J. Eng. Technol. Manage. 38 (2015) 53–70

Contents lists available at ScienceDirect

Journal of Engineering and Technology Management

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

Development of a patent roadmap through the Generative Topographic Mapping and Bass diffusion model

Yujin Jeong [a], Keeeun Lee [a], Byungun Yoon [a,*], Rob Phaal [b]

[a] Department of Industrial & Systems Engineering, School of Engineering, Dongguk University, 26, Pil-dong 3-ga, Chung-gu, Seoul 100-715, Republic of Korea

[b] Centre for Technology Management, Institute for Manufacturing, Department of Engineering, University of Cambridge, 17 Charles, Babbage Road, Cambridge, UK

ARTICLE INFO

Article history:
Received 27 August 2014
Received in revised form 28 July 2015
Accepted 23 August 2015
Available online 1 September 2015

Keywords:
Patent roadmap
Generative Topographic Mapping (GTM)
Bass diffusion model
Patent planning and forecasting
Visualization

ABSTRACT

This paper aims to present a novel concept roadmap—the patent roadmap—and suggest an advanced patent roadmapping process, based on the Generative Topographic Mapping (GTM) and Bass diffusion model. The process for patent roadmapping is composed of two modules: Developing the GTM-based patent map and determining the appearance time of the emerging patent through Bass model. The result of this research is meaningful knowledge from analyzing a vast store of patent data with quantitative methods and automated tools. It can serve as an effective patent planning tool, and proposes a strategic Research and Business Development (R&BD) for both firms and governments.

1. Introduction

In the present age, intangible intellectual property has gained control of the global marketplace, and the world has entered an era of limitless competition through the development of intellectual property rights in the global marketplace through such means as "patent wars" (Chia, 2012; O'Hearn, 2000). The smartphone industry has become a big story and the international patent infringement trial between Apple and Samsung attracted public attentions. This battle lasted the longest, included the most amounts of products, happened in the most countries, and not finished yet. Apple accused Samsung of infringing on its patents covering the general appearances while Samsung sued Apple copied patents covering 3G networking, MP3 playback, and a method for recording a user's place in a gallery of images. Another battle was a battle between Research in Motion (RIM) who makes BlackBerry, and Nokia. It is not an issue of copying designs, but of using essential patents without paying royalties to the patent holder (Charlton, 2012). These trends make firms perceive the importance of acquisition and exploitation of intellectual property for their innovative invention. Accordingly, most of the enterprises actively tend to consider patent management methods such as a patent portfolio, and patent planning as key areas of business activities, making attempts to manage their intellectual property strategically. In this way, patent disputes have occurred internationally in frequent succession to defend their rights regarding new technologies and the establishment of a solid foothold in world markets such as the display and smartphone industries (Chia, 2012; Lloyd et al., 2011). Besides, the advent of patent trolling has accelerated to bring about patent wars, which shows that the value of intellectual property is superior to any actual assets.

* Corresponding author.
E-mail address: postman3@dongguk.edu (B. Yoon).

http://dx.doi.org/10.1016/j.jengtecman.2015.08.006
0923-4748/

For research institutes and technology development departments, it is effective to explore new products and novel marketplaces, beginning with patent planning and successive technology development—technology-driven planning. In particular, a well-planned patent in the R&D planning stage can play important role in preventing from imitation by competitors and lead to successful technology commercialization, resulting in making huge profits and creating additional value. Therefore, it is necessary to emphasize the importance of patent planning and forecasting rather than technology forecasting itself.

Patents are an output of the R&D activities of firms, research institutes, and universities including both technological specifications and commercial information such as enterprise information, assignee and inventor information, and cooperative relationships (Ernst, 1997). Since patents are a representative source for the changeover of technology developments in chronological order (Brockhoff, 1991), many studies have been conducted by utilizing the number of patent applications, citation relationships, claims, and patent families (Chen et al., 2010; Curran and Leker, 2011; Ponomarev et al., 2014) in different ways such as through trend analysis (Bigwood, 1997), network analysis (Choi and Park, 2009; Goetze, 2010; Verspagen, 2007; Wartburg et al., 2005), and principal component analysis (PCA) (Lee et al., 2009). Bigwood (1997) tracked the rate of publication of patents over time by utilizing trend analysis, and provided information concerning technological maturity and a corporation's technological strategies. Kash and Rycoft (2000) suggested a framework offering insight into the three distinctive patterns of innovation that are evident in the evolution of six technologies. The first one—the normal pattern—was represented as a stable network and repeated improvements of the same basic technology design through a self-organizing network (SON). Other patterns were transformation and transition patterns defined by new technologies and new networks. These new technologies might be an output of modifications to established technologies or the creation of fundamentally new designs. Wartburg et al. (2005) applied patent citation analysis for the measurement of inventive progress, thereby determining technological change, and contributing to the discussion with in-depth methodological reflection and the potential of citation network analysis for explaining technological changes. In particular, Roepke and Moehrle (2014) suggested the concept of technology DNA for characterizing technological fields by means of patent classifications and analytical coding.

However, previous studies have focused on existing indicators such as the number of citation and the number of patent application for identifying trajectory of technology development (Hidalgo and Gabaly, 2012). Besides, they have conducted research studies that are relevant to patent forecasting rather than planning. Thus, previous studies could forecast the trends of patent applications quantitatively or just explore potential patents or technology for a certain period, rather than performing stepwise patent planning in detail. Therefore, it is appropriate to establish a patent roadmap leading to a new patent planning and forecasting approach. Therefore, this study aims to suggest a patent roadmapping approach through the Generative Topographic Mapping (GTM) and Bass model. The suggested process consists of two modules—planning contents of promising patents and forecasting the appearance time of new patents. In planning the content of a patent, the GTM technique which is widely used as a visualization tool (Son et al., 2012; Jeong and Yoon, 2013), is applied to establish a patent map for the purpose of exploring patent vacuums. They are considered as promising nodes in patent roadmaps, and then, appropriate timing when they will be applied as patents is forecasted through the Bass model. In particular, the Bass model is chosen to consider factors influencing technology innovation because it is composed of several parameters reflecting imitation and innovation effects when emerging technology is adopted or diffused. The proposed approach depends on quantitative analysis with historical patent data and gains insight into patents that represents an ample source for analyzing technology and markets. Patent data serves as a critical source for identifying undeveloped areas and deriving contents of new patents in this study. Since existing patents are able to provide trajectory of technology development until now, it is necessary to analyze and review historical data despite of outdated data in order to avoid duplication with previous invention and extrapolate recent trends.

This research will suggest contents and timing of novel patent application through quantitative analysis and systematic framework, while examining vacuum technologies. Furthermore, it will become a baseline for strategic patent planning and technological developments for researchers in diverse research organizations and firms. For many enterprises and governments, patent roadmaps will serve as a tool for supporting resource allocation to promising areas for the commercialization of research and technology.

This paper is structured as follows: Section 2 contains review-related studies on patent planning, including patent roadmaps and both GTM and Bass models. The overall specific process to develop patent roadmaps will be explained in Section 3. In Section 4, a proposed methodology for patent roadmapping will be illustrated using mobile communication technology, and conclusions from this research will suggest directions for future study in Section 5.

2. Background

2.1. Patent planning and patent roadmap

The importance of patent planning has increased continuously in light of recent circumstances in which firms have struggled for patent infringement protection in the display and smartphone industries. In particular, the patent provides a legal right whereby no one can imitate the content of the inventions mentioned in the patents, so it is critical to occupy a valuable patent in advance to dominate the market and remain competitive. Moreover, the advent of patent trolling has raised both the value of intellectual property and the importance of planning for valuable patents. Nevertheless, a few studies have attempted to plan for patents and define the concept of the patent roadmap. In our previous research, the patent roadmap was a novel concept, defined as a tool that could not only forecast patents that will be registered in the near future, but plan the content of the prospective patents while determining the timing of their appearance in the market (Jeong and Yoon, 2014). Relatively little attention was paid to the concept of the patent roadmap including planning and forecasting, compared to the concepts of the technology roadmap and technology planning. Although a few have occasionally used the term "patent roadmap" (Keefe, 2013; Plotkin, 2014), there were no definitive tangible representation of that concept. There have been different perspectives for defining the meaning of patent roadmaps as a tool for analyzing the current state of technological development, preventing patent infringement between firms. Most have used the term "patent roadmap" in the context of technology planning and forecasting methods such as the development of a technology roadmap, not focusing on patents themselves and identifying technology development states and patterns (Chengjian and Lucheng, 2009). Chengjian and Lucheng (2009) roughly referred to a patent roadmap as a method for technology analysis, particularly describing the state of the development of technology based upon a large amount of patent information. In particular, they utilized a patent management map as one branch of a technology roadmap, and used that systematic method to analyze patent data in a different direction, including the total number of patents, the countries, competitors, cite index, etc., contributing some references to the strategies of China's Carbon NanoTube–Flat Panel Display (CNT–FED) R&D institutions. Although it developed a patent management map to identify patent development patterns and stages, it is insufficient for visualizing and forecasting future strategic directions in R&D activities for new technological development.

Another view of patent roadmaps is to come up with an effective counterplan for patent infringement by suggesting legal countermeasures that include reviews, such as prosecution and appeals (Keefe, 2013). A law firm can designate patent roadmaps as a strategic plan for patent protection based on a firms' technologies and goals, so it makes individual firms able to defend themselves against their competitors' patents (Plotkin, 2014). The other study was to establish the Institute for Information Technology Advancement (IITA) in South Korea, the patent roadmap of the Global System for Mobile Communication (GSM) for the purpose of observing the trends of GSM technology and investigating a counterplan for dealing with patent infringement. However, it was limited, not by patent planning and forecasting, but by patent infringements.

2.2. Generative Topographic Mapping (GTM)

The GTM model was first introduced by Bishop and Svensén (1998) as a probabilistic re-formulation of the self-organization map (SOM) (Bishop and Svensén, 1998; Bishop et al., 1998) in terms of using probabilistic methods, based on Bayesian theory. The SOM is an unsupervised neural network, mapping high dimensional input data onto a usually two-dimensional output space, while preserving the relations between the data points (Pampalk, 2001). Although the SOM inspired by neuroscience has been regarded as a well-known machine learning method, and has been employed for data clustering and visualization (Andrade et al., 2005), it has shortcomings when compared with GTM. The SOM does not include a cost function, and there is a shortage in the theoretical basis for choosing the learning rate parameter schedules and neighborhood parameters when trying to ensure topographic ordering. There is no general evidence of convergence—that is, the model cannot define a probability density (Bishop et al., 1998; Pampalk, 2001). In an attempt to overcome these drawbacks, GTM was suggested as a form of non-linear latent variable model for which the parameters of the model could be determined using the EM algorithm (Bishop and Svensén, 1998; Bishop et al., 1998).

