, BM: BM domain (705 class)

Table 5.2 Core BMs based on their roles in knowledge flows

Role	Description
Local knowledge utilizer	A high degree of internal absorption and a low degree of internal diffusion
Specialized knowledge supplier	A high degree of internal diffusion and a low degree of internal absorption
Knowledge consolidator	A high degree of internal absorption and diffusion
Knowledge boundary spanner	A high degree of external absorption and a low degree of external diffusion
General knowledge supplier	A high degree of external diffusion and a low degree of external absorption
Knowledge platform provider	A high degree of external absorption and diffusion

- Local knowledge utilizer: BMs in this group absorb a relatively large amount of knowledge within their domain. They are able to exploit knowledge existing in the same domain and generate mainly incremental innovations (Karvonen and Kässi, 2013).
- Specialized knowledge supplier: BMs in this group diffuse a relatively large amount of knowledge within their domain. They have technological dominance in their domain, exerting a significant impact on other BMs.

- Knowledge consolidator: BMs in this group absorb and diffuse a relatively large amount of knowledge within their domain. They can facilitate technological progress of BMs through internal circulation of knowledge (Ko et al., 2014)
- Knowledge boundary spanner: BMs in this group absorb a relatively large amount of knowledge beyond their domain. They have ability to recognize the value of new external knowledge, and then assimilate and utilize such knowledge, which can lead to a broadening of their knowledge base and improvement in innovation performance (Escribano et al., 2009).
- General knowledge supplier: BMs in this group diffuse a relatively large amount of knowledge beyond their domain. They can be considered a major invention since they can havesubstantial effects on, or can even spawn, innovations in other technological domains through knowledge transfer (Nemet, 2012).
- Knowledge platform provider: BMs in this group absorb and diffuse a relatively large amount of knowledge beyond their domain. They have access to a variety of knowledge sources and at the same time their knowledge is being applied broadly in many domains. Such inter-sectoral flows of knowledge plays a particularly important role in radical innovations (Nemet, 2012).

5.4. Empirical analysis and results

5.4.1. Data collection

The business method patents for social commerce were initially collected from 705 class of the USPTO database using the search terms. After removing irrelevant patents, we established a final dataset consisting of 264 business method patents (See. Since social commerce is a recent form of e-commerce, the collected patents are also relatively new, registered between 2000 and 2015.We extracted frequent terms from the final dataset using a java-based software and then manually removed irrelevant or insignificant terms. This text-mining procedure resulted in a keyword set composed of 227 terms. Lastly, a total of 5,140 backward citations and 4,361 forward citations were obtained.

5.4.2. Classification of BMs in social commerce

The 264 x 264 patent similarity matrix was constructed by measuring textual similarities between all pairs of patents, as shown in Figure 5.3. The cosine similarity generates values between 0 and 1, and the higher the values, the similar the patents are.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		**6141653**	**6161095**	**6266649**	**6332135**	**6336105**	**6338050**	**6584451**	**6963850**	**7006999**	**7072846**	**7076452**	**7082407**
2	**6141653**	1.000	0.201	0.034	0.901	0.935	0.993	0.102	0.219	0.000	0.000	0.152	0.031
3	**6161095**	0.201	1.000	0.019	0.134	0.200	0.200	0.000	0.181	0.000	0.000	0.201	0.156
4	**6266649**	0.034	0.019	1.000	0.023	0.038	0.034	0.000	0.266	0.051	0.218	0.046	0.130
5	**6332135**	0.901	0.134	0.023	1.000	0.899	0.894	0.116	0.136	0.022	0.000	0.081	0.021
6	**6336105**	0.935	0.200	0.038	0.899	1.000	0.928	0.096	0.218	0.067	0.000	0.161	0.062
7	**6338050**	0.993	0.200	0.034	0.894	0.928	1.000	0.101	0.217	0.000	0.000	0.150	0.031
8	**6584451**	0.102	0.000	0.000	0.116	0.096	0.101	1.000	0.075	0.064	0.101	0.087	0.030
9	**6963850**	0.219	0.181	0.266	0.136	0.218	0.217	0.075	1.000	0.121	0.162	0.203	0.161
10	**7006999**	0.000	0.000	0.051	0.022	0.067	0.000	0.064	0.121	1.000	0.116	0.000	0.104
11	**7072846**	0.000	0.000	0.218	0.000	0.000	0.000	0.101	0.162	0.116	1.000	0.052	0.000
12	**7076452**	0.152	0.201	0.046	0.081	0.161	0.150	0.087	0.203	0.000	0.052	1.000	0.063
13	**7082407**	0.031	0.156	0.130	0.021	0.062	0.031	0.030	0.161	0.104	0.000	0.063	1.000
14	**7174301**	0.180	0.142	0.026	0.092	0.193	0.187	0.049	0.142	0.000	0.119	0.789	0.035
15	**7184970**	0.106	0.249	0.104	0.058	0.166	0.105	0.000	0.355	0.290	0.000	0.162	0.258
16	**7209931**	0.076	0.151	0.011	0.040	0.060	0.075	0.000	0.234	0.034	0.000	0.121	0.094
17	**7254552**	0.034	0.057	0.400	0.023	0.091	0.034	0.000	0.264	0.226	0.059	0.068	0.387
18	**7254559**	0.171	0.220	0.029	0.076	0.176	0.170	0.164	0.225	0.000	0.132	0.817	0.059
19	**7343323**	0.247	0.109	0.012	0.154	0.240	0.261	0.031	0.136	0.000	0.000	0.099	0.102
20	**7363243**	0.099	0.230	0.022	0.053	0.118	0.098	0.000	0.154	0.197	0.034	0.099	0.102
21	**7409362**	0.258	0.180	0.041	0.182	0.216	0.256	0.000	0.140	0.030	0.000	0.082	0.000
22	**7469232**	0.000	0.000	0.027	0.000	0.000	0.035	0.000	0.000	0.000	0.000	0.000	0.000
23	**7487114**	0.061	0.121	0.037	0.033	0.081	0.060	0.047	0.113	0.054	0.042	0.853	0.101
24	**7512544**	0.200	0.128	0.012	0.134	0.184	0.206	0.000	0.087	0.000	0.000	0.092	0.048
25	**7571121**	0.000	0.035	0.391	0.000	0.014	0.000	0.101	0.436	0.303	0.366	0.021	0.109
26	**7580854**	0.200	0.128	0.012	0.134	0.184	0.206	0.000	0.087	0.000	0.000	0.092	0.048

Figure 5.3 Patent similarity matrix

The matrix-based clustering technique was then applied to the patent similarity matrix to generate clusters, each of which represent a particular business technologies. Among the generated clusters, small-sized clusters having less than three patents were excluded. We consequently obtained 17 clusters consisting of 121 patents in total (See Appendix B). The keywords in each cluster were divided into the two categories, and then clusters were examined and labelled based on the keywords, as shown in Table 5.3.

