Technological Forecasting & Social Change 91 (2015) 264–279

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Development of data-driven technology roadmap considering dependency: An ARM-based technology roadmapping

Youngjung Geum [a], HyeonJeong Lee [b], Youngjo Lee [c], Yongtae Park [d,*]

[a] Department of Industrial and Systems Engineering, Seoul National University of Science and Technology, Seoul, Republic of Korea
[b] Vic21Task, LG Electronics, Seoul, Republic of Korea
[c] Department of Statistics, Seoul National University, Seoul, Republic of Korea
[d] Department of Industrial Engineering, School of Engineering, Seoul National University, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 30 October 2013
Received in revised form 2 March 2014
Accepted 6 March 2014
Available online 13 April 2014

Keywords:
Technology roadmap
TRM
Dependency
Association rule mining
ARM

ABSTRACT

The active incorporation of business data has become a vital process in the recent business environment. Despite the potential utility of massive database, technology roadmap, a well-known strategic planning method, still remains a subjective and qualitative method conducted by some experts. Even if some studies have tried, previous research lacks a dependency measure that can be used between layers, which is a critical part of technology roadmaps. This paper therefore suggests an association rule mining (ARM)-based technology roadmap to identify the relationship between different layers. The use of ARM fits the purpose, in terms of capturing the dependency information. Two types of roadmap are developed: a keyword portfolio map and a keyword relational map. In the keyword portfolio map, four types of keyword pairs are identified according to their support and confidence. In the keyword relational map, a 2-dimensional map is developed using support as an intra-layer affinity relationship and confidence as an inter-layer dependency relationship.

1. Introduction

We live with data all around the world. With the help of information technology (IT), many kinds of business transactions occur in electric form, which causes a massive database to be generated and consumed [17,19]. With the rise of the Internet, streams of logs and user data generated from many information systems give insight not only into the operation of a system itself, but also into the behavior patterns of users [5]. In online marketplaces, customer needs and requirements are represented in the form of customer reviews or forums [3]. In addition, the products and services of a firm can be identified in many types of documents, such as service descriptions in a mobile open market, or product/service manuals.

Quite naturally, the active incorporation of these massive databases becomes an imperative and vital process for recent business environment [4,13,20]. Data can be utilized in two ways. The first usage is related to the reactive process such as a bibliometric analysis, analyzing historical patterns based on the statistical approach. The second usage deals with more prominent issue, the proactive process. The proactive process is related to the forecasting and planning process in which firms forecast future trends and plan their strategies. This means that current planning methods should actively incorporate business data into their planning procedures. The technology roadmap, which is a representative and prominent tool for the strategic planning [14,22], is no exception.

However, the development of technology roadmaps still remains a subjective and qualitative task conducted by only some experts [15]. Expert knowledge, of course, still plays a decisive role, and may be more desirable due to the strategic

* Corresponding author at: Department of Industrial Engineering, School of Engineering, Seoul National University, San 56-1, Shilim-Dong, Kwanak-Gu, Seoul 151-742, Republic of Korea. Tel.: +82 2 880 8358; fax: +82 2 889 8560.
E-mail addresses: yjgeum@seoultech.ac.kr (Y. Geum), missydoris@snu.ac.kr (H. Lee), youngjo@snu.ac.kr (Y. Lee), parkyt1@snu.ac.kr (Y. Park).

http://dx.doi.org/10.1016/j.techfore.2014.03.003
0040-1625/

nature of technology roadmaps [11]. However, the active incorporation of business data cannot simply be neglected since it can deliver essential, unexpected, and effective information [10,15].

Responding to the needs of data-driven roadmapping, there has been some research into the analytic approach to technology roadmapping, measuring the relationship between layers in the technology roadmap [15,16,24,26,27]. The discussion in these studies has mainly focused on the quantification of relationships between planning elements, which is a critical step in technology roadmapping.

However, previous studies on the quantification of relationships in the technology roadmap surprisingly converged on a common limitation: the lack of dependency measurement. Previous studies have focused on measuring the relationship using keyword similarity, by simply calculating the frequency of keyword occurrence [15,16,24,26,27]. However, the tenet of technology roadmaps stems from identifying links among the market, service, product, and technologies of a specific firm — in other words, links between different layers. This means, measuring "causal" relationships between layers, i.e., measuring "dependency" is the core and critical information to be incorporated.