The GTM has been utilized in a wide range of fields, because its probabilistic settings enable an extent for hierarchical structures (Tino and Nabney, 2002) and for novel topographic mapping methods such as the Topographic Products of Experts (ToPoE) and the Harmonic Topographic Mapping (HaToM) (Fyfe, 2007). The former was based on the products of experts, so they trained the weights to maximize the probability of the data under this product of experts, and the latter is based on the harmonic average to converge to better solutions than K-means or from the mixture of experts. Above this, the GTM technique effectively works in clustering and visualization (Grambmeier and Rudolph, 2002; Jeong and Yoon, 2013; Olier and Vellido, 2008; Son et al., 2012; Yang and Zhang, 2001), missing data imputation (Carreira-Perpiñán, 2000; Vellido, 2006), adaptive regularization (Vellido et al., 2003), modeling discrete data (Girolami, 2002), and detecting and handling robust outliers (Bullen et al., 2003). In particular, the GTM is used as one methodology for clustering and data visualization, since it allows more flexibility than PCA by adopting soft clustering through the responsibility of each data point. Meanwhile, the expression of the data structure by SOM is limited on a given node set by a winner-take-all selection method, and the relationship between the data and node sets is ambiguous (Pampalk, 2001; Vesanto, 1999). The GTM was applied for the clustering of motor unit action potentials and provided a method of visualizing serving as a machine-learning tool. Consequently, this research compared the performance of GTM with that of three other clustering methods (Andrade et al., 2005). The other study dealt with GTM in analyzing the web customer data and compared their relative merit on the clustering and visualization with other known methods such as SOM or PCA (Yang and Zhang, 2001).

Moreover, the GTM technique has been widely used as a method for patent mapping, and it was useful in identifying vacuums in an objective way, compared to the PCA, which has been regarded as a qualitative approach when finding vacuums and clusters. Son et al. (2012) proposed a GTM-based patent map that aimed to automatically detect and interpret technology vacuums, providing a grid-based two-dimensional map in which each patent was mapped to the relevant grid. Jeong and Yoon (2013) suggested a method that derived essential patents through GTM-based standard and patent maps. It included a systematic process that identified vacuums on a standard map in a specific technological field and enabled analysts to find candidates for promising essential patents, rather than relying on experts.

The GTM has been in the limelight because it is an effective method of visualizing information among bibliometric analysis tools, which have usually been used in the process of summarizing document characteristics into tables for statistical analysis (Kostoff, 1997; Morris et al., 2002). In particular, if a latent space is a two-dimensional space, then it can be used as a tool for data visualization, so high-dimensional data may be visualized in this lower dimensional space (Andrade et al., 2005). A two-dimensional GTM-based patent map is appropriate for visualizing technological information and enables the identification of the status of patenting activities. Furthermore, GTM algorithms can be extended to model time-series data by incorporating them as emission density, which is called a GTM through time (GTM-TT) model. It is the temporal version of GTM, and is comprised of hidden Markov models where the hidden states are given by the latent points in the GTM algorithm, and the emission probabilities are defined as the GTM mixture distribution (Bishop et al., 1997). A GTM-TT model is effective at analyzing multivariate time series exploration through clustering and visualization for simultaneous time series (Son et al., 2013). However, this research might as well use GTM rather than GTM-TT, because we aim to find novel patenting areas through GTM and forecast appearance times with a Bass diffusion model, and GTM-TT is more appropriate for investigating the transition of content for patents per time interval.

2.3. Bass diffusion model

There have been many attempts to describe phenomena such as market diffusion, technology substitution, and innovation (Bass, 1969; Bretschneider and Mahajan, 1980; Fisher and Pry, 1971; Mahajan, 1979; Mansfield, 1961; Norton and Bass, 1987; Pearl and Reed, 1920; Winsor, 1932). As a result, a diverse range of mathematical models have been produced under different assumptions, such as logistic models (Kucharavy and Guio, 2011), Gompertz (Winsor, 1932), Bass model (Bass, 1969; Norton and Bass, 1987), and Fisher-Pry (Fisher and Pry, 1971). The main findings can be summarized as the construction of bell-shaped curves, depicting the frequency of adoption against time and S-shaped curves, when the cumulative numbers of adopters are considered.

The Bass model is one of the most well-known and widely used models of first-purchase demand, and it was first introduced for capturing the growth of new durable products or innovations caused by the diffusion effect (Bass, 1969). It is a model for the adoption timing of an innovation, and has been central to subsequent developments (Norton and Bass, 1987). The Bass model has a behavioral rationale that is consistent with studies in social science literature on the adoption and diffusion of innovation, and is based on a simple premise about the hazard function (conditional probability that an adoption will occur at time t, given that an adoption has not yet occurred). Thus, set $f(t)$ as the probability of adoption at time t, or the fraction of the ultimate potential that adopts the innovation at time t, and $F(t)$ as the fraction of the ultimate potential that has been adopted by time t. Then the fundamental premise is that the likelihood of adoption at time t given that it has not yet occurred is:

$$\frac{f(t)}{1-F(t)} = p + qF(t)$$

It has three parameters: The coefficients of innovation and imitation, and the ultimate market potential, because there is an assumption that adoptions of the product are made by "innovators" and "imitators." The coefficient of innovation is denoted as p, and of imitation is q, and the ultimate potential is usually called m. The importance of innovators is greater at first, but diminishes monotonically with time, while the imitation effects increase with time (Norton and Bass, 1987). The above formulation and the corresponding adoption rate function are given by formulation as below:

$$f(t) = p + (q-p)F(t) - q[F(t)]^2$$

There are many applications of the Bass model in several areas such as marketing (Goldenberg et al., 2001; Mahajan et al., 1990; Tellis, 2007), industrial technology (Bardhan and Chanda, 2007; Cheng, 2012; Rao and Kishore, 2009), pharmaceuticals (Leigh and Yorke-Smith, 2011), and consumer durable goods markets (Bass, 1969). There are many applications of the Bass model in several areas such as marketing (Goldenberg et al., 2001; Mahajan et al., 1990; Tellis, 2007), industrial technology (Bardhan and Chanda, 2007; Cheng, 2012; Rao and Kishore, 2009), pharmaceuticals (Leigh and Yorke-Smith, 2011), and consumer durable goods markets (Bass, 1969). Although Bass model is not directly applied to implement technology roadmap or patent roadmap, it has been utilized to model a phenomenon of technology substitution under the circumstance where are "discontinuities" that one technology is replaced by another, following a time behavior pattern described by sequence of learning or S-shaped curves (Linton, 2002; Kim et al., 2009; Michalakelis et al., 2010; Lim et al., 2012; Song et al., 2015). Linton (2002) proposed an approach for forecasting disruptive and discontinuous innovation by reflecting differences caused by constant markets and learning curve effects with transformed Bass model. Michalakelis et al. (2010) suggested a methodology for describing innovation diffusion, exploiting the effects of generation substitution with dynamic ceiling and then evaluated the proposed methodology over historical data of 2G and 3G mobile services. Lim et al. (2012) also forecasted 3G mobile subscription by using a stochastic frontier analysis which was utilized to measure the relative market potential of 2G service, and the Bass diffusion model to estimate parameters and potential market size. Song et al. (2015) devised a new diffusion model that is called as the hybrid Bass-Markov model for a dual-type device-based service, which is offered by two different types of devices and markets. The environment of type I service is new to the market and the other has existed competing services. For example, these services (e.g. Long-Term Evolution) can be offered by two different types of devices, mobile phones and other emerging phones.

Among these attempts, most studies focused on forecasting the number of potential market while little has dealt with forecasting timing for replacement between incumbent technology and new one. Thus, this paper makes an effort to apply the Bass diffusion model to derive accurate appearance timing of nodes in the patent roadmap. Bass model will be utilized to estimate three parameters, specifically the ultimate potential of innovation as the maximum amount of patent applications when developing technology. Then, they will be fitted to a Bass cumulative curve, based upon the estimated parameters and dealing with the forecasting duration for technology development and the appearance time of future patents by regarding them as a distribution of patent applications with cut-off criteria.

3. Research framework

3.1. Basic concept

This paper aims to suggest a novel process for developing a patent roadmap, based on the GTM and Bass diffusion models, comprising of two modules—planning for the content of new patents through GTM, and forecasting their appearance time based upon the Bass model. At first, a two-dimensional patent map is developed by applying the GTM technique, and it enables the identification of the current state of patent applications while simultaneously pinpointing vacuums in the GTM-based patent map. All patent vacuums are evaluated for whether they are worth developing in the near future, and among them, the highly evaluated patent vacuums will be considered as candidates that have a strong likelihood of requiring intensive development. It is possible to derive the content of future patents and infer nodes in the patent roadmap. After planning novel patents, their appearance time is estimated from the Bass model, which is applied to forecast the maximum amount of patent applications in the technology field of interest. Since there is no accurate time point that stands as significant start and end points in technological development, this paper adopts a specific percentage of the distribution as the most important point in the timeline of technological development. The market potential, which is one of parameters in the Bass model, is regarded as the upper limit point when technology is developed. The parameters, including the ultimate potential and the coefficients of innovation and imitation enable the estimation of the growth curve, which is called the survival distribution of the technology in this paper, in a form of a probability table plotted against time.

3.2. Overall process

The overall roadmapping process is as shown in Fig. 1, and the framework is composed of three modules: (i) Developing the GTM-based patent map, (ii) Forecasting the appearance times of patents, and (iii) Developing the patent roadmap. In the first module, a patent map is developed by GTM and text-mining techniques, based on collected patent data, and vacuums are identified from the GTM-based patent map, resulting in the contents of new predicted patents and the titles of nodes in the patent roadmap. The second module attempts to forecast the appearance time of patents by the Bass model. This module aims to determine the start and end points of patent development, and then predict the appearance times of new patents. In other words, two modules are pursued to determine the title, length, and location at the roadmap of nodes, which are regarded as critical elements of the patent roadmap. After determination, existing patents are arranged in chronological order, depending on their application year, and new patents, which are inferred from vacuums in the GTM-based patent map, are put into the roadmap in accordance with the results of their forecasted appearance time. Simultaneously, the length and location of each node are adjusted by their appearance time and the duration of patent development derived from the second module.