C1 (socially-relevant ads) generates advertisements, including contents, messages, stories and so on, for a user by employing social networking information of the user and various algorithms or models. The advertisements can also be targeted to social network contacts of the user. C2 (sharing information on social network platforms) enables a user to view shopping information of other users to whom the user is connected or to share item information on a network-based social platform. C3 (group-buying) is related to mechanisms for promoting sales by offering volume discounts on the condition that a minimum number of buyers would make the purchase. Group-buying is one of the mainstream business models in social commerce (Lee et al., 2016). C4 (gathering of rating data from social networking sites), allows suppliers to gather item scores evaluated and rated by participants of social networking sites. C5 (creation/distribution of contents by users) represents various methods supporting generation of contents by users and commercial use of contents on a network. C6 (social network based recommendation) includes methods for providing item recommendations to a user based on social networking information of the user. The recommendations or information on an item can also be provided by members of a social network associated with the user. C7 (social network enabled review system) represents review engines that include a social network engine. These engines can receive, store and retrieve reviews based upon the users' relationship to the authors of the reviews, as well as the subject. C8 (electronic word-of-mouth) facilitates social network members to communicate with and transmit information to others about advertisers, virtual advertisements, incentives and promotion. C9 (patient community management) includes methods and systems of healthcare management particularly for a community of patients. This allows service providers to collect or impart information to the community of patients, relating to treatment regimens. C10 (community based marketplace) creates a digital marketplace in which the community members can interact and facilitate commercial property transactions by exchanging accurate and standardized information. C11 (social networking services integrated with gift card services) integrates gift card services for mobile devices and social networking services so that social networking profiles of users can be used in providing services. C12 (online promotions through social network platforms) includes systems for online promotions integrated with social network platforms, providing viral features such as friend invite features and newsfeeds. C13 (community based negotiation) represents negotiation engines that create and a community for buyers and sellers having similar interests and promote interactions and negotiations among the participants. C14 (managing supply/demand data from community sites) retrieves and analyze information from a community of market participants for activities related to supply-demand management, such as planning and forecasting. C15 (gathering of community member data) collects member data from the community group associated with a user based on the members' social connection with the user and their lists of purchased items. C16 (transaction processing based on social networking) authenticates and processes requests of users received via social networking websites. C17 (interactive assistant on social network platform) employs branded virtual characters across multiple social network platforms to provide various services and information.

Table 5.3 Keywords of BM clusters

Cluster	Label	Keywords
C1	Socially-relevant ads	advertisement, target, profile, algorithm, predictive model, personal, display, criteria, personalized content, personalized message, item, social network contact, electronic catalog, customize, interest, interface, prefer, retrieve, social networking system, communicate, like
C2	Sharing information on social network platforms	information, view, item, list, profile, network-based social platform, share, communication, search
C3	Group-buying	social pricing, volume discount, incentive, group, mechanism, promotion, social network, interact
C4	Gathering of rating data from social networking sites	item, software, website, participate, product, promotion, incentive, product, social networking site, evaluate, rate, share, score
C5	Creation and distribution of contents by users	content, content network, online entertainment, create, stream, distribute, user-contributed, rate
C6	Social network based recommendation	select, online site, offering, member, associated, purchase, recommend, item, plurality
C7	Social network enabled review system	engine, receive, retrieve, store, privacy, database, sort, filter, social network engine, review engine, rank, rate, relationship
C8	Electronic word-of-mouth (e-WOM)	electronic commerce, advertiser, virtual advertise, coupon, mobile, promotion, incentive, group, information, media platform, transmit, communicate
C9	Patient community management	treatment regimen, information, compliance, pharmaceutical, protocol, assist, patient, cloud, community, interaction, refer, message, review

C10	Community based marketplace	commercial real estate, accurate, standardize, model, information, digital marketplace, transact, exchange, analytic, database, communicate, community, interact
C11	Social networking services integrated with gift card services	service, credit, gift card, member, mobile phone, social networking service, account, loyalty, profile, track, interest, message
C12	Online promotions through social network platforms	promotion, social network-based platform, widget, display, server, application, sponsor, incentive, prize, banner, survey, viral feature, invite, newsfeed, notify, message
C13	Community based negotiation	negotiations engine, buyer, seller, create, participate, sponsor, order, offer, database, process, transact, administrate, authoring, history, pricing, record, retrieve, track, community, interact, search, evaluate
C14	Managing supply/demand data from community sites	engine, market, data, database, supply, demand, logic, information, process, activity, contract, forecast, interface, manage, business, planning, retrieve, server, communicate, community, collaborate, search
C15	Gathering of community member data	item, group, data, list, member, associate, register, demand, community, communicate
C16	Transaction processing based on social networking	data, store, facility, authenticate, process, social networking website, processor, request, plurality, mobile phone, interface, communicate
C17	Interactive assistant on social network platform	network platform, virtual character, brand, service, product, advisor, transact, sell, insurance, educate, information, notify

5.4.3. Identification of core BMs

The BM clusters were evaluated using the four indicators (i.e., internal absorption, internal diffusion, external absorption and external diffusion) to decide which ones are the most significant in terms of knowledge flows (See Appendix C). Since only 121 patents were included in the BM clusters, we selected the associated backward (3,283) and forward citations (3,295) from the initial dataset. The indicator values for each cluster were then computed with the citation information and the results were visualized in maps, as shown in Figure 5.4 and 5. These maps were constructed in two dimensions, where the horizontal axis represents knowledge absorption and the vertical axis represents knowledge diffusion. It should be noted that we plotted data on log scales for better presentation. The BMs, represented by a circle, were positioned into four different quadrants, taking the mean values as a criterion. The size of a circle indicates the number of patents in the corresponding BM cluster. Using the positions of the BMs on the maps, core BMs were identified and their roles were classified. The overall results are summarized in Table 5.4.

We first investigated core BMs that play major roles in internal knowledge flows (Figure 5.4). The bottom-right quadrant of the map contains local knowledge utilizers. C1 (socially-relevant ads) belongs to this quadrant as it exploits a large amount of knowledge within its domain. On the opposite, BMs in the top-left quadrant are a specialized knowledge supplier. C13 (community based negotiation), C6 (social network based recommendation), C10 (community based marketplace) and C9 (patient community management) are included in this group. These BMs diffuse a large amount of knowledge within their domain although their size is relatively small or moderate. The top-right quadrant, which represents a knowledge consolidator, contains C3 (group-buying). C3 (group-buying) facilitates internal circulation of knowledge, exchanging a large amount of knowledge within its domain. Taking account that the map is in log scale, its position implies that it is especially capable of utilizing internal knowledge, compared to the other BMs.

Using the map in Figure 5.5, we identified and examined core BMs in terms of external knowledge flows. The bottom-right quadrant of the map contains BMs that absorb a large amount of knowledge beyond their domain. C3 (group-buying) and C1 (socially-relevant ads) are included in this quadrant and called a knowledge boundary spanner. General knowledge suppliers, which disseminate a large amount of knowledge to other domains, are positioned in the top-left quadrant of the map. C6 (social network based recommendation), C13 (community based negotiation) and C2 (sharing information on social network platforms) are found to be in this group. Finally, knowledge platform providers are included in the top-right quadrant. The knowledge platform providers, namely C9 (patient community management) and C10 (community based marketplace), exchange a large amount of knowledge beyond their domain.

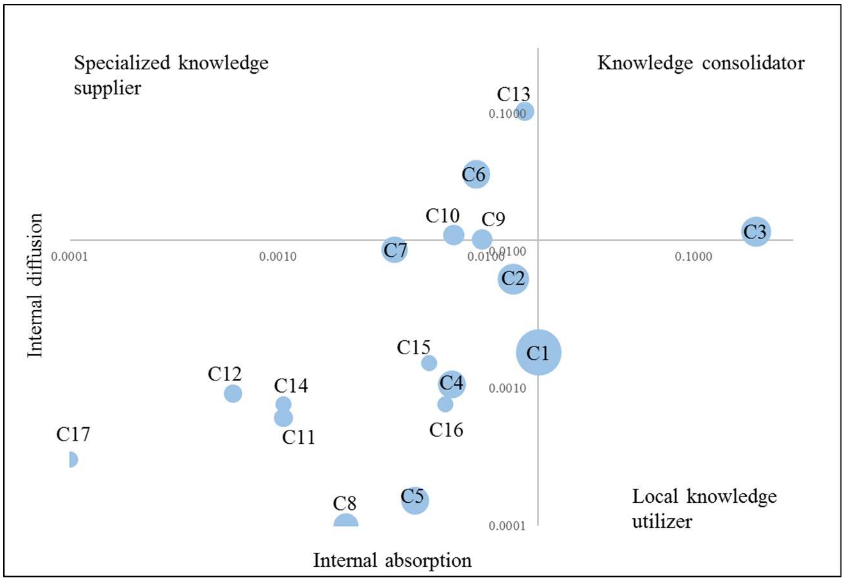


Figure 5.4 Positioning of BM clusters with relation to internal absorption and diffusion