The use of association rule mining (ARM) fits the purpose. The essence of ARM is the identification of relationships and potential associations from huge amounts of data [1,9,23]. The rules used can be effective in uncovering unknown relationships which can be an inspiration for further decision making [9]. ARM is particularly capable of capturing dependency information using the confidence measure, which can be effectively implemented in the technology roadmap.

In response, this paper suggests an association rule mining (ARM)-based technology roadmap to identify the relationship between different layers, facilitating the development process of a data-driven technology roadmap. Using ARM, two types of associations are measured and employed: support and confidence. Support measures the ratio of the number of transactions that include two specific items, which can be expressed as item affinity. Confidence measures the ratio of the number of transactions containing a specific item, given in transactions containing another item, which can represent the dependency relationship. The advantage of expressing the dependency relationship in ARM is that it provides an excellent methodological sufficiency for the technology roadmap, whose core value stems from the dependency relationship between each layer.

The remainder of this study is organized as follows. A literature review deals with the theoretical and methodological background of this paper: the era of big data, technology roadmap, and ARM. The proposed approach considers how the data-driven technology roadmap can be developed and how ARM is used for this purpose. The overall process and detailed procedures are introduced. To illustrate how the proposed approach works, a simple case study is conducted. Finally, the contributions and limitations of this paper are provided in the conclusion.

2. Related works

2.1. Technology roadmap

The technology roadmap has long been considered a prominent tool for the strategic planning of technology [14,22]. It enables a firm to carry out its R&D activities in a systematic manner, laying out explicit plans about what technologies to develop, when and how. Fig. 1 illustrates the generic structure of a technology roadmap as a time-based chart, comprising a number of layers [22].

Some researchers view a technology roadmap as a visualization tool for the strategic management of technology. Although there are various definitions of the term 'roadmap', the key features and benefits typically relate to visualization and communication [26]. Many practitioners and researchers visualize or summarize such information to achieve a variety of benefits. Yoon et al. [27] present four techniques to structure technological information for technology roadmapping: summarization, information extraction, clustering and navigation.

There have been several attempts to find the most effective way to build technology roadmaps. Bray and Garcia [2] suggested three phases: preliminary activity, roadmap development, and follow-up activity. Groenveld [6] developed a seven-stage process. The development of 'T-Plan' supports the swift initiation of roadmapping in three stages: planning, roadmapping and roll-out [21], and a modified T-Plan process has also been introduced with five key modules [8]. Lee and Park [14] suggested a framework for customizing the technology roadmapping process according to its specific purposes, suggesting eight formats of roadmaps.

These roadmapping processes commonly include an important step: identifying the relationship between layers. Since the technology roadmap is a multiple-layered chart including market, product and technology layers, identification of relationships between layers is of importance for identifying the "when and how" strategy. Kostoff and Schaller [10] noted the need for measuring functional relationships in technology roadmaps. Due to the inherent uncertainties and evolving requirement changes in large programs, the structure of technology roadmap should be flexible enough to incorporate the dynamics. This denotes the importance of linked functional relationships that reflect changes at any node of technology roadmap to the whole layers of the technology roadmap [10]. The task of identifying relationships between layers, however, is mostly dependent on expert judgment [10,15]. Due to the rise of big data, a proactive process of incorporating the data is becoming increasingly important.

2.2. ARM

ARM is one of the most important and well researched techniques of data mining. It identifies associations among a

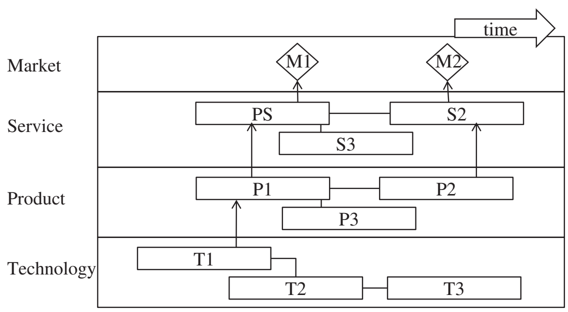

Fig. 1. Structure of technology roadmap.

set of product items frequently purchased together, and is used as a widespread approach to market basket analysis [1,18]. It aims to extract interesting correlations, frequent patterns, associations or casual structures among sets of items in the transaction databases or other data repositories [1,12]. The main advantages of ARM are the ease of application and interpretation. The results of the method can be easily understood because they are expressed through if–then rules.