3.3. Detailed process

3.3.1. Module 1: Developing a patent map based on GTM

Patents are a significant source of the acquisition of technical information, and the core subject of analysis, because they are the outcomes of R&D activity. Therefore, this paper utilizes patent documents from the USPTO patent database, including the titles of inventions, their abstracts, detailed description, claims, and citation information. After collecting patent documents, keywords are extracted by text mining techniques that convert unstructured data into meaningful structured data. As shown in Fig. 2, at first, all words are derived by truncating all sentences in patent documents and then unnecessary words including prepositions and conjunctions that have no relevance to analyzing the contents of the patents are removed. Second, words are arranged in accordance with occurrence frequency preferentially and then strained by the Term Frequency–Inverse Document Frequency (TF–IDF) method which exhibits how important a word is to a document in a collection. When extracting keywords, synonyms are treated as much as possible, adverting to both a general dictionary and Wordnet. Simultaneously, it requires that domain experts take part in a filtering process to increase accuracy and robustness about keyword extraction despite of automatic procedures that derive meaningful and representative words. Domain experts are able to suggest a cut-off value of occurrence frequency and TF-IDF and adjust a definitive list of keywords, considering importance of them in the field. Following that, a keyword vector is generated by whether a specific keyword has

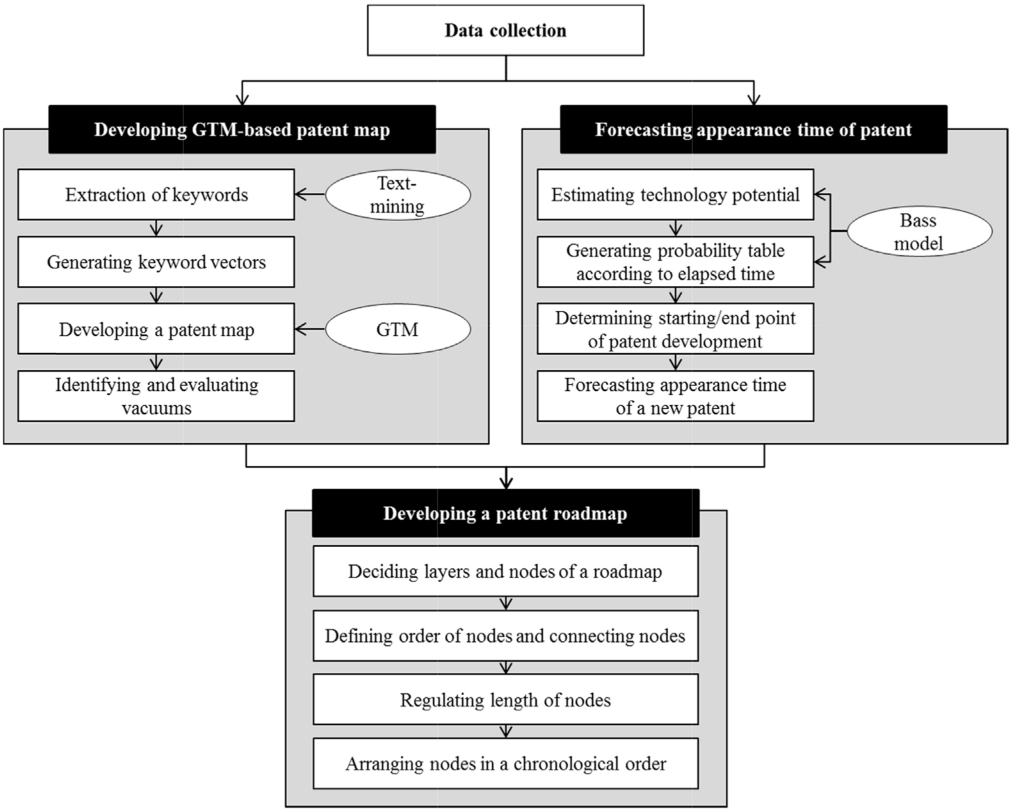

Fig. 1. Research framework.

occurred in the patent documents, as shown in Fig. 3. The first column is made up of keywords and the first row is comprised of collected patent documents. The intersection of matrix for keyword vectors is filled with binary values: 0 when a keyword has not appeared in a patent document or 1 when a keyword has occurred in a patent document, making it possible to discriminate the characteristics of patents by keyword vectors.

Depending on the generated keyword vector, the patent map is developed by utilizing the GTM that can visualize information in the form of a grid. The keyword vector generated in the prior step has high dimensionality, because each keyword corresponds to a dimension; for example, if there are 50 keywords, then each keyword vector has 50 dimensions. Although the setting of parameters corresponds to the choice of a prior distribution over possible models, which is generally a difficult task, the imperative is the determination, because the smoothness of the mapping is controlled by both the following parameters we have chosen. The basis function parameters are able to control the smoothness of the mapping, so more or narrower basis functions enable more flexible mapping, while fewer or broader basis functions make mapping smoother. They are selected to be radially symmetric Gaussians where the centers are distributed over a uniform grid in latent space (Svensén, 1999). The width parameter of the basis functions defines the distance between the basis functions; moreover, it also requires the selection of latent space sample points. If there are a few sample points with respect to the number of basis functions, the Gaussian mixture centers in the data become relatively independent, and the desired smoothness properties may become lost. However, a large number of sample points increases the computational cost, and there is one parameter to set for training, which is called the "weight regularization factor," which represents the degree of weight decay applied during training. A small degree of weight regularization is generally advisable in practice, as it prevents the weights from growing too large, since a finite number of latent and data points are used. Otherwise, smoothness imposed by the basis function parameter could result. Therefore, the GTM-based patent map has been developed as shown in Fig. 4.

After developing a patent map based on the GTM method, patent vacuums on the map can be identified and then, the value of patent vacuums needs to be evaluated. They can be considered in two ways: A vacuum that has the potential to be developed in the near future despite its current low occurrence versus a wasteland caused by its low value and technical infeasibility. Patent vacuums can be transformed to the original keyword vector through inverse mapping following identification, to provide the original meaning of the patent vacuum. The characteristics of inverse mapping that differentiate the GTM approach from other models enables projection from latent space into the data space (Bishop et al., 1998). Thus, a limitation of previous studies, the manual and subjective interpretation of patent vacuums can be eliminated

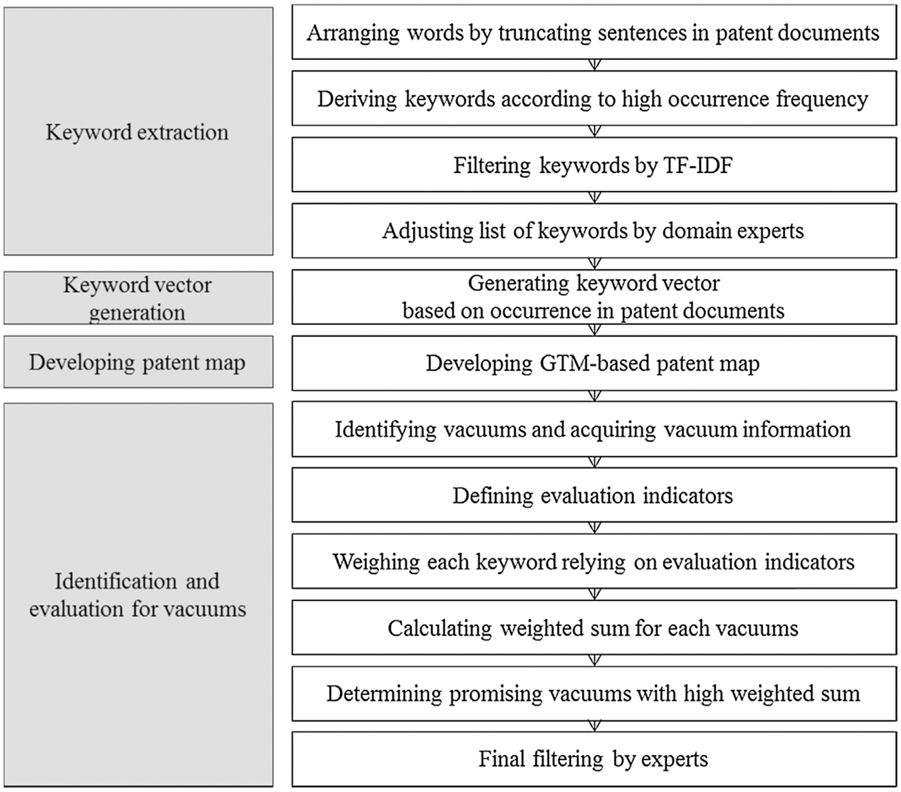

Fig. 2. A process for developing GTM-based patent map and vacuum evaluation.

by the automatic and objective interpretation. Consequently, keyword vectors are identified by inversely mapping patent vacuums in latent space into new vectors in data space, as illustrated in Fig. 3. Since there is no definitive method in determining a threshold value to include keywords, it can be decided depending on the purpose of research. In brief, if the threshold value is low, identified patent vacuums consist of many keywords (Fig. 5).

When evaluating these vacuums, it is necessary to reflect their technical characteristics as well as the novelty of technology to derive new technological content considered as nodes in the patent roadmap. In order to evaluate vacuums, indicators for assessing keywords that represent each vacuum are utilized in terms of their technical features and types of technology. Indicators to evaluate technical features are composed of interoperability with existing technology, promise, originality, and practicality. Keywords are also evaluated in accordance with the type of technology—base technology, applied technology, and technology for improving performance/effectiveness/efficiency. Considering how the keyword reflects the above features and corresponds to the specific type of technology, weights for keywords are measured with seven Likert scales, as explained in Table 1. Then, they are multiplied to vectors of keywords from extracted GTM results, and finally, the weighted sum of each vacuum can then be calculated. Vacuums with a weighted sum lower than the cut off value are considered wilderness and removed, the rest become the contents of nodes in the patent roadmap and can potentially be patented by researchers.