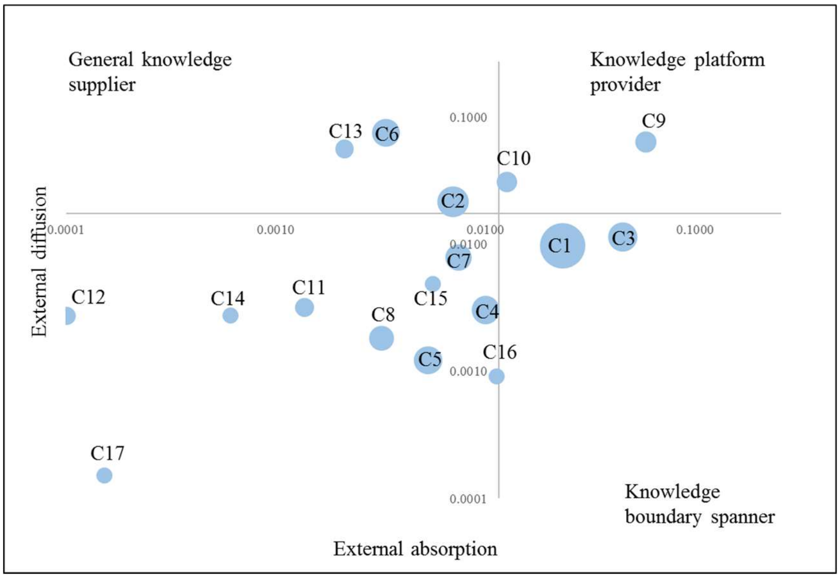


Figure 5.5 Positioning of BM clusters with relation to external absorption and diffusion

Table 5.4 Roles of core BMs

Role	Description	Core BM
Local knowledge utilizer	High IA-Low ID	C1
Specialized knowledge supplier	High ID-Low IA	C13, C6, C10, C9
Knowledge consolidator	High IA-High ED	C3
Knowledge boundary spanner	High EA-Low ED	C3, C1
General knowledge supplier	High ED-Low EA	C6, C13, C2
Knowledge platform provider	High EA-High ED	C9, C10

5.5. Discussion

5.5.1. Core BMs in social commerce

Our results exhibit that C1 (socially-relevant ads), C2 (sharing information on social network platforms), C3 (group-buying), C6 (social network based recommendation), C9 (patient community management), C10 (community based marketplace) and C13 (community based negotiation) are the core BMs in social commerce. C1 (socially-relevant ads) play two different roles, a local knowledge utilizer and knowledge boundary spanner, meaning that it has an ability to utilize both internal and external knowledge. Such 'learning' ability is correlated with productivity in R&D or innovation activities (Duguet and MacGarvie, 2005). Therefore, C1 (socially-relevant ads) is likely to be a more innovative area in social commerce and this may be the reason for the large number of patents in this cluster. On the contrary, C6 (social network based recommendation) and C13 (community based negotiation) score high on both internal and external knowledge diffusion. They are knowledge suppliers which have a considerable and pervasive effect across technological domains. C13 (community based negotiation) has a particularly dominant position in internal knowledge diffusion. C2 (Sharing information on social network platforms), classified as a general knowledge supplier, concentrates on distributing knowledge beyond its domain. That is, knowledge in this social commerce BM is applied in various technological domains. On the other hand, C9 (patient community management) and C10 (community based marketplace) are specialized knowledge suppliers which transfer a large amount of knowledge to other BMs. When it comes to external knowledge flows, these BMs play the role of knowledge platform providers. In particular, C9 (patient community management) exhibits the greatest openness for inter-domain knowledge flows. These knowledge platform providers enable a better penetration and diffusion of innovation across technological domains. C3 (group-buying), as a knowledge consolidator, can facilitate an overall technological progress of BMs through internal knowledge exchange. At the same time, as a knowledge boundary spanner, it explores new and different ideas about social commerce design, concepts and development by gaining access to external knowledge sources.

5.5.2. Methodological limitations and alternatives

One of the important steps in the proposed approach is to classify BMs in social commerce based on a textual similarity. The proposed approach uses the cosine similarity which measures similarity between term frequency vectors generated by the classic vector space model (VSM). In the VSM, the dimensions represent the terms in the context, and the component values represent their frequencies However, a limitation of the standard VSM is that it cannot cope with semantically related terms (Moen and Marsi, 2013). That is, it relies solely on matching the terms present in the documents.

Semantic similarity measures, such as Latent Semantic Analysis (LSA), Hyperspace Analogue to Language (HAL), Generalized Latent Semantic Analysis (GLSA), Cross-Language Explicit Semantic Analysis (CL-ESA) and Pointwise Mutual Information - Information Retrieval (PMI-IR), can be used to overcome such limitation. They are a widely used approach to the core problem of language understanding (Niraula et al., 2013). The most popular technique of semantic similarity is LSA that automatically derives meaning representations in the form of latent concepts by statistical computations applied to a large corpus (Bíró, 2009; Gomaa and Fahmy, 2013; Niraula et al., 2013). LSA uses dimensionality reduction as a means of accessing latent distributional similarities between terms (Moen and Marsi, 2013).

Latent Dirichlet Allocation (LDA) is also frequently used for measuring sematic similarity of texts. LDA belongs to the broader category of methods called topic models as it regards texts as distribution over topics, which are groups of semantically related words, and presents words as a vector of contributions to topics (Niraula et al., 2013). In this sense, LDA represents multiple meanings of a word explicitly while LSA does not (Niraula et al., 2013). Also, LDA has been proposed to address some limitations of the earlier technique, Probabilistic Latent Semantic Analysis (PLSA). PLSA is not a fully generative model, particularly at the level of documents, because there is no straightforward solution to assign probability to a previously unseen document (Bíró, 2009). Another issue is that the number of parameters to be estimated grows linearly with the number of training documents, which in turn will result in overfitting (Bíró, 2009).

These sematic similarity measures are often considered a more advanced technique that outperforms other term-based similarity measures. However, this is still controversial: some early studies suggested that LSA can improve results, yet recent studies suggested otherwise (Moen and Marsi, 2013). Likewise, comparisons between LSA and LDA regarding performance seems open to interpretation. One found that LDA-based methods and LSA-based methods yield competitive results (Niraula et al., 2013). The other, however, found that LDA is apparently superior to LSA (Bíró, 2009). This implies that a method should be carefully selected through comparison of different techniques.

5.6. Conclusion

Social commerce is a dominant trend in e-commerce, and firms look for ways to leverage social features for commercial gain. Since technology is mainly responsible for business processes, practices, and operations, technology-based BMs are becoming crucial enablers of new business models in social commerce. Therefore, comprehensive and objective analysis of BMs would provide valuable insights into the current innovations in social commerce. This paper proposes a framework for identifying core BMs in social commerce. Specifically, we define core BMs as the drivers of knowledge flows as they can play a critical role in the whole evolution process. We also perform empirical analysis, applying the proposed approach to BM patents from the USPTO.

We believe this study makes three significant contributions. First, this study focuses on the technological aspect of business models in social commerce. Despite its importance in the growth of social commerce, technology-based BMs have not been discussed in the existing social commerce studies. To our knowledge, this study is the first attempt to propose a systematic framework directed at investigating BMs in social commerce. Second, the proposed framework provides a means to objectively assess BMs in social commerce. BMs are evaluated using four knowledge flow indicators – internal absorption, internal diffusion, external absorption and external diffusion – that are based on patent citations. The patent citation data has widely used as an effective measure of knowledge flows in economics and business research. Third, this study classifies core BMs in terms of their roles in knowledge flows. Identifying core BMs merely based on the indicator scores does not provide much information. We further categorize roles of core BMs into six types – local knowledge utilizer, specialized knowledge supplier, knowledge consolidator, knowledge boundary spanner, general knowledge supplier and knowledge platform provider – since they have quite different implications. By refining their roles, we can better understand core BMs and can obtain more reliable and insightful information.