ARM generates association rules that represent interesting relationships among items in a given data set [7]. A rule is generally expressed as {X → Y}, which means that the transaction including item X also includes item Y. In ARM, there are three major types of association rules: support, confidence, and lift, as represented below.

$$\begin{aligned}&\mathrm{Support}(X{\rightarrow}Y) = P(X{\cap}Y) = \frac{N(X{\cap}Y)}{N}\\&\mathrm{Confidence}(X{\rightarrow}Y) = \frac{P(X{\cap}Y)}{P(X)} = P(Y|X) = \frac{N(X{\cap}Y)}{N(X)}\\&\mathrm{Lift}(X{\rightarrow}Y) = \frac{P(Y|X)}{P(Y)} = \frac{N(X{\cap}Y)N}{N(X)N(Y)}\end{aligned} \tag{1}$$

Firstly, support measures the ratio of the number of transactions that include both items X and Y. This implies the usefulness of discovered rules through the probability of co-occurrence of items X and Y in a given set. Secondly, confidence measures the ratio of the number of transactions containing item Y, given in transactions containing item X. Therefore, confidence can be represented as the conditional probability of Y given X. Statistically, this denotes the certainty of the rule. Finally, lift measures the statistical dependence between items X and Y, calculated by dividing the confidence by the probability P(Y). If the lift value is greater than one, it shows a positive correlation.

To generate association rules is to discover all the association rules that have support greater than, or equal to, a minimum support (minsup) threshold and confidence greater than, or equal to, a minimum confidence (minconf) threshold. The association rules must satisfy two conditions:

$$\begin{aligned}&\mathrm{Support}(X{\rightarrow}Y) \geq \min\ \sup\\&\mathrm{Confidence}(X{\rightarrow}Y) \geq \min\ \mathrm{conf}.\end{aligned} \tag{2}$$

3. ARM-based technology roadmap

3.1. Assumptions

The tenet of this paper starts from the fact that business documents can be effectively utilized for identifying the current trends and planning for the future. Therefore, each business document is represented as a set of keywords. Each layer consists of several keywords that represent the characteristics of products, services, and technologies. The relationships between layers are identified using ARM: the support and confidence of a keyword set. This information is further represented as a form of network graph to show the quick but clear view.

To develop the data-driven technology roadmap, some assumptions should be made. This paper employs two types of assumptions: document-related assumptions and keyword-related assumptions, as shown in Table 1.

Table 1
Assumptions for data-driven technology roadmaps.

Type	Assumption
Document-related assumptions	· A specific document employed contains both product-related and service-related keywords · The keyword represents the contents of documents well · Frequency and existence of keyword attributes represent the importance of each keyword
Keyword-related assumptions	· If a keyword P1 and a keyword P2 frequently occurs simultaneously, keyword P1 and keyword P2 have an affinity relationship · If a keyword P1 happens frequently when keyword P2 happens, keyword P1 and keyword P2 have a dependency relationship

3.2. Overall process

This paper suggests a data-driven technology roadmap which provides a quantitative assessment of relationships by employing ARM. ARM and technology roadmapping have complementary characteristics in providing a data-driven technology roadmap in terms of their characteristics, purpose, operational procedures, and analytic objects, as shown in Table 2. For this reason, the challenge of marrying the technology roadmap and ARM appears worth the effort.