	Keyword *1*	Keyword *2*	Keyword *3*	…	…	Keyword *n-1*	Keyword *n*
Patent *1*	1	0	0	1	1	0	0
Patent *2*	0	0	1	0	0	1	0
Patent *3*	1	1	1	1	0	1	0
…	0	0	0	0	1	1	1
…	0	1	1	0	0	1	0
Patent *m-1*	1	1	1	0	0	0	0
Patent *m*	1	1	0	1	1	1	1

Fig. 3. An example of keyword vector.

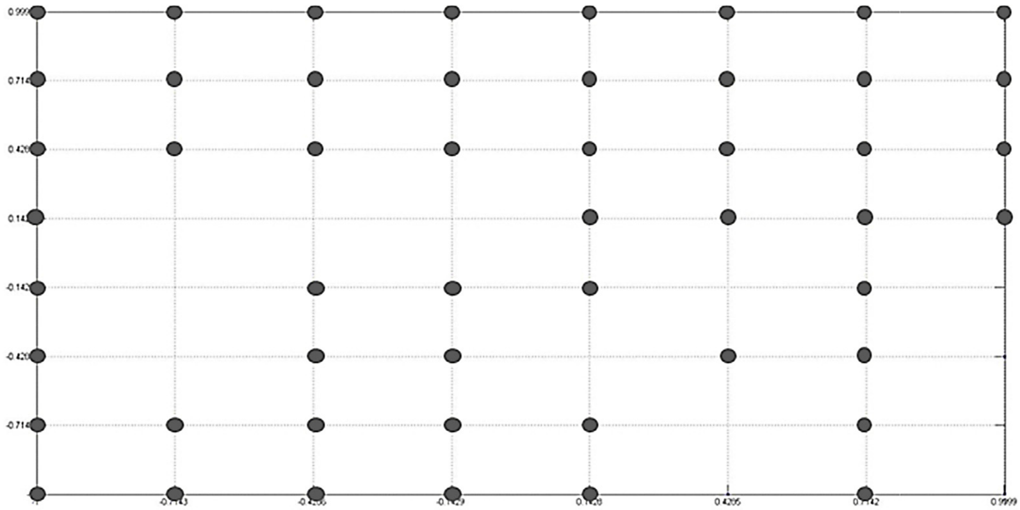

Fig. 4. An example of GTM result.

In order to validate the results of vacuum extraction, domain experts participate again to evaluate vacuums and are encouraged to judge value of vacuums by the different viewpoints – technical feasibility and market feasibility as shown in Table 2. Assessment of technical feasibility consists of status of patent application, potential for patenting, and applicability of technology, which aims to evaluate whether a patent corresponding to specific vacuums is able to be realized technically. Similarly, market feasibility seeks to evaluate markets into which a patent can enter by being embedded to technologies or products. At this time, market size and growth rate are evaluated, and a possibility to create new markets or expand market is assessed by expert's viewpoints. Like this, vacuums are scored by 5-point scale and then a weighted sum technique is applied in finally filtering vacuums out.

3.3.2. Module 2: Forecasting the appearance time of patents by the Bass model

In order to forecast the appearance time of new patent nodes, this research takes advantage of the Bass model that is one of the most widely used diffusion models, because it can predict the size of ultimate market as time passes. In other words, the results of estimation show the ultimate size that a specific market is able to grow based on historical data. In this way, this study can forecast appearance time and duration of technology development by analyzing remaining proportion to the ultimate size of technology after predicting the ultimate size of technology.

Since it might be a hasty generalization to regard the first patent application in a specific technology area as the starting point of technology development, it is critical to identify a starting as well as end time points in technology development. Besides, when developing technology and patents, there is no accurate time representing when technology has already been completely developed, unlike medical areas such as the birth and death of patients or other subjects. Thus, the Bass model is

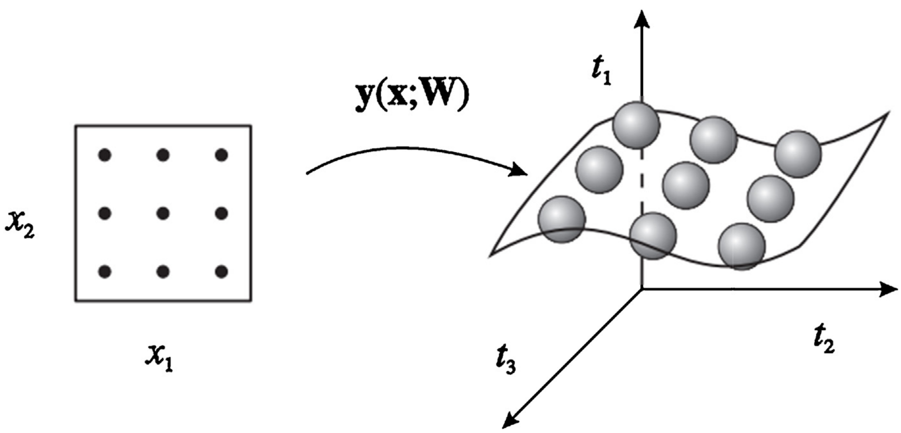

Fig. 5. Mapping vacuums in latent space into new vectors in data space (Bishop and Svensén, 1998).

Table 1
Indicators and Criteria for deciding weights of keyword.

Weight	Criteria for deciding weights of keywords						
	Technical index				Type of technology		
	3G/4G	Promising	Originality	Practicality	Enhancement technology	Applied technology	Base technology
7	0	1	1	1	1	1	1
6	0	1	1	1	1	0	1
5	1	1	0	1	0	0	1
4	1	0	1	1	1	1	1
3	1	0	1	1	0	1	0
2	1	0	0	1	1	0	0
1	1	0	0	0	0	0	1

Table 2
Criteria to filter vacuums in the secondary validation process.

Items	Sub-items	Contents of evaluation
Technical feasibility	Status of patent application	Whether similar patent exists for avoiding patent infringement and leading to avoidance design
	Potential for patenting	The extent of novelty and progressivity for patenting
	Applicability	The extent of application to different fields of technology
Market feasibility	Market size	The size of potential market which patent will be able to enter by applying to technology/product
	Growth rate	The growth rate of potential market which patent will be able to enter by applying to technology/product
	Possibility to create new market or expand market	The extent of creating new market or expanding existing market

selected to estimate survival probabilities and the ultimate market potential, which serves as the potential of technology development in this research, is estimated through patent applications, and it is regarded as the death point of the technology as shown in Fig. 6.

First, three parameters in the Bass model are estimated with the annual number of patent applications. Among them, the ultimate market potential is redefined as the upper limit of patent applications, and from that, the overall distribution of technology development can be forecast. There is an assumption that the number of patent applications represents the degree of advancements in technology. Then, the start and end points of technological development are determined by the estimated distribution, and they are deducted from the criteria percentage. In this research, this criterion is decided as 5% and 95%, since they are mostly utilized as confidence intervals applied to indicate the reliability of an estimate or significant test through statistics. The start point is the time at which significant patents for the development of a given technology

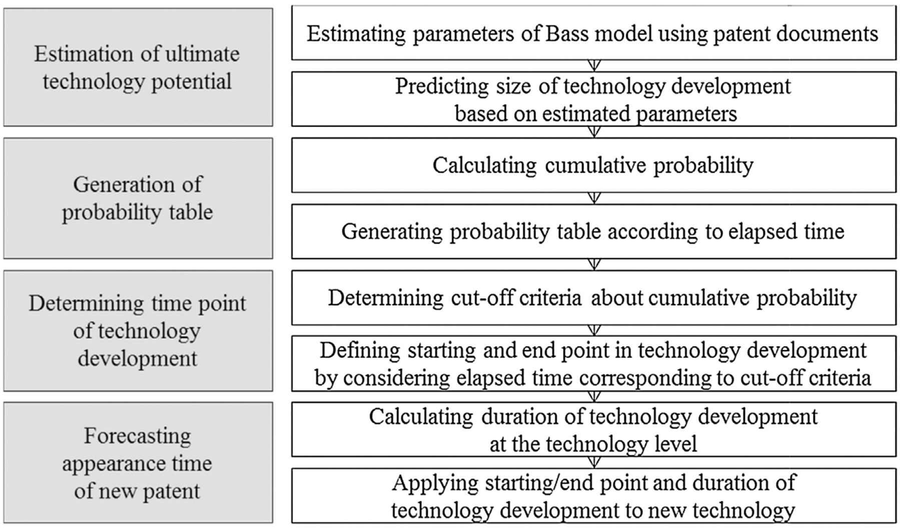

Fig. 6. A process determining appearance time of new patent.

begin to emerge, and the end point is the time at which the technological performance begins to contract and enter the declining stage in the technology lifecycle. As a result, the duration for technological development is set as time from when significant or meaningful developments appear in accordance with the elapsed time for reaching five percent distribution after the first patent application. Likewise, the end point of patent applications is determined from the elapsed time corresponding to ninety-five percent of the distribution. Consequently, the time gap between an incumbent patent and the emergence of a patent is extracted by applying the elapsed year, relying on a pre-determined duration for technology development in the prior steps.

3.3.3. Module 3: Development of the patent roadmap

It is necessary to define the layer, nodes, and times that are related to the location and length of nodes to develop the patent roadmap. As shown in Fig. 8, the results are derived from a GTM-based patent map and estimation by the Bass model, vacuum information, and duration of patent development, which decide these critical elements of the patent roadmap. Vacuums evaluated as promising and valuable patent areas become nodes in the patent roadmap, and each node is titled and inferred through keyword information involved in the corresponding vacuums with the aid of patent documents and relevant additional technical reports. Since it is necessary to define the order of patent nodes before arranging nodes and regulating the length of nodes, technological similarities are investigated by referring keywords, resulting in connecting the relevant nodes. Then, these nodes are preferentially arranged in chronological order according to the granted time of similar patents and the length and location of patent nodes are adjusted by duration of patent development.

3.4. Advantages of the proposed approach

Suggested roadmapping processes make it possible to forecast more accurate appearance times and the contents of promising patents with the aid of the Bass model and GTM quantitatively. While previous research has concentrated on the application time gap between relevant patents concerning incumbent technology and their application to new technology (Jeong and Yoon, 2014), the proposed approach intends to reflect patent development trends for predicting the appearance time of new patents using the Bass model, which was used to determine meaningful time points (start and ending point) when developing patents and technology. Moreover, this research has relied on not comparing analogous technology, but the prior generation of technology that is involved in the same technology platform. For this reason, it will become an alternative to solving disjunctions by analyzing analogous technology and applying it to the new technology. GTM has been widely utilized as a tool for information visualization, and its effectiveness has been established continuously. In this paper, GTM is also regarded as an effective technique for the detection of novel patenting areas by visualizing keyword information and clustering patents according to their use of similar keywords. Thus, this process helps plan and establish strategies for developing technologies and applying patents systematically, not depending on their subjective and time-consuming meeting. In other words, the proposed approach intends to take advantage of knowledge acquired from quantitative techniques and the huge amount of patent data that has been accumulated over a long time.