Despite the contributions, there are some limitations that should be noted and require further examination and additional research. First, our data is very limited due to difficulty in specifying an appropriate data retrieval query. The data is collected using some search terms, such as “social commerce,” “social network,” “community-based” and so on. However, there must be social commerce related patents that does not contain such terms in their abstracts, and thus a better data collection technique is required. Second, the indicators for assessing BMs in terms of knowledge flows can be further elaborated. Some examples can be indicators that are related to breadth of knowledge interactions, such as the number of patent classes from which a BM absorb knowledge and the number of patent classes to which a BM diffuse knowledge. Finally, knowledge network analysis or more micro-level analysis of BMs can improve our understanding of technological structure in social commerce.

Chapter 6. Conclusion

6.1. Summary and contributions

The overall objective of this study was to answer the questions related to exploring flows of technological knowledge in business model innovation using BM patents. The first research question was: *How can flows of technological knowledge in BM patents be measured?* Answering this question required finding a way to overcome a major drawback in the current methodology. Patent citation analysis, when used alone, cannot adequately estimate the actual extent of knowledge flows. To address this issue, this paper proposed an approach that integrates the patent citation analysis and text mining technique. The degree of knowledge flows through BM patents was measured by the proposed approach and it was found that BM patents actively participate in stimulating knowledge flows for business model innovation.

The second research question was: *What patterns of technological knowledge flows through BM patents can be identified?* This study applied a dynamic approach since static analysis cannot capture time-varying processes of knowledge flows in BM patents. An HMM was used with two different citation measures to identify diverse temporal patterns of knowledge flows in BM patents, and a clustering technique was then applied to find major patterns.

The third research question was: *How can BM patents be positioned based on the patterns of technological knowledge flows?* As a response to this question, this study proposed a systematic framework directed at investigating different roles BM patents in knowledge flows. It should be noted that the framework was specifically designed for investigating BM patents related to social commerce which is a dominant trend in e-commerce today. The empirical analysis classified BM patents into several clusters (or groups) based on their technical characteristics. The core clusters were then identified with citation-based indicators and their roles were specified according to the knowledge flow patterns.

This study expands the field of business model innovation research by linking the concept of business model innovation with technological development and knowledge flows. It presents a perspective on business model innovation that differs from, but also complements, those presented by the perspective focusing on economics and business strategy. Another important contribution is in the methodological use of patent indicators and other effective techniques in understanding business model innovation. Specific contributions of each topic are summarized in Table 6.1.

Table 6.1 Research contributions

Topic	Contributions
A structured approach to explore knowledge flows through technology-based business methods	Proposed a structured approach that integrates two complementary methods, patent citation analysis and text mining technique; Enhanced an understanding of the value and function of technology-based BMs in business model innovation
Identifying dynamic knowledge flow patterns of	Attempted to model dynamic patterns of knowledge flows of BMs without any

business method patents with a Hidden Markov Model	arbitrary hypothesis or bias by employing an HMM, which is a very flexible tool for characterizing various temporal patterns; Identified major knowledge flow patterns through clustering, which would help to understand overall trends in business technology
Identifying core business methods in social commerce from a knowledge flow perspective	Proposed a framework that provides a means to objectively assess BMs in social commerce based on patent citations; Specified several roles of core BMs in terms of knowledge flows

Based on the findings in this thesis, we can derive some implications for firm managers and policy makers. Firms can use BM patents and the approaches presented in this study clarify their current technological position in terms of business model innovation and formulate appropriate strategies. For example, a firm can compare knowledge flows of its BM patents with that of others. If its BM patents show relatively low-level flows, this indicates that both productivity in innovation activities of the firm and the value of its BM patents are likely to be low. In such case, managers should come up with right strategies to leverage external knowledge and develop BMs that are widely applicable. With the proposed approaches, firms also can identify potentially productive and valuable technologies involved in business model innovation and thus can make more informed investment decisions. In addition, the results from our study implies that policy for stimulating business model innovation should be made based on a concrete analysis of knowledge flow patterns in BM patents. For example, analyzing inter-domain knowledge flows can help policy makers to identify technological domains that are closely related to business model innovation. The policy makers then can propose more R&D cooperation stimulating policy towards these domains to increase efficiency of R&D and business model innovation as well.

6.2. Limitations and future research

This study has some limitations that should be addressed in future research. The scope of patent data used in each case study is limited to a specific BM patent area. The case study of the first research theme concentrated on BM patents related to postage metering system and that of the second theme used BM patents in 16 subclasses related to secure transactions. Data from a wider range should be used to understand an overall impact of technological knowledge flows in business model innovation. Also, the dataset for the third research theme is very limited due to difficulty in specifying an appropriate data retrieval query and thus a better data collection technique is required.

Second, the research methodologies in this study can be further elaborated. The first theme provided positioning maps of BM patents with relation to knowledge flows, yet they only showed a snapshot of the phenomenon. Creating maps for different time periods will present the evolving stages of knowledge flows brought about by BMs. The proxy measures of knowledge flows used in the second and third themes can be further specified depending on the focus of research. Moreover, the extensions of an HMM, such as a hidden-semi Markov model and mixed HMM, can be employed as they have some advantages over a simple HMM.

Finally, the three research themes presented in this study cannot answer all the questions related to technological knowledge flows in business model innovation. Future research should expand the scope of research related to knowledge flows in business model innovation to gain important and practical insights into technological trends in the business world. For example, this study mainly focuses on *ex post* analysis, but more *ex ante* analysis would help firms to forecast trends in business model innovation. In addition, further research is needed to understand more deeply the co-evolution process of business model innovation and technological innovation.

Bibliography

Allison, J. R., & Tiller, E. H. (2003). The business method patent myth. *Berkeley Technology Law Journal*, 987-1084.

Altman, R. M. (2007). Mixed hidden Markov models: an extension of the hidden Markov model to the longitudinal data setting. *Journal of the American Statistical Association*, *102*(477), 201-210.

Amit, R., & Zott, C. (2012). Creating value through business model innovation. *MIT Sloan Management Review*, *53*.

Anderson, M., Sims, J., Price, J., & Brusa, J. (2011). Turning "Like" to "Buy" social media emerges as a commerce channel. *Booz & Company Inc*, *2*(1), 102-128.

Autant-Bernard, C., Fadairo, M., & Massard, N. (2013). Knowledge diffusion and innovation policies within the European regions: Challenges based on recent empirical evidence. *Research Policy*, *42*(1), 196-210.

Baden-Fuller, C., & Haefliger, S. (2013). Business models and technological innovation. *Long range planning*, *46*(6), 419-426.

Baeza-Yates, R., & Frakes, W. B. (Eds.). (1992). *Information Retrieval: Data Structures & Algorithms*. Prentice Hall.

Baghdadi, Y. (2016). A framework for social commerce design. *Information Systems*, *60*, 95-113.

Bagley, M. A. (2000). Internet business model patents: Obvious by analogy. *Michigan Telecommunications Technology Law Review 7,* 253-288.

Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. V. (2013). Digital business strategy: toward a next generation of insights. *Mis Quarterly*, *37*(2), 471-482.

Bilmes, J. A. (1998). A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. *International Computer Science Institute*, *4*(510), 126.

Bilmes, J. A. (2006). What HMMs can do. *IEICE TRANSACTIONS on Information and Systems*, *89*(3), 869-891.

Bíró, I. (2009). Document classification with latent dirichlet allocation. *Unpublished Doctoral Dissertation*, *Eotvos Lorand University*, 4.

Blunsom, P. (2004). *Hidden Markov Models*. The University of Melbourne, Department of Computer Science and Software Engineering, Available online: http://www.cs.mu.oz.au/460/2004/materials/hmm-tutorial.pdf (accessed in September 2015).

Boulton, R. E., Libert, B. D., & Samek, S. M. (2000). A business model for the new economy. *Journal of Business Strategy*, *21*(4), 29-35.