The overall process of this paper is composed of five steps, as shown in Fig. 2. Firstly, data utilized for the technology roadmap is identified. For each layer, appropriate data sources are identified to provide the data-driven technology roadmap. The second step is to identify the keyword list for each layer. Based on the text-mining analysis for each document, keywords which represent the document are identified, and corresponding keyword vectors are identified in turn. Following on from the keyword-vector, the ARM technique is employed for measuring keyword relationships and dependency. Support measures the co-occurrence of two keywords whereas confidence measures the dependency of two keywords. After measuring each rule, two types of map are developed: a keyword portfolio map and a keyword relational map. In the keyword portfolio map, four types of keyword pairs are identified according to their support and confidence. In the keyword relational map, a 2-dimensional map is developed using support as an intra-layer affinity relationship and confidence as an inter-layer dependency relationship.

3.3. Detailed procedure

3.3.1. Data collection

First and foremost, the development of a data-driven technology roadmap starts with data selection. Each layer of the

Table 2
Characteristics of technology roadmap and ARM.

Method	Technology roadmap	ARM
Characteristics	Qualitative & visionary	Quantitative & analytic
Purpose	Planning	Operations
Operated by	Mostly conducted by expert judgments	Mostly conducted by automatic algorithms
Object	Relationship between elements	Relationship between elements

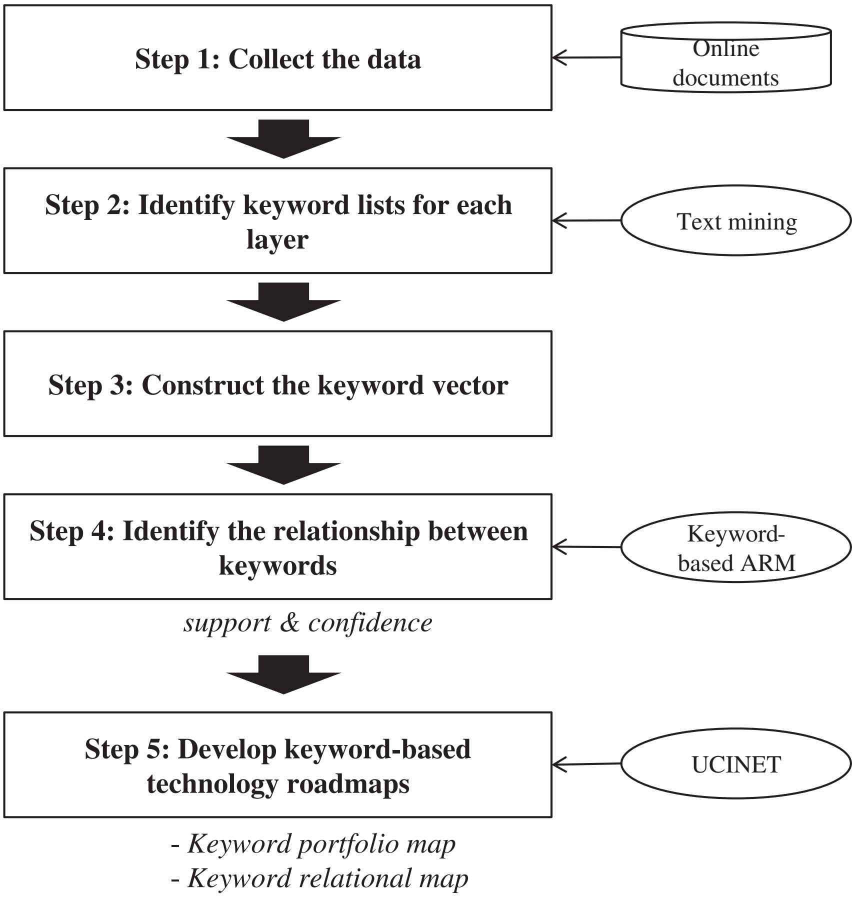

Fig. 2. Overall process of this paper.

technology roadmap – market, service, product, and technology – should be represented by a set of keywords. Therefore, selecting appropriate data sources to represent the innate characteristics of each layer is of great importance.

Many types of data are eligible for developing roadmaps. For the market layer, firms want to know recent trends in the market environment, current customer needs, and potential customers. To get this information, customer forums, where customers actively discuss current services or the current marketplace can be a good starting point. In addition, trend magazines or news sites can also be excellent data sources. For the product/service layer, firms want to know current product/service functions and promising functions that customers want. To achieve this, product/service description documents or customer reviews can be a good data source. Similarly, for the technology layer, patents and academic papers could provide essential information for technological trends and future issues. Table 3 shows the types of layer-specific data sources for technology roadmaps.