4. Illustration: patent roadmapping for mobile communication technology

4.1. Data

This research has chosen mobile communication technology to illustrate our proposed approach. It has evolved over the past three decades at a rapid pace and is making new leaps forward to the next generation—the 5th generation (5G)—of mobile communication technology. There have been many attempts to set up new standards for the next generation of technology and market it as their own technology in various countries with diverse organization and research institutes. This is in response to rising demand for quicker and more massive information exchange; thus, we have selected technology related to wireless communication—Wideband Code Division Multiple Access (WCDMA) and Long Term Evolution (LTE)—as they are representative standards for 3G and 4G mobile communication technologies, which have become winners among their competition. Based on 3G and 4G mobile communications, this study estimates the appearance time of 5G technologies and simultaneously plans for the content of requisite and potential patents when developing 5G mobile communication technologies, based upon the patented data relevant to 4G technologies. In particular, there are various types of nodes, since each telecommunication technology may comprise diverse elemental technologies. Thus, this research has made attempts to estimate 3G technology and the components of 4G technologies such as MIMO (Multiple-Input-Multiple-Output) and SDR (Software Defined Radio).

Our study has collected patent data relevant to mobile communication technology from the United States Patent and Trade Office (USPTO) database. Although many countries possess their own intellectual property offices, the USPTO database is commonly utilized when analyzing patents. The largest percentage of patent application is calculated from the USPTO database among national patent offices because most of the inventors tend to apply their inventions to USPTO after applying their homeland patent office by priority. Applying a patent to USPTO assures that markets of United States is secured and become a base of advance with exercising technical leverage. For the purpose of investigating the contents of patents, full-text documents including the title, abstract, claim, and description have been collected, and 11,791 patents related to 4G mobile communication technology were collected as shown in Table 3. Among them, patents that were not relevant to

Table 3
The number of collected patent for each technology.

Technology		The number of relevant patents	The number of irrelevant patents
3rd generation (3G)	WCDMA (Wideband Code Division Multiple Access)	1972	63
4th generation (4G)	MIMO (Multiple-Input Multiple Output)	730	51
	SDR (Software Defined Radio)	4273	133
	OFDMA (Orthogonal Frequency	571	34
	LDPC (Low Density Parity Check)	1902	329
	UWB (Ultra WideBand)	3329	561
	Smart antenna	986	89

mobile communication were removed. The 3G technology involved 1972 patents. Among them, 63 irrelevant patents were removed. In the case of 4G technology, MIMO technology included 51 irrelevant patents and SDR possessed 133 irrelevant patents, and UWB had the most irrelevant patents (561 patents) as following Table 3.

4.2. Results

4.2.1. The GTM-based patent map

Prior to developing our patent map, we extracted 49 keywords through text mining with consideration for their frequency and the characteristics of the technology with validation by domain experts. Two experts who are affiliated to a research institution related to telecommunication technology with more than five years of experiences in that field, took part in the validation process. After extracting the keywords and data preprocessing, GTM was applied to develop a patent map to identify vacuums. The parameters for constructing the GTM-based patent map were pre-defined through sensitivity analysis with adjustable reliability. Based upon five parameters, a two-dimensional patent map was established, as shown in Fig. 7. The map in Fig. 5 was composed of 13 by 13 grids which show the number of latent points, and the number of the basis function which controls the smoothness of the mapping was defined as 49. This parameter generally was chosen to be radially symmetric Gaussians in which the centers are distributed on a uniform grid in latent space. The width of the basis function which determines the distance between the basis function and is necessary to select latent space sample points was 2.0, and the weight regularization factor was 0.001, which sets for training and regulates the degree of weight decay applied during training. The GTM algorithm was repeated by 50 to make the map smooth and get better results. Both axes ranged from -1 to 1, since the projection of data was decided according to the posterior modes that represent maximum posterior probability estimation in Bayesian statistics. Cells that have no point are considered vacuums, and they may consist of either a singular vacuum or a vacuum that includes multiple vacuums, resulting in thirty cells of vacuum in the GTM-based patent map. Because the inverse function of GTM helps us extract keyword information, we could get information about the vacuum areas, and this will make it possible to forecast the contents of future patents, as shown in Table 4.

Then, all the keywords were given weights by their evaluation indicators, resulting in weighted sums, multiplying it with the value of keyword vectors involved in each vacuum. The case study of telecommunications selected the first quartile of the weighted sum for vacuums as a cut-off value, and four vacuum cells were removed. Since vacuums with relatively small weighted sum may not be considered as promising nodes in the roadmap, a criterion showing the biggest difference between

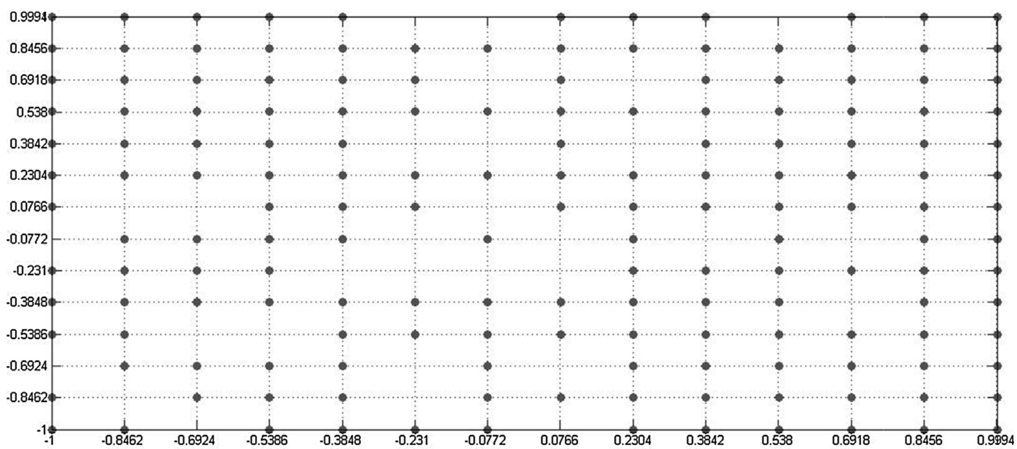

Fig. 7. GTM-based patent map.

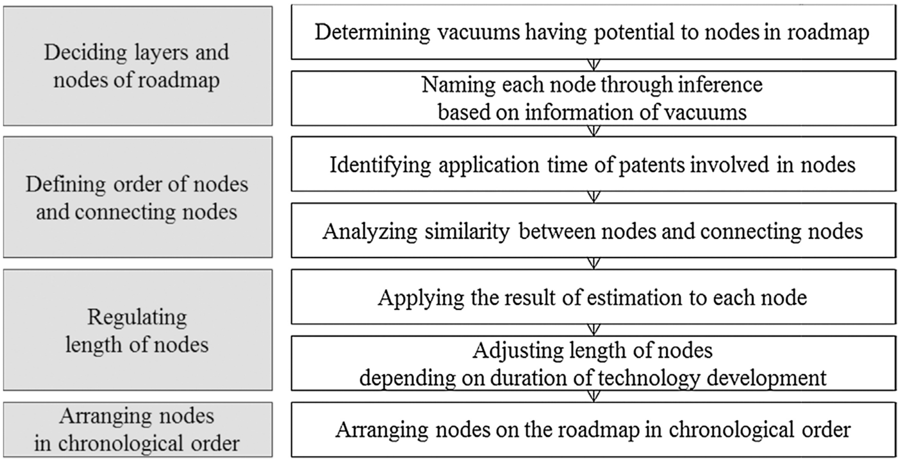

Fig. 8. A process establishing definitive patent roadmap.

Table 4
Keyword information involved in the vacuum cells.