Busalim, A. H. (2016). Understanding social commerce: A systematic literature review and directions for further research. *International Journal of Information Management*, *36*(6), 1075-1088.

Calado, P., Cristo, M., Gonçalves, M. A., de Moura, E. S., Ribeiro-Neto, B., & Ziviani, N. (2006). Link-based similarity measures for the classification of Web documents. *Journal of the American Society for Information Science and Technology*, *57*(2), 208-221.

Chang, S. B. (2012). Using patent analysis to establish technological position: Two different strategic approaches. *Technological Forecasting and Social Change*, *79*(1), 3-15.

Chang, S. B., Lai, K. K., & Chang, S. M. (2009). Exploring technology diffusion and classification of business methods: Using the patent citation network. *Technological Forecasting and Social Change*, *76*(1), 107-117.

Chesbrough, H. (2011). Bringing open innovation to services. *MIT Sloan Management Review*, *52*(2), 85-90.

Choi, C., & Park, Y. (2009). Monitoring the organic structure of technology based on the patent development paths. *Technological Forecasting and Social Change*, *76*(6), 754-768.

Crespi, G., Criscuolo, C., Haskel, J. E., & Slaughter, M. (2008). *Productivity*

Growth, Knowledge Flows, and Spillovers (No. w13959). National Bureau of Economic Research.

Curty, R. G., & Zhang, P. (2013). Website features that gave rise to social commerce: a historical analysis. *Electronic Commerce Research and Applications*, *12*(4), 260-279.

Damanpour, F., & Damanpour, J. A. (2001). E-business e-commerce evolution: perspective and strategy. *Managerial Finance*, *27*(7), 16-33.

Desyllas, P., & Sako, M. (2013). Profiting from business model innovation: Evidence from Pay-As-You-Drive auto insurance. *Research Policy*, *42*(1), 101-116.

Duguet, E., & MacGarvie, M. (2005). How well do patent citations measure flows of technology? Evidence from French innovation surveys. *Economics of Innovation and New Technology*, *14*(5), 375-393.

Dumont, M., & Tsakanikas, A. (2001). Knowledge spillovers through R&D networking. *Innovative Networks: Co-operation in National Innovation Systems*, *209*.

Érdi, P., Makovi, K., Somogyvári, Z., Strandburg, K., Tobochnik, J., Volf, P., & Zalányi, L. (2013). Prediction of emerging technologies based on analysis of the US patent citation network. *Scientometrics*, *95*(1), 225-242.

Escribano, A., Fosfuri, A., & Tribó, J. A. (2009). Managing external knowledge flows: The moderating role of absorptive capacity. *Research Policy*, *38*(1), 96-105.

Fink, G. A. (2014). Hidden Markov Models. In *Markov Models for Pattern Recognition* (pp. 71-106). Springer London.

Ghassempour, S., Girosi, F., &Maeder, A. (2014). Clustering multivariate time series using hidden Markov models. *International Journal of Environmental Research and Public Health*, *11*(3), 2741-2763.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1991). *Growth in Cities* (No. w3787). National Bureau of EconomicResearch.

Glenisson, P., Glänzel, W., Janssens, F., & De Moor, B. (2005). Combining full text and bibliometric information in mapping scientific disciplines. *Information Processing & Management*, *41*(6), 1548-1572.

Gomaa, W. H., & Fahmy, A. A. (2013). A survey of text similarity approaches. *International Journal of Computer Applications*, 68(13).

Grandon, E. E., & Pearson, J. M. (2004). Electronic commerce adoption: an empirical study of small and medium US businesses. *Information & Management*, *42*(1), 197-216.

Guan, J., & Chen, Z. (2012). Patent collaboration and international knowledge flow. *Information Processing & Management*, *48*(1), 170-181.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, 16-38.

Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Han, M. C., & Kim, Y. (2016). Can Social Networking Sites Be E-commerce Platforms?. *Pan-Pacific Journal of Business Research*, *7*(1), 24.

Healey, P., Rothman, H., & Hoch, P. K. (1986). An experiment in science mapping for research planning. *Research Policy*, *15*(5), 233-251.

Ho, M. H. C., Lin, V. H., & Liu, J. S. (2014). Exploring knowledge diffusion among nations: a study of core technologies in fuel cells. *Scientometrics*, *100*(1), 149-171.

Hu, A. G., & Jaffe, A. B. (2003). Patent citations and international knowledge flow: the cases of Korea and Taiwan. *International Journal of Industrial Organization*, *21*(6), 849-880.

Huang, Z., & Benyoucef, M. (2013). From e-commerce to social commerce: A close look at design features. *Electronic Commerce Research and Applications*, *12*(4), 246-259.

Huang, Z., & Benyoucef, M. (2015). User preferences of social features on social commerce websites: An empirical study. *Technological Forecasting and Social Change*, *95*, 57-72.

Huggins, R., & Johnston, A. (2010). Knowledge flow and inter-firm networks: The influence of network resources, spatial proximity and firm size. *Entrepreneurship & Regional Development*, *22*(5), 457-484.

Hunt, R. M. (2001). You can patent that? Are patents on computer programs and business methods good for the new economy?. *Business Review*, (Q1), 5-15.

Jaffe, A. B., Trajtenberg, M., & Fogarty, M. S. (2000). *The Meaning of Patent Citations: Report on the NBER/Case-Western Reserve Survey of Patentees* (No. w7631). National Bureau of Economic Research.

Janssens, F. (2007). *Clustering of Scientific Fields by Integrating Text Mining and Bibliometrics*, (Ph.D. thesis,) Faculty of Engineering, Katholieke Universiteit Leuven, Belgium. (Retrived from http://hdl. Handle.net/1979/847).

Janssens, F., Glänzel, W., & De Moor, B. (2008). A hybrid mapping of information science. *Scientometrics*, *75*(3), 607-631.

Janssens, F., Tran Quoc, V., Glänzel, W., & De Moor, B. (2006, October). Integration of textual content and link information for accurate clustering of science fields. In *Proceedings of the I International Conference on Multidisciplinary Information Sciences & Technologies (InSciT2006). Current Research in Information Sciences and Technologies. Volume I* (pp. 615-619). Springer.

Jim Wu, Y. C. (2005). Unlocking the value of business model patents in e-commerce. *Journal of Enterprise Information Management*, *18*(1), 113-130.

Jones, W. P., & Furnas, G. W. (1987). Pictures of relevance: A geometric analysis of similarity measures. J*ournal of the American Society for Information Science*, *38*(6), 420.

Karlsson, C., & Gråsjö, U. (2014). Knowledge flows, knowledge externalities, and regional economic development. In *Handbook of Regional Science* (pp. 413-437). Springer Berlin Heidelberg.

Karvonen, M., & Kässi, T. (2013). Patent citations as a tool for analysing the early stages of convergence. *Technological Forecasting and Social Change*, *80*(6), 1094-1107.

Kim, C., Kim, S., & Kim, M. S. (2011). Identifying relationships between technology-based services and ICTs: a patent analysis approach. *World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, *5*(12), 1948-1952.

Ko, N., Yoon, J., & Seo, W. (2014). Analyzing interdisciplinarity of technology fusion using knowledge flows of patents. *Expert Systems with Applications*, *41*(4), 1955-1963.

Kreng, V. B., & Tsai, C. M. (2003). The construct and application of knowledge diffusion model. *Expert Systems with Applications*, *25*(2), 177-186.

Laudon, K.C., & Traver, C.G. (2013) *E-commerce 2013: Business, Technology, Society*, 9 ed., Pearson Prentice Hall, New Jersey.

Lee, C., Kim, J., Kwon, O., & Woo, H. G. (2016). Stochastic technology life cycle analysis using multiple patent indicators. *Technological Forecasting and Social Change*, *106*, 53-64.

Lee, C., Kim, J., Noh, M., Woo, H. G., & Gang, K. (2017). Patterns of technology life cycles: stochastic analysis based on patent citations. *Technology Analysis & Strategic Management*, *29*(1), 53-67.