3.3.2. Identify keyword lists for each layer

After selecting data sources, the next step is to select appropriate keywords for each layer. This is based on the assumption that a document can be expressed by a set of relevant keywords. A number of studies on knowledge discovery in texts have been based on this assumption [25].

This process is conducted via two steps. The first step is the quantitative collection of keywords by employing a text-mining technique. In this step, keywords are extracted from documents containing customer needs and product/service features. The text-mining package, TextAnalyst 2.1, is used to extract the keywords. However, keywords extracted with the text mining technique are insufficient to describe the market or service characteristics. For this purpose, as the second step, a qualitative filtering process – screening and repetition process – is conducted. In this step, expert judgment plays a key role in defining the keywords. Experts again select appropriate keywords and filter the keywords based on their judgment. Overly general or insignificant keywords are excluded. The final set of keywords is derived, having accounted for abbreviations, synonyms, and singular and plural forms of words. Note that keywords are domain-specific as well as firm-specific.

3.3.3. Construct the keyword vector

The next step is to construct the keyword vector for each document. Consider a keyword vector for measuring relationships between the product layer and the service layer. Suppose that there are d documents. In d documents, we

Table 3
Types of data for technology roadmap.

Layer	Required information	Required data	Available source (ex. mobile service)
Market	- What customers want to receive - Who is the potential customer - How the social trends changes	- Customer opinion - Customer review - Customer idea for new service - News and reports - Trend magazine	http://itunes.apple.com/kr/app/twitter/id333903271?mt=8 http://discussions.apple.com/ http://appcomments.com http://modmyi.com/forums http://iphoneapplicationlist.com/forum http://www.iphonewoners.com http://appstorehq.com
Service	- What function the current service provides - What utility the future customer wants	- Service description - Service manual (for installation or use)	http://appshopper.com http://iphoneappsplus.com http://appstorehq.com https://market.android.com
Product	- What function the current product provides - What utility the future customer wants	- Product description - Product manual (for installation or use)	http://appshopper.com http://reviews.cnet.com http://www.techradar.com/reviews/phones/mobile-phones/ http://thesmartphoneforum.com/ http://forum.brighthand.com/
Technology	- What aspects the technology provides - What technological aspects the future products wants	- Technology description - Patents - Academic papers	http://uspto.gov http://thesmartphoneforum.com/ http://forum.brighthand.com/

identify k product keywords and n service keywords. In this situation, the keyword vector is constructed, as shown in Eq. (1). Here, $p_{dk} = 1$ means that a document d contains a product keyword k, and $p_{dk} = 0$ means that a document does not. Similarly, $s_{dk+n} = 1$ denotes that a document d contains a service keyword n, and $s_{dk+n} = 0$ means that a document does not.

$$\begin{bmatrix} document\ 1 \\ document\ 2 \\ \cdots \\ document\ d \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} & s_{1k+1} & s_{1k+2} & \cdots & s_{1k+n} \\ p_{21} & p_{22} & \cdots & p_{2k} & s_{2k+1} & s_{22} & \cdots & s_{2k+n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{d1} & p_{d2} & \cdots & p_{dk} & s_{dk+1} & s_{dk+2} & \cdots & s_{dk+n} \end{bmatrix} \tag{1}$$

where

d number of documents
k number of product keywords
n number of service keywords
$pij = 1$ a document i contains a product keyword j
$sij = 1$ a document i contains a service keyword j

3.3.4. Identify the relationships of keywords

The next step is to identify the relationships between different layers. In this paper, two types of association rules are identified from ARM: support and confidence. Support measures the affinity between two keywords, which means "how closely the two keywords are interrelated." When developing a technology roadmap, this can be applied to the horizontal axis, which shows the recent trends of specific products or services. Measuring affinity can also provide patterns of keyword generation, if applied with the dynamic analysis. This characteristic fits for the development of each layer, since each layer provides the relationships of each product or service from a static perspective. Each layer can also represent the evolutionary change of a certain product line or service line.