Vacuum cell number	Keyword information
Cell 8	LTE, uplink, antenna, downlink, handover, eNB, scheduling, interference, PDCP, MAC, MME, modulation, OFDM, authentication, RRC, HARQ, backhaul, QoS, MIMO, RAN, multicast, reference signal, multiple access, method, apparatus, TDD
Cell 12	LTE, uplink, antenna, downlink, handover, eNB, scheduling, PDCP, MAC, MME, modulation, OFDM, MBMS, authentication, RRC, PDCCH, DRX, NAS, QoS, RAN, encryption, CRC, integrity, resource allocation, method, apparatus, TDD
Cell 21	LTE, uplink, antenna, downlink, eNB, scheduling, OFDM, RRC, RAN, reference signal, method, apparatus, TDD
Cell 27	LTE, uplink, downlink, handover, scheduling, PDCP, MAC, MME, modulation, OFDM, RRC, PDCCH, DRX, NAS, RAN, CRC, mobility management, resource allocation, method, apparatus
Cell 35	LTE, uplink, antenna, downlink, scheduling, OFDM, RRC, HARQ, reference signal, method, apparatus
Cell 39	LTE, downlink, handover, scheduling, PDCP, MAC, OFDM, RRC, PDCCH, HARQ, DRX, RAN, method, apparatus
Cell 42	LTE, OFDM, RRC, PDCCH, RAN, encryption, multicast, mobility management, method, apparatus
Cell 53	LTE, OFDM, RRC, Bluetooth, DRX, QoS, RAN, method, apparatus
Cell 65	LTE, modulation, OFDM, RRC, Bluetooth, RAN, WLAN, method, apparatus
Cell 71	LTE, uplink, antenna, downlink, interference, modulation, OFDM, Wifi, authentication, ASN, HARQ, MIMO, WLAN, encryption, multicast, DHCP, QPSK, multiple access, method, apparatus, interoperability
Cell 75	uplink, antenna, downlink, interference, modulation, OFDM, QoS, multiple access, method, TDD,
Cell 78	LTE, uplink, antenna, downlink, handover, scheduling, interference, modulation, OFDM, Wifi, Bluetooth, backhaul, RAN, method, apparatus
Cell 79	LTE, uplink, antenna, downlink, handover, scheduling, interference, modulation, OFDM, Wifi, Bluetooth, RAN, WLAN, method, apparatus
Cell 82	LTE, antenna, handover, eNB, PDCP, MAC, MME, RRC, QoS, RAN, WLAN, method, apparatus
Cell 83	LTE, uplink, antenna, downlink, handover, eNB, PDCP, MAC, MME, authentication, RRC, NAS, QoS, RAN, integrity, mobility management, method, apparatus
Cell 85	uplink, antenna, downlink, interference, MAC, modulation, OFDM, Wifi, ASN, MIMO, WLAN, multicast, DHCP, QPSK, multiple access, method, apparatus, interoperability
Cell 87	uplink, antenna, downlink, interference, MAC, modulation, OFDM, QPSK, multiple access, method, apparatus, TDD, interoperability
Cell 89	uplink, antenna, downlink, scheduling, interference, MAC, OFDM, Wifi, Bluetooth, WiMAX, method, TDD, interoperability
Cell 91	LTE, uplink, antenna, downlink, handover, scheduling, interference, modulation, OFDM, Wifi, Bluetooth, WiMAX, backhaul, QoS, RAN, method, apparatus, interoperability
Cell 93	LTE, uplink, antenna, downlink, handover, modulation, OFDM, Wifi, Bluetooth, WiMAX, backhaul, RAN, WLAN, method, apparatus
Cell 106	LTE, uplink, antenna, downlink, handover, modulation, Wifi, Bluetooth, WiMAX, backhaul, RAN, method, apparatus, interoperability
Cell 107	LTE, uplink, antenna, downlink, handover, OFDM, Wifi, WiMAX, backhaul, RAN, method, apparatus
Cell 110	LTE, uplink, antenna, downlink, handover, eNB, MME, RRC, RAN, WLAN, mobility management, method
Cell 112	LTE, uplink, downlink, handover, MME, OFDM, Wifi, authentication, Bluetooth, WiMAX, backhaul, NAS, RAN, WLAN, integrity, mobility management, method, apparatus, SGW
Cell 115	uplink, downlink, modulation, OFDM, QoS, QPSK, multiple access, method, apparatus, interoperability
Cell 117	uplink, downlink, scheduling, interference, modulation, OFDM, method, interoperability
Cell 134	LTE, uplink, antenna, downlink, method, apparatus
Cell 141	antenna, scheduling, interference, modulation, OFDM, Wifi, Bluetooth, QoS, MIMO, WLAN, SDR, UWB, method, apparatus, power amplifier
Cell 162	LTE, method, apparatus,
Cell 164	LTE, ASN, method

upper vacuums and others is chosen to filter vacuums out. In particular, vacuum cells that were adjacent to other cells are grouped as one vacuum, on the basis that they could have similar content through calculating the cosine similarity between the keyword vectors. All vacuums are extended to the content of nodes in the roadmap by the literature and interviews with domain experts, simultaneously searching research papers, patents, or trend reports through combining keywords involved in those vacuums in Table 5.

4.2.2. Prediction of the appearance time and duration of patent developments

From the collected patent data, we have estimated parameters, including the maximum of the potential in technology development (m), innovation coefficient (p), and imitation coefficient (q), as shown in Table 6. Estimated parameters are applied to predict the size of technology development by elapsed time and provide information about the applicable amount of patents in technology fields through simulation as time passes as shown in Table 6 and Fig. 9. Because the technologies in telecommunications are considered a rapidly growing area, they have slimmer time intervals. Since this research sets the time interval as quarter of a year, it enables to predict the size of technology development and the length of nodes in the patent roadmap by the quarter. Thus, it means that it is possible to forecast time more concretely by the quarterly period than by the yearly period.

In the case of 3G mobile communications, the time for the development of significant technology is approximately 8.5 years after the first patent application resulting from five percent of the distribution, as shown in Fig. 9. Similarly, the status of patent development can reach ninety-five percent in the Bass diffusion distribution after a lapse of about three-quarters, or 21 years. In short, 3G telecommunication technologies first appeared in the first quarter of 1996, which was 8.5 years after the first patent application in the third quarter of 1987, and has completed in the first quarter of 2009 applying the elapsed time, 21.5 years, as shown in Fig. 9.

As mentioned above, 4G telecommunication consists of core element technologies—Multiple-Input-Multiple-Output (MIMO), Software Defined Radio (SDR), Orthogonal Frequency Division Multiplexing/OFD Multiple Access (OFDM/OFDMA), Low density Parity Check (LDPC), Ultra WideBand (UWB), and smart antenna. We have estimated the maximum technological potential of all categories of 4G telecommunication in Table 6 and Fig. 5. UWB technology includes all kinds of technology related to communication and networks in a wide bandwidth range, and can transmit in a manner that does not interfere with conventional narrowband and carrier waves used in the same frequency bands. The relevant patents were acquired in the late 1980s in military-related areas, and they have grown since the late 1990s. Patenting activities have continued since the first quarter of 2002, and will continue to the first quarter of 2024, because UWB technology has a wide range of applicable areas in diverse industries such as mobile communications and military and radar areas.

There was an attempt to apply patents related to SDR technology in the second quarter of 1994, and more patents that are meaningful have been applied from the second quarter of 2005, which is a time interval corresponding to five percent of the estimated probability distribution, continuing until the second quarter of 2019. SDR technology facilitates the implementation of some functional modules in radio systems, such as modulation/demodulation, signal generation,

Table 5
Patent vacuum information and title of nodes in the roadmap.

Vacuum	Corresponding cells	Weighted sum	Title of node
V1	C8	90.05406066	Multiplexing including reference-signal structure (UL/DL)
V2	C12	80.73789708	Flexible DTX and DRX in a wireless communication system
V3	C21	48.83390018	Method of transmitting scheduling reference signal
V4	C27	46.76132393	PDCCH Resource allocation
V5	C35	32.61829272	Smart antenna
V6	C39	44.88802081	LTE power saving, Battery consumption
V7	C42	50.51685664	Method of receiving multimedia broadcast/multicast service in cell-based wireless communication system
V8	C65, C78, C79, C93, C106, C107	46.60385399	Accessing Cloud Computing Resources over 4G LTE
V9	C71, C85	68.49135589	Multi-user resource allocation and medium access control (MAC) overhead reduction for mobile worldwide interoperability for microwave access (WiMAX) systems
V10	C75, C89	43.25060656	LTE-TDD (Time Division Duplex)
V11	C87	40.89232207	OFDMA (OFDM/FDMA)
V12	C91	52.77293445	Cooperating base station set selection and network reconfiguration in limited backhaul networks
V13	C82, C83	38.05228278	Radio Resource Control and mobility robustness
V14	C110	30.39518834	mobility and handoff management-V2V
V15	C112	56.46827728	SGW (Serving Gateway)-mobility management
V16	C115	29.83517017	SCMA (Sparse Code Multiple Access) for downlink multiple access of 5G networks
V17	C117	33.50386415	uplink/downlink scheduling algorithm
V18	C141	48.5758716	SDR minimizing latency

Table 6
Estimated parameters in bass model.

Parameters	3G mobile communication	4G mobile communication					
		MIMO	SDR	OFDMA	LDPC	UWB	Smart Antenna
m	5630.6438	6414.2763	1710.8499	1295.3271	4036.3718	7260.6019	1783.9824
p	0.0003193	0.0002489	0.0000724	0.0015432	0.0001607	0.0000893	0.00046581
q	0.1123193	0.1722489	0.1050724	0.1765432	0.1071607	0.0660893	0.06446581
Significance level	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*

* Significant ($p < 0.001$).

coding, and link layer protocols in software. Since SDR technology has key features such as reconfigurability, ubiquitous connectivity, and interoperability, it can be used in a wide range of radio applications such as Bluetooth, WLAN, GPS, Radar, WCDMA, and GPRs, and will be extended to both 4G and 5G telecommunication.

4.2.3. Patent roadmap

A patent roadmap consists of nodes, time, and layers, and they are completed by the content of potential patents that result from a GTM-based patent map and the start and end points of technology development in 3G and 4G telecommunication areas by the Bass model. The proposed patent roadmap has two layers: Technology and patents. The technology layer is filled with telecommunication technology in each generation, and they are connected with nodes that have potential for patenting in areas in the patent layer, according to the interrelationship between technology and patents. Nodes are named after searching and the combination of keyword information involved in vacuum cells that are drawn from a patent map by GTM, as shown in Table 5. All nodes are arranged by a sequence of patent developments in chronological order, with consideration for the time gap between existing technologies estimated through the Bass model; thus, a definitive patent roadmap can be established, as shown in Fig. 11. In particular, a rectangular node means that patents have been actively applied in relevant areas, and a sharp-pointed rectangular patent node with solid line is one that has been implemented relatively recently and will conduct patent applications continuously in the future on a more active scale than previously (e.g. node "SDR maximizing latency", "OFDMA"). Nodes with dotted lines represent novel patenting areas and requisite fields to the implementation of 5G mobile communication, based on the existing 3G and 4G telecommunication. The nodes "SDR minimizing latency" and "Smart antenna" are uniquely placed in the patent roadmap pertaining to the corresponding technology, SDR technology, and Smart antenna technology, so that they follow the duration for the development of patents resulting from the Bass model. In Fig. 10, it is

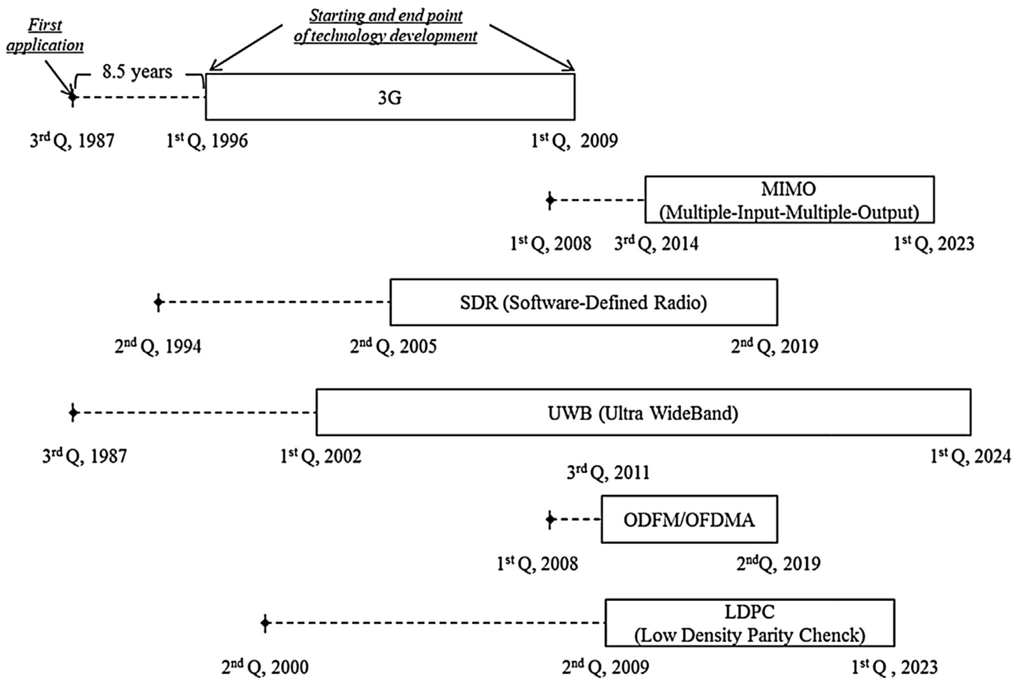

Fig. 9. The result of estimating appearance time in previous technology.