Lee, C., Park, H., & Park, Y. (2013). Keeping abreast of technology-driven business model evolution: A dynamic patent analysis approach. *Technology Analysis & Strategic Management*, *25*(5), 487-505.

Lee, H. J., Lee, S., & Yoon, B. (2011). Technology clustering based on evolutionary patterns The case of information and communications technologies. *Technological Forecasting and Social Change*, *78*(6), 953-967.

Lee, S., Lee, H. J., & Yoon, B. (2012). Modeling and analyzing technology innovation in the energy sector Patent-based HMM approach.

Computers & Industrial Engineering, *63*(3), 564-577.

Lee, Y. K., Kim, S. Y., Chung, N., Ahn, K., & Lee, J. W. (2016). When social media met commerce: a model of perceived customer value in group-buying. *Journal of Services Marketing*, *30*(4), 398-410.

Lee, Y. S., & Cho, S. B. (2011). Activity recognition using hierarchical hidden Markov models on a smartphone with 3D accelerometer. In *Hybrid Artificial Intelligent Systems* (pp. 460-467). Springer Berlin Heidelberg.

Leydesdorff, L. (1987, December). Words and Co-Words as Indicators of the Intellectual Organization of the Sciences. In *EASST workshop in Amsterdam* (*December 1987*).

Li, C., & Biswas, G. (1999). Temporal pattern generation using hidden Markov model based unsupervised classification. In *Advances in Intelligent data analysis* (pp. 245-256). Springer Berlin Heidelberg.

Liang, T. P., & Turban, E. (2011). Introduction to the special issue social commerce: a research framework for social commerce. *International Journal of Electronic Commerce*, *16*(2), 5-14.

Lin, X., Li, Y., & Wang, X. (2017). Social commerce research: Definition, research themes and the trends. *International Journal of Information Management*, 37(3), 190-201.

Liu, X., Jiang, S., Chen, H., Larson, C. A., & Roco, M. C. (2015). Modeling knowledge diffusion in scientific innovation networks an institutional comparison between China and US with illustration for nanotechnology. *Scientometrics*, *105*(3), 1953-1984.

Lukach, R., & Plasmans, J. (2002). Measuring knowledge spillovers using patent citations: Evidence from Belgian firms' data, CESifo.

Lukach, R., & Plasmans, J. (2005). International knowledge flows from and into a small open economy: Patent citation analysis. *Beyond Borders: Internationalisation of R&D and Policy Implications for Small Open Economies*, 331-357.

Ma, Z., & Lee, Y. (2008). Patent application and technological collaboration in inventive activities: 1980–2005. *Technovation*, *28*(6), 379-390.

Macdissi, C., & Negassi, S. (2002). International R&D spillovers: an empirical study. *Economics of Innovation and New Technology*, *11*(2), 77-91.

McGill, M., Koll, M., & Noreault, T. (1979). An evaluation of factors affecting document ranking by information retrieval systems, Final Report to the National Science Foundation. NSF-IST-78-10454.

Meagher, K., & Rogers, M. (2004). Network density and R&D spillovers. *Journal of Economic Behavior & Organization*, *53*(2), 237-260.

Mergner, S. (2009). *Applications of State Space Models in Finance: An Empirical Analysis of the Time-Varying Relationship Between Macroeconomics, Fundamentals and Pan-European Industry Portfolios*. Universitätsverlag Göttingen.

Moen, H., & Marsi, E. (2013, July). Cross-lingual random indexing for information retrieval. In *International Conference on Statistical Language and Speech Processing* (pp. 164-175). Springer Berlin Heidelberg.

Morris, J. W. (2014). Non-Practical Entities: Business Method Patents and the Digitization of Culture. *Critical Studies in Media Communication*, *31*(3), 212-229.

Mounteney, S. (2002). European Patent Office Update. Marks & Clerk, IP Europe, January.

Narin, F., & Olivastro, D. (1998). Linkage between patents and papers: An interim EPO/US comparison. *Scientometrics*, *41*(1-2), 51-59.

National Science Board (US). (2002). *Science & Engineering Indicators* (Vol. 1). National Science Board.

Nelson, A. J. (2009). Measuring knowledge spillovers: What patents, licenses and publications reveal about innovation diffusion. *Research Policy*, *38*(6), 994-1005.

Nemet, G. F. (2012). Inter-technology knowledge spillovers for energy technologies. *Energy Economics*, *34*(5), 1259-1270.

Ng, C. S. P. (2013). Intention to purchase on social commerce websites across cultures: A cross-regional study. *Information & Management*, *50*(8), 609-620.

Niraula, N., Banjade, R., Ştefãnescu, D., & Rus, V. (2013, July). Experiments with semantic similarity measures based on lda and lsa. In *International Conference on Statistical Language and Speech Processing* (pp. 188-199). Springer Berlin Heidelberg.

No, H. J., & Park, Y. (2010). Trajectory patterns of technology fusion: Trend analysis and taxonomical grouping in nanobiotechnology. *Technological Forecasting and Social Change*, *77*(1), 63-75.

No, H. J., An, Y., & Park, Y. (2015). A structured approach to explore knowledge flows through technology-based business methods by integrating patent citation analysis and text mining. *Technological Forecasting and Social Change*, *97*, 181-192.

Osterwalder, A. (2004). *The business model ontology: A proposition in a design science approach*, Institut d'Informatique et Organisation, 173, University of Lausanne, Ecole des Hautes Etudes Commerciales HEC, Lausanne, Switzerland.

Ovans, A. (2000). Can you patent your business model. *Harvard Business Review*, *78*(4), 16.

Pagani, M. (2013). Digital business strategy and value creation: Framing the dynamic cycle of control points. *Mis Quarterly*, *37*(2), 617-632.

Pateli, A. G., & Giaglis, G. M. (2005). Technology innovation-induced business model change: a contingency approach. *Journal of Organizational Change Management*, *18*(2), 167-183.

Plasmans, J., & Lukach, R. (2010). The Patterns of Inter-firm and Inter-industry Knowledge Flows in the Netherlands, CESifo Working Paper Series No. 3057.

Priego, J. (2003). A Vector Space Model as a methodological approach to the Triple Helix dimensionality: A comparative study of Biology and

Biomedicine Centres of two European National Research Councils from a Webometric view. *Scientometrics*, *58*(2), 429-443.

Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. In *Proceedings of the IEEE*, *77*(2), 257-286.

Rausch, E. (2003). International Patenting of Internet-Related Business Methods. *InfoBrief*, *NSF*, 03-314.

Salton, G., & McGill, M. J. (1983). Introduction to modern information Philadelphia, PA. American Association for Artificial Intelligence retrieval.

Sammarra, A., & Biggiero, L. (2008). Heterogeneity and specificity of Inter-Firm knowledge flows in innovation networks. *Journal of Management Studies*, *45*(4), 800-829.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461-464.

Spieth, P., Schneckenberg, D., & Ricart, J. E. (2014). Business model innovation–state of the art and future challenges for the field. *R&D Management*, *44*(3), 237-247.

Stephen, A. T., & Toubia, O. (2009). Explaining the power-law degree distribution in a social commerce network. *Social Networks*, *31*(4), 262-270.

Stephen, A. T., & Toubia, O. (2010). Deriving value from social commerce networks. *Journal of Marketing Research*, *47*(2), 215-228.

Trajtenberg, M., Henderson, R., & Jaffe, A. (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology*, *5*(1), 19-50.

Tsai, C. M. (2008). Integrating intra-firm and inter-firm knowledge diffusion into the knowledge diffusion model. *Expert Systems with Applications*, *34*(2), 1423-1433.

USPTO (2016). Manual of Patent Classification

Van Raan, A., & Tijssen, R. (1993). The neural net of neural network research: An exercise in bibliometric mapping. *Scientometrics*, *26*(1), 169-192.

Verbeek, A., Debackere, K., & Luwel, M. (2003). Science cited in patents: A geographic "flow" analysis of bibliographic citation patterns in patents. *Scientometrics*, *58*(2), 241-263.

Verspagen, B., & De Loo, I. (1999). Technology spillovers between sectors. *Technological Forecasting and Social Change*, *60*(3), 215-235.

Viterbi, A. J. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE transactions on Information Theory*, *13*(2), 260-269.

Vladimir, Z. (1996). Electronic commerce: structures and issues. *International Journal of Electronic Commerce*, *1*(1), 3-23.

Von Wartburg, I., Teichert, T., & Rost, K. (2005). Inventive progress measured by multi-stage patent citation analysis. *Research Policy*, *34*(10), 1591-1607.

Wagner, S. (2008). Business method patents in Europe and their strategic use—evidence from franking device manufacturers. *Economics of Innovation and New Technology*, *17*(3), 173-194.

Wang, C. C., Huang, M. H., & Chen, D. Z. (2012). The evolution of knowledge spillover and company cluster in semiconductor industry. *Journal of the Knowledge Economy*, *3*(2), 109-124.

Wang, C., & Zhang, P. (2012). The evolution of social commerce: The people, management, technology, and information dimensions. *Communications of the Association for Information Systems*, *31*(5), 1-23.

Wang, X., Zhao, Y., Liu, R., & Zhang, J. (2013). Knowledge-transfer analysis based on co-citation clustering. *Scientometrics*, *97*(3), 859-869.

Weill, P., & Woerner, S. L. (2013). Optimizing your digital business model. *MIT Sloan Management Review*, *54*(3), 71.

Wirtz, B., Göttel, V., & Daiser, P. (2016). Business Model Innovation:

Development, Concept and Future Research Directions. *Journal of Business Models*, *4*(1).

Wu, Y. C. J. (2005). Unlocking the value of business model patents in e-commerce. *Journal of Enterprise Information Management*, *18*(1), 113-130.

Yadav, M. S., De Valck, K., Hennig-Thurau, T., Hoffman, D. L., & Spann, M. (2013). Social commerce: a contingency framework for assessing marketing potential. *Journal of Interactive Marketing*, *27*(4), 311-323.

Yang, G. C., Li, G., Li, C. Y., Zhao, Y. H., Zhang, J., Liu, T., Chen, D.Z., & Huang, M. H. (2015). Using the comprehensive patent citation network (CPC) to evaluate patent value. *Scientometrics*, *105*(3), 1319-1346.

Zhang, J., & Korfhage, R. R. (1999). A distance and angle similarity measure method. *Journal of the Association for Information Science and Technology*, *50*(9), 772.

Zhang, J., & Rasmussen, E. M. (2001). Developing a new similarity measure from two different perspectives. *Information Processing & Management*, *37*(2), 279-294.

Zott, C., Amit, R., & Massa, L. (2011). The business model: recent developments and future research. *Journal of Management*, *37*(4), 1019-1042.

Zucchini, W., & MacDonald, I. L. (2009). *Hidden Markov Models for Time Series: An Introduction Using R* (Vol. 22). Boca Raton: CRC press.

Appendix

Appendix A. Social commerce patents

Subclass	Title of subclass	Patents
705/14.16	Referral award system	8504417, 8650070, 8688515
705/14.19	Giving input on a product or service or expressing a customer desire in exchange for an incentive or reward	8560385
705/14.25	Based on user history	8825523
705/14.27	Frequent usage incentive system (e.g., frequent flyer miles program, point system, etc.)	7184970, 8788332
705/14.33	Method of redeeming a frequent usage reward	8458016
705/14.39	Online discount or incentive	8595064
705/14.4	Advertisement	8452655, 8504421, 8583480, 8775247, 8799068, 8812360
705/14.41	Determination of advertisement effectiveness	8671019
705/14.49	Targeted advertisement	7870026, 7941339, 8504423, 8527344, 8751305, 9020835
705/14.5	Based on event or environment (e.g., weather, festival, etc.)	8131593, 8423409
705/14.52	Based on statistics	8682723
705/14.53	Based on user history	8489458, 8788340, 9009065
705/14.54	User search	8612293, 8615434
705/14.6	Based upon Internet or website rating	8768772

705/14.64	Wireless device	8386318
705/14.66	Based on user profile or attribute	8359237, 8438062, 8484083, 8600812, 8744911, 8942993
705/14.67	Personalized advertisement	8548855, 8554627
705/14.68	Period of advertisement exposure	9037486
705/14.73	Online advertisement	8589235
705/16	Including point of sale terminal or electronic cash register	8571937
705/2	Health care management (e.g., record management, ICDA billing)	6161095, 7904307, 8239215, 8290788, 8781866
705/26.1	Electronic shopping	7409362, 7752081, 7752082, 7756756, 7761342, 7761343, 7822646, 7933810, 7945482, 7970657, 7970660, 7970661, 7996270, 8001010, 8117080, 8140402, 8224707, 8266002, 8266007, 8275666, 8285598, 8355955, 8386329, 8392270, 8392271, 8401918, 8417577, 8438069, 8484089, 8484092, 8494914, 8494915, 8510172, 8548865, 8560397, 8589242, 8589247, 8606643, 8620765, 8666825, 8666826, 8666836, 8676661, 8706560, 8712861, 8775262, 8788358, 9020839, 9043227
705/26.25	Regulated	8055552
705/26.3	Auction	8762221, 8781913
705/26.35	Buyer or seller confidence or verification	7917406, 8548870
705/26.4	Request for offers or quotes	8521611, 8666842
705/26.41	Third party assisted	7865400, 8095430

705/26.43	Representative agent	7076452
705/26.44	Neutral agent	7254552
705/26.5	Item configuration or customization	8209238
705/26.61	Item investigation	7657458
705/26.62	Directed, with specific intent or strategy	7343323, 8566177
705/26.7	Item recommendation	7082407, 7571121, 7970665, 7974889, 8095432, 8224714, 8271352, 8433620, 8484098, 8626608, 8706566, 8738468, 8781915
705/26.8	List (e.g., purchase order, etc.) compilation or processing	9037503
705/27.1	Shopping interface	7881975, 8112324, 8244599, 8515832, 8630921, 8838484
705/27.2	Graphical representation of item or shopper	7487114
705/3	Patient record management	8015033, 8032399, 8650046
705/30	Accounting	7840457
705/300	Collaborative creation of a product or a service	8374972, 8719173
705/313	Real estate	7174301
705/319	Social networking	7707122, 7933843, 8175980, 8224756, 8311948, 8315953, 8326769, 8326770, 8473422, 8489515, 8504484, 8620828, 8650131, 8775323, 9037515
705/320	Human resources	8688595
705/325	Personal security, identity, or safety	8117133

705/347	Business establishment or product rating or recommendation	8510232, 8756168, 8990124
705/35	Finance (e.g., banking, investment or credit)	7945498, 7953654, 8160943, 8224727, 8229819, 8234193, 8386352, 8386353, 8396772, 8473386, 8538846, 8554655, 8589266, 8626627, 8635135, 8655762, 8660924, 8843406
705/36R	Portfolio selection, planning or analysis	7783547, 8676689
705/37	Trading, matching, or bidding	6584451, 7640204, 7672897, 8244623, 8301545, 9026471
705/38	Credit (risk) processing or loan processing (e.g., mortgage)	8458084, 8458085, 8732073, 8838498
705/39	Including funds transfer or credit transaction	8326752, 8423459, 8554670, 8577799, 8700526, 8738522, 8849714, 8918339
705/4	Insurance (e.g., computer implemented system or method for writing insurance policy, processing insurance claim, etc.)	8060386, 8788298
705/400	FOR COST/PRICE	8583564
705/41	Having programming of a portable memory device (e.g., IC card, "electronic purse")	8447690
705/42	Remote banking (e.g., home banking)	8688578
705/44	Requiring authorization or authentication	8571989, 8612351, 8671056
705/51	Usage protection of distributed data files	7254559, 7716136, 7783575, 8209261
705/52	Usage or charge determination	8892471