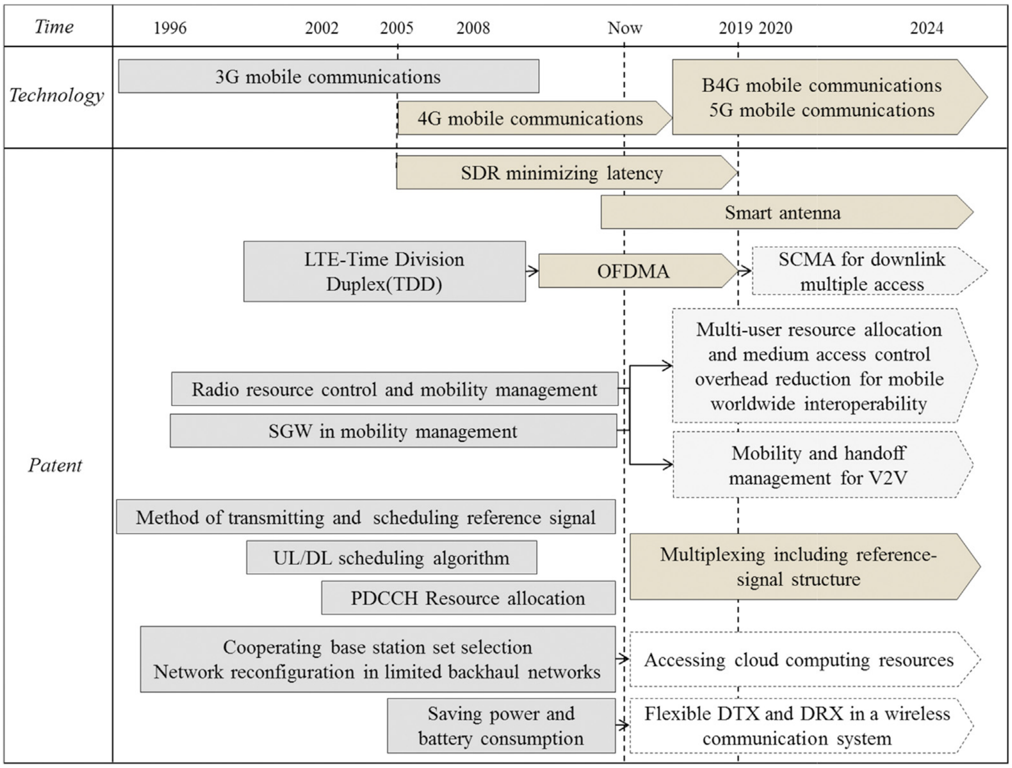

Fig. 10. A definitive patent roadmap.

judged that SDR technology started to apply patents from 2005, and the application for meaningful patents would be terminated in 2019, thus a node "SDR minimizing latency" can be mapped for that duration in the patent layer. It can be proved that there were many attempts to implement SDR technology in a wide range, and that technology that will be developed will link to the next generation of telecommunication to improve the efficiency and effectiveness of transmission and communication.

Three nodes relevant to OFDM/OFDMA technology—"LTE-Time Division Duplex (TDD)," "OFDMA" and "SCMA for downlink for multiple access"—were extracted and arranged by order of technological development. An "LTE-Time Division Duplex (TDD)" node first appeared with trials to decide the standards for 4G telecommunication, especially Long-term Evolution (LTE) technology. Then, the "OFDMA" node occurred and will sustain until 2019 according to the time gap and duration from the Bass model. A node "Sparse Code Multiple Access (SCMA) for downlink multiple access" is arranged as a successor of the above nodes. It is a new frequency domain non-orthogonal multiple-access technique that will improve the spectral efficiency of wireless radio. This patent will have the limelight as a core technology for the implementation of 5G mobile communications to increase the 5G wireless cellular network. It will provide an advantage for MU-SCME over other existing multiplexing techniques such as MU-MIMO, in which sensitivity to channel aging and the high overhead of channel knowledge feedback are obstacles for their practical implementation in the real world (Nikopour et al., 2014).

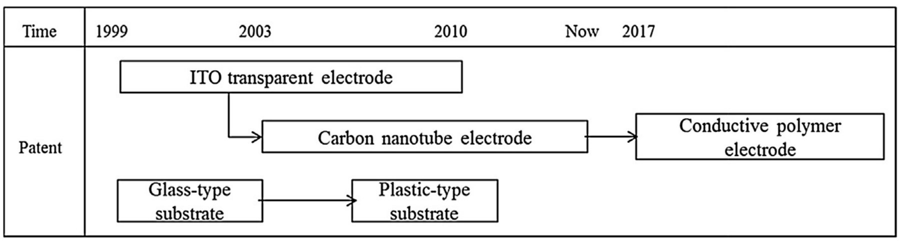

Fig. 11. The partly established patent roadmap in the display industry.

In order to evaluate the roadmapping approach, this study additionally established the patent roadmap according to the proposed approach in the field of display which has quite clear subsequence of technology development such as evolution from LCD (Liquid Crystal Display) to OLED (Organic Light Emitting Diodes) as shown in Fig. 11. As a result, there was a little difference in terms of appearance time and duration due to different attributes and types of technology, but the sequences of development derived from the suggested approach was similar to real cases of technology development in the display industry. For example, electrodes composed of carbon nanotube or conductive polymers have been in the limelight because the ITO transparent electrode for developing OLED has defectives which is not flexible as well as has possibility for exhaustion of resources. Even though this approach is evaluated by applying to other technology area, there is another alternative and it will be dealt with in the next section.

5. Conclusion and future research

This study proposed a new roadmapping process for patents through the GTM and Bass model. The GTM technique was utilized for identifying potential patent vacuums and foreseeing the content of promising research areas. It was effective for investigating the status of patent applications and exploring vacuum areas in the development of technology. Vacuums extracted in the GTM-based patent map are regarded as potential areas requiring the development of emerging technologies. The Bass model was applied to forecast the maximum amount of patent applications in related technological areas and the approximated appearance time of new patent nodes. The approach was helpful for strategically estimating the appropriate time for patent applications by reflecting the trends of technological development, overcoming disadvantages such as difficulties in accurately discovering the timing of the appearance of new patents.

This research contributes to the fields of patent planning and forecasting, and knowledge discovery through data mining in two ways. First, this study has attempted to define an advanced concept of patent roadmaps for the purpose of patent forecasting and planning from a conceptual perspective. Second, the GTM and Bass model, which are regarded as quantitative approaches, provided meaningful information such as vacuums in GTM-based patent maps and the appearance time of novel patenting areas estimated by the Bass model. The proposed approach has an advantage to identify incremental innovation based on historical patent data, which is mapped by the GTM technique, exploring undeveloped areas. In other words, this approach is superior to investigate incremental innovation because it relies on analysis of historical data, and patterns from patent data are applied to extract contents of promising patents.

The proposed roadmapping process will be critical for exploring directions for R&D activities, because it can analyze the status of patent development and vacuum areas, subsequently finding strategic fields for researchers and strategists. Thus, it can play an important role in establishing strategies for patent applications and R&D activities in research institutes and firms' R&D departments. The patent roadmap can serve as a tool for arranging a new budget from a viewpoint of government because it becomes a critical evidence to prioritize R&D programs with reducing time and efforts to select promising areas, overcoming the disadvantages of quite subjective participation of domain experts through the systematic roadmapping process. At the nation level, it enables to being more competitive in the global market or to create a new market through well-planned patents and technologies resulting from active R&D efforts for predicted areas in advance by the established patent roadmap. For example, government may focus on the promising fields derived from patent roadmapping and encourage for firms and research institutions to conduct R&D activities through enormous R&D investment. Based upon high R&D intensity, firms and institutions not only develop technology but also commercialize them by embedding them to products and/or services, leading to create a new market or extend existing market. Thus, it can realize the successful commercialization of patents by enabling to focus on promising areas that correspond to well-defined patents, making huge profits and adding value.

Although we have proposed a roadmapping process based on a quantitative methodology, future research should focus on implementing patent roadmaps by applying the proposed process. At first, validating vacuums is necessary after evaluating weighted sum by reflecting technological characteristics and the types of technology. In this study, several vacuums were removed by the cut-off value, but these may be promising areas. Second, it might be difficult to find entirely new patenting areas, because the contents of nodes in the roadmap have been derived based on patents that have already been applied. In other words, exploring radical innovation will be harder than incremental innovation in the proposed approach. Thus, further efforts should be made to overcome these limitations of this study by complementing them with additional processes. Furthermore, the proposed approach will be evaluated by comparison to a patent roadmap developed in traditional ways such as workshop, Delphi technique through using a classification matrix which is widely used as the accuracy measure. It may lead to more objective validation for the patent roadmap derived from the suggested approach.

Acknowledgment

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2014R1A1A2054892).

References

Andrade, A.O., Nasuto, S., Kyberd, P., Sweeney-Reed, C.M., 2005. Generative topographic mapping applied to clustering and visualiation of motor unit action potentials. BioSystems 82 (3), 273–284.
Bardhan, A.K., Chanda, U., 2007. A model for adoption of successive generations of a high technology product. Int. J. Technol. Mark. 2, 53–64.
Bass, F.M., 1969. A new product growth for model consumer durables. Manag. Sci. 15 (5), 215–227.
Bigwood, M.P., 1997. Patent trend analysis: incorporate current year data. World Patent Inf. 19 (4), 243–249.
Bishop, C.M., Hinton, G.E., Strachan, I.G.D., 1997. GTM through time. In: Proceedings of Conference GTM through time, Cambridge, UK.
Bishop, C.M., Svensén, M., 1998. GTM: the generative topographic mapping. Neural Comput. 10 (1), 215–234.
Bishop, C.M., Svensén, M., Williams, C.K.I., 1998. Developments of the generative topographic mapping. Neurocomputing 21, 203–224.
Bretschneider, S.I., Mahajan, V., 1980. Adaptive technological substitution models. Technol. Forecast. Soc. Changes 18, 129–139.
Brockhoff, K., 1991. Competitor technology intelligence in German companies. Ind. Mark. Manag. 20 (2), 91–98.
Bullen, R.J., Cornford, D., Nabney, I.T., 2003. Outlier detection in scatterometer data: neural network approaches. Neural Netw..
Carreira-Perpiñán, M.A., 2000. Reconstruction of sequential data with probabilistic models and continuity constraints. Adv. Neural Inf. Process. Syst. 12, 1–7.
Charlton, A., 2012. Patent Wars 2012: Six Key Battles in the Technology Industry. International Business Times, United Kingdom.
Chen, Y.-H., Chen, C.-Y., Lee, S.-S., 2010. Technology forecasting of new clean energy: the example of hydrogen energy and fuel cell. Afr. J. Bus. Manag. 4 (7), 1372–1380.
Cheng, A.-C., 2012. Exploring the relationship between technology diffusion and new material diffusion: the example of advanced ceramic powders. Technovation 32 (3–44), 163–167.
Chengjian, G., Lucheng, H., 2009. The study on CNT-FED for emerging technology forecasting by using patent management map. In: Proceedings of Conference The study on CNT-FED for Emerging Technology Forecasting by using Patent Management Map, Ternopil, Ukraine.
Chia, T.H., 2012. Fighting the smartphone patent war with RAND-Encumbert Patents. Technol. Law J. 27, 210–240.
Choi, C., Park, Y., 2009. Monitoring the organic structure of technology based on the patent development paths. Technol. Forecast. Soc. Changes 76, 754–768.
Curran, C.-S., Leker, J., 2011. Patent indicators for monitoring convergence – examples from NFF and ICT. Technol. Forecast. Soc. Changes 79 (2), 256–273.
Ernst, H., 1997. The use of patent data for technological forecasting: the diffusion of CNC-Technology in the machine tool industry. Small Bus. Econ. 9, 361–381.
Fisher, J.C., Pry, R.H., 1971. A Simple substitution model of technological change. Technol. Forecast. Soc. Changes 3, 75–88.
Fyfe, C., 2007. Two topographic maps for data visualization. Data Min. Knowl. Disc 14, 207–224.
Girolami, M., 2002. Latenet variable models for topographic organisation of discrete and strictly positive data. Neurocomputing 48, 185–198.
Goetze, C., 2010. An empirical enquiry into co-patent networks and their stars: the case of cardiac pacemaker technology. Technovation 30 (7–8), 436–446.
Goldenberg, J., Libai, B., Muller, E., 2001. Using complex system analysis to advance marketing theory development: modeling heterogenetiy effects on new product growth through stochastic cellular automata. Acad. Mark. Sci. Rev. 2001 (9) .
Grambmeier, J., Rudolph, A., 2002. Techniques of cluster algorithms in data mining. Data Min. Knowl. Disc 6, 303–360.
Hidalgo, A., Gabaly, S., 2012. Use of prediction methods for patent and trademark applications in Spain. World Patent Inf. 34 (1), 19–29.
Jeong, S., Yoon, B., 2013. A systematic approach to exploring an essential patent linking standard and patent maps: application of generative topographic mapping (GTM). Eng. Manag. J. 25 (1), 48–57.
Jeong, Y., Yoon, B., 2014. Development of patent roadmap based on technology roadmap by analyzing patterns of patent development. Technovation, http://dx.doi.org/10.2016/k.technovation.2014.03.001.
Kash, D.E., Rycoft, R.W., 2000. Patterns of innovating complex technologies: a framework for adaptive network strategies. Res. Policy 29, 819–831.
Keefe, S.L., 2013. Patenting in the United States. Civil Engineering & Patent Law, VA, United States.
Kim, D.H., Shin, Y.G., Park, S.S., Jang, D.S., 2009. Forecasting diffusion of technology by using Bass model. AIP Conf. Proc. 1148 (1), 149–152.
Kostoff, R.N., 1997. The Handbook of Research Impact Assessment DTIC Report ADA 296021, 7th ed. Office of Naval Research, VA, United States.
Kucharavy, D., Guio, R.D., 2011. Logistic substitution model and technological forecasting. Procedia Eng. 9, 402–416.
Lee, S., Yoon, B., Park, Y., 2009. An approach to discovering new technology opportunities: keyword-based patent map approach. Technovation 29, 481–497.
Leigh, L., Yorke-Smith, N., 2011. An adaptation of the bass new product diffusion model for multiple purchases of capital items. In: Proceedings of Conference an Adaptation of the Bass New Product Diffusion Model for Multiple Purchases of Capital Items, Paris, France.
Lim, J., Nam, C., Kim, S., Lee, E., Lee, H., 2012. Forecasting 3G mobile subscription in China: a study based on stochastic frontier analysis and a Bass diffusion model. Telecommun. Policy 36 (19), 858–871.
Linton, J.D., 2002. Forecasting the market diffusion of disruptive and discontinuous innovation. IEEE Trans. Eng. Manag. 49 (4), 365–374.
Lloyd, M., Spielthernner, D., Mokdsi, G., 2011. The Smartphone Patent Wars. Griffith Hack, Australia.
Mahajan, V., 1979. First-purchase diffusion models of new-product acceptance. Technol. Forecast. Soc. Changes 15 (2), 127–146.
Mahajan, V., Muller, E., Bass, F.M., 1990. New product diffusion models in marketing: a review and directions for research. Am. Mark. Assoc. 54 (1), 1–26.
Mansfield, E., 1961. Technical change and the rate of imitation. Econometrica 29 (4) .
Michalakelis, C., Varoutas, D., Sphicopoulos, T., 2010. Innovation diffusion with generation substitution effects. Technol. Forecast. Soc. Changes 77 (4), 541–557.
Morris, S., DeYong, C., Wu, Z., Salman, S., Yemenu, D., 2002. DIVA: a visualization system for exploreing document databases for technology forecasting. Comput. Ind. Eng. 43 (4), 841–862.
Nikopour, H., Yi, E., Bayesteh, A., Au, K., Hawryluck, M., Baligh, H., Ma, J., 2014. SCMA for downlink multiple access of 5G wireless networks. Globecom 2014, 3940–3945.
Norton, J.A., Bass, F.M., 1987. A diffusion theory model of adoption and substitution for successive generations of high-technology products. Manag. Sci. 33 (9), 1069–1086.
O'Hearn, T.J., 2000. The reason for the patent wars. Bus. Horizon 43 (4), 33–42.
Olier, I., Vellido, A., 2008. Advances in clustering and visualization of time series using GTM through time. Neural Netw. 21 (7), 904–913.
Pampalk, E., 2001. Limitations of the SOM and the GTM. Department of Medical Cyberetics and Artificial Intelligence, University of Vienna.
Pearl, R., Reed, L.J., 1920. On the rate of growth of the population of the United States since 1790 and its mathematical representation. Proc. Natl. Acad. Sci. U. S. A. 6 (6), 275–288.
Plotkin, R., 2014. Patent Roadmaps. Robert Plotkin, P.C., MA, United States.
Ponomarev, I.V., Williams, D.E., Hackett, C.J., Schnell, J.D., Haak, L.L., 2014. Predicting highly cited papers: a method for early detection of candidate breakthroughs. Technol. Forecast. Soc. Changes 81 .
Rao, K.U., Kishore, V.V.N., 2009. Wind power technology diffusion analysis in selected states of India. Renew. Energy 34 (4), 983–988.
Roepke, S., Moehrle, M.G., 2014. Sequencing the evolution of technologies in a system-oriented way: the concept of technology-DNA. J. Eng. Technol. Manag. 32 (4/6), 110–128.

Son, C., Geum, Y., Park, Y., 2013. How to identify the trends of services: GTM-TT service map. Expert Syst. Appl. 40 (8), 2956–2965.
Son, C., Suh, Y., Jeon, J., Park, Y., 2012. Development of a GTM-based patent map for identifying patent vacuums. Expert Syst. Appl. 39 (3), 2489–2500.
Song, Y., Lee, S., Zo, H., Lee, H., 2015. A hybrid Bass-Markov model for the diffusion of a dual-type device-based telecommunication service: the case of WiBro service in Korea. Comput. Ind. Eng. 79, 85–94.
Svensén, M., 1999. In: Group NCR (Eds.), The GTM Toolbox – User's Guide. Aston University, Birmingham, UK.
Tellis, G.J., 2007. A critical review of marketing research on diffusion of new products. Rev. Mark. Res. 3, 39–80.
Tino, P., Nabney, I., 2002. Hierarchical GTM: constructing localized nonlinear projection maniforlds in a principled way. IEEE Trans. Pattern Anal. Mach. Intell. 24 (5), 639–656.
Vellido, A., 2006. Missing data impultation through GTM as a mixture of t-distributions. Neural Netw. 19, 1624–1635.
Vellido, A., El-Deredy, W., Lisboa, P.J.G., 2003. Selective smoothing of the generative topographic mapping. IEEE Trans. Neural Netw. 14 (4), 847–852.
Verspagen, B., 2007. Mapping technological trajectories as patent citation networks: a study on the history of fuel cell research. Adv. Complex Syst. 10 (1), 93–115.
Vesanto, J., 1999. SOM-based data visualization methods. Intell. Data Anal. 3, 111–126.
Wartburg, I.V., Teichert, T., Rost, K., 2005. Inventive progress measured by multi-stage patent citation analysis. Res. Policy 34, 1591–1607.
Winsor, C.P., 1932. The gompertz curve as a growth curve. Proc. Natl. Acad. Sci. U. S. A. 18 (1), 1–8.
Yang, J., Zhang, B., 2001. Customer data mining and visualization by generative topographic mapping methods. In: Proceedings of Conference Customer data mining and visualization by generative topographic mapping methods, Freiburg, Germany.