705/64	Secure transaction (e.g., EFT/POS)	7469232
705/7.11	Operations research or analysis	8214236, 8255248
705/7.14	Skill based matching of a person or a group to a task	8195498, 8380554, 8423392
705/7.16	Schedule adjustment for a person or group	8676626
705/7.25	Needs based resource requirements planning and analysis	7650294
705/7.29	Market data gathering, market analysis or market modeling	6266649, 6963850, 8024214, 8296175, 8332256, 8671012, 8856019
705/7.31	Market prediction or demand forecasting	7363243, 7512544, 7580854, 8793154
705/7.32	Market survey or market poll	7072846, 7827054, 8249915, 8321261, 8521580, 8744900, 8756097
705/7.33	Market segmentation	7689452, 8583471, 8600797
705/7.34	Location or geographical consideration	8209217, 8880421
705/74	Anonymous user system	7006999
705/80	ELECTRONIC NEGOTIATION	6141653, 6332135, 6336105, 6338050

Appendix B. Social commerce patent clusters

Cluster	Size	Patents
C1	24	8527344, 8583471, 8600797, 8768772, 8775247, 8781913, 8812360, 9020835, 8355955, 8706566, 9020839, 8484083, 8560385, 8589242, 8615434, 8438062, 8744911, 8266007, 8712861, 7870026, 7974889, 8504421, 8612293, 8799068

C2	11	7945482, 8001010, 8392270, 8392271, 8417577, 8438069, 8484098, 8560397, 9037503, 9043227, 7082407
C3	10	8140402, 8285598, 8401918, 8494914, 8494915, 8589247, 8620765, 7672897, 7970661, 8266002
C4	9	7689452, 8321261, 8521580, 8620736, 8626608, 8671012, 8744900, 8756097, 9037515
C5	9	7783575, 7827054, 8209261, 8311948, 8751305, 8423409, 8666826, 8688515, 8793154
C6	9	8489515, 8510232, 7756756, 7970665, 8271352, 8738468, 6266649, 8095432, 8386329
C7	8	7409362, 7657458, 7752081, 7752082, 7761342, 7761343, 7822646, 7881975
C8	7	8756168, 8775243, 8799060, 8825523, 8930236, 8452655, 8583480
C9	5	6161095, 8015033, 8032399, 8290788, 8650046
C10	5	7076452, 7174301, 7254559, 7487114, 7640204
C11	4	7953654, 8396772, 8554655, 8655762
C12	4	8229819, 8234193, 8538846, 8626627
C13	4	6141653, 6332135, 6336105, 6338050
C14	3	7512544, 7580854, 7650294
C15	3	7996270, 8095430, 8706560
C16	3	8160943, 8224727, 8386353
C17	3	8060386, 8635135, 8660924

Appendix C. Indicator values for clusters

Cluster	Internal absorption	Internal diffusion	External absorption	External diffusion
C1	0.0179	0.0018	0.0233	0.0097
C2	0.0135	0.0062	0.0070	0.0216

C3	0.1985	0.0138	0.0450	0.0114
C4	0.0068	0.0011	0.0100	0.0030
C5	0.0046	0.0002	0.0053	0.0012
C6	0.0090	0.0362	0.0033	0.0751
C7	0.0036	0.0102	0.0074	0.0079
C8	0.0021	0.0000	0.0032	0.0018
C9	0.0096	0.0122	0.0582	0.0638
C10	0.0070	0.0131	0.0126	0.0310
C11	0.0011	0.0006	0.0014	0.0032
C12	0.0006	0.0009	0.0000	0.0027
C13	0.0154	0.1040	0.0021	0.0561
C14	0.0011	0.0008	0.0006	0.0027
C15	0.0053	0.0015	0.0056	0.0049
C16	0.0064	0.0008	0.0112	0.0009
C17	0.0000	0.0003	0.0002	0.0002
Average	0.0178	0.0120	0.0116	0.0175

초 록

정보통신기술(ICT)이 새로운 비즈니스 모델 개발에 폭넓게 활용됨에 따라 비즈니스 모델 개념은 혁신 및 기술경영 분야에서 점차 중요한 연구 주제로 부상하고 있다. 인터넷 기술의 발전은 비즈니스 환경의 변화와 혁신을 주도하며 전자상거래 시대의 개막을 알렸다. 이에 따라 상당부분, 때로는 대부분의 상업 활동이 인터넷 플랫폼 상에서 가능해 지면서 정보통신기술은 새로운 비즈니스 모델 창출에 보다 핵심적인 요소가 되어가고 있다. 기업들은 비즈니스 모델 구현에 바탕이 되는 구체적인 방법 개발에 정보통신기술을 보다 적극적으로 활용하고 이를 특허로 보호하기 위해 많은 노력을 기울이고 있다. 이러한 특허는 Business model 특허 또는 Business method (BM) 특허로 불리며 디지털 기술과 소프트웨어 기술을 기반으로 하는 다양한 상업적 기법을 포함한다. BM 특허는 비즈니스 모델 혁신과 관련된 중요한 기술적 지식을 포함하고 있으며, 지식의 전달을 통해 비즈니스 모델 개발과 혁신을 촉진시킬 수 있다는데 큰 장점이 있다. 기술적 지식 흐름이 비즈니스 모델 혁신에 중요한 영향을 미침에도 불구하고 비즈니스 모델 혁신, 기술 그리고 지식 흐름 등의 세 가지 개념을 연계한 연구는 아직 미흡한 상황이다. 비즈니스 모델의 기술적 기반, 비즈니스 기술을 통한 지식 흐름, 비즈니스 모델의 정량적 분석 등의 연구 주제는 기존 연구에서 거의 다루어 지지 않고 있다. 따라서 본 학위논문의 목적은 BM 특허 분석을 기반으로 비즈니스 모델 혁신과 관련된

기술 지식의 흐름을 탐색하고자 한다.

본 논문은 세 가지 연구 주제로 이루어져 있다. 첫 번째 연구는 비즈니스 모델 혁신의 기술 지식 흐름을 측정하는 체계적인 방법론을 개발하고자 한다. 제시된 방법론은 특허 인용 분석과 텍스트 마이닝(Text mining) 기법 등 두 가지 상호보완적인 방법론을 통합하여 설계된다. 이를 활용하여 BM 특허로 인한 지식 흐름의 양을 측정하고, BM 특허가 비즈니스 모델 혁신에의 지식 흐름을 촉진시킨다는 점을 실증 분석을 통해 알아본다.

두 번째 연구는 비즈니스 모델 혁신의 기술 지식 흐름 패턴을 파악하고자 한다. 본 연구는 시간에 따라 변화하는 지식 흐름의 과정을 나타내고자 동적 접근법을 채택하였다. 구체적인 방법론으로 은닉 마르코프 모델(Hidden Markov Model), 특허 인용 분석 및 클러스터링(Clustering)기법을 활용하여 BM 특허를 통한 기술 지식 흐름의 시간적 패턴을 탐색한다.

세 번째 연구는 BM 특허의 역할을 비즈니스 모델 혁신의 지식 흐름 관점에서 다루고 있다. 본 연구는 소셜 커머스(Social commerce) 혁신을 위한 지식 흐름을 촉진시키는 BM 특허들의 역할을 체계적으로 분석할 수 있는 접근법을 제시한다. 특허 인용 데이터를 기반으로 하는 다양한 지표들을 활용하여 주요 BM 특허들을 파악하고 지식 흐름 패턴을 바탕으로 각각의 역할을 유형화한다.

본 학위논문은 비즈니스 모델 혁신의 개념을 기술적 발달과 지식 흐름과 연계하고 특허 인용 분석과 다양한 방법론을 활용한 체계화된 분석 방법론을 제시함으로써 비즈니스 모델 혁신의

기술적 측면에 대한 이해를 넓힐 수 있다는 점에서 의의를 갖는다.

주요어: 비즈니스 모델 혁신, 지식 흐름, 비즈니스 기술, BM 특허, 특허 인용 분석
학 번: 2013-30315