In contrast, confidence measures the dependency between two elements, which means "how frequently a specific keyword occurs when another keyword occurs." This can be considered as the conditional probability of Y given X, which is closely related to the dependency of two keywords. This can be effectively applied for measuring the relationship between different layers (e.g. relationship between product layer and product layer), since it mainly reflects dependency between the two. For example, the development of a product element is mainly dependent on the existence of a technology element, which implies the need to measure the dependency between the two. Confidence is therefore effectively applied for the development of the vertical axis of a technology roadmap by identifying the causal relationship between layers. This can be represented as Eq. (2).

$$\begin{aligned} \text{Support}(p \rightarrow s) &= P(p \cap s) = \frac{N_{p \cap s}}{N_{total}} \\ \text{Confidence}(p \rightarrow s) &= \frac{P(p \cap s)}{P(p)} = P(p|s) = \frac{N_{p \cap s}}{N_s} \end{aligned} \tag{2}$$

where

N_p number of documents that contain product keyword p
N_s number of documents that contain service keyword s
$N_{p \cap s}$ number of documents that contain both product keyword p and service keyword s
N_{total} number of total document

The keyword vector represented in Eq. (1) is now converted into the revised format to measure support for the intra-layer relationship and confidence for the inter-layer relationship, as shown in Eq. (3). Here, $s_{k,n}$ means the support value of keyword k and keyword n, whereas $c_{k,n}$ denotes the confidence value of keyword n given keyword k.

$$\begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,k} & c_{1,k+1} & c_{1,k+2} & \cdots & c_{1,k+n} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,k} & c_{2,k+1} & c_{2,k+2} & \cdots & c_{2,k+n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ s_{k,1} & s_{k,2} & \cdots & s_{k,k} & c_{k,k+1} & c_{k,k+2} & \cdots & c_{k,k+n} \\ c_{k+1,1} & c_{k+1,2} & \cdots & c_{k+1,k} & s_{k+1,k+1} & s_{k+1,k+2} & \cdots & s_{k+1,k+n} \\ c_{k+2,1} & c_{k+2,2} & \cdots & c_{k+2,k} & s_{k+2,k+1} & s_{k+2,k+2} & \cdots & s_{k+2,k+n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{k+n,1} & c_{k+n,2} & \cdots & c_{k+n,k} & s_{k+n,k+1} & s_{k+n,k+2} & \cdots & s_{k+n,k+n} \end{bmatrix} \quad (3)$$

where

k	number of product keywords
n	number of service keywords
si,j	support between keyword i and keyword j
ci,j	confidence between keyword i and keyword j

3.3.5. Develop the technology roadmap

Based on the results of ARM, a technology roadmap is developed. In this study, two types of map are developed: a keyword portfolio map and a keyword relational map. In the keyword portfolio map, four types of keyword pairs are identified according to their support and confidence. In the keyword relational map, a 2-dimensional map is developed using support as an intra-layer affinity relationship and confidence as an inter-layer dependency relationship.

3.3.5.1. Keyword portfolio map. Based on the association result, this paper suggests a keyword portfolio map to represent the level of confidence and support within each keyword pair, as shown in Fig. 3. The matrix is divided into four quadrants: interactive keywords, causal keyword, family keywords, and unrelated keywords.

1) *Interactive*
'Interactive' quadrant is characterized by high confidence and high support for a certain keyword pair, implying that those two keywords have a high level of co-occurrence and a high level of dependency. This means that those keywords are closely associated in terms of their characteristics and functions. In other words, those keywords are additionally highly dependent in terms of their occurrence. Keyword pairs in this quadrant therefore require co-development when firms plan new products or services.
2) *Causal*
The second type is a causal keyword set. Keyword pairs in this quadrant are not so tightly coupled in terms of their occurrence. However, there exists dependency between two keywords, i.e., when a specific keyword A exists, keyword B is frequently likely to exist. In this case, keyword A plays an important role in providing B, or supports an additional utility for providing B. This means that the causal relationship between two keywords should be considered when developing new products or services.
3) *Family*
The third quadrant is named the family keyword set, which is characterized by high support and low confidence for a certain keyword pair, implying that those two keywords have a high level of co-occurrence and low level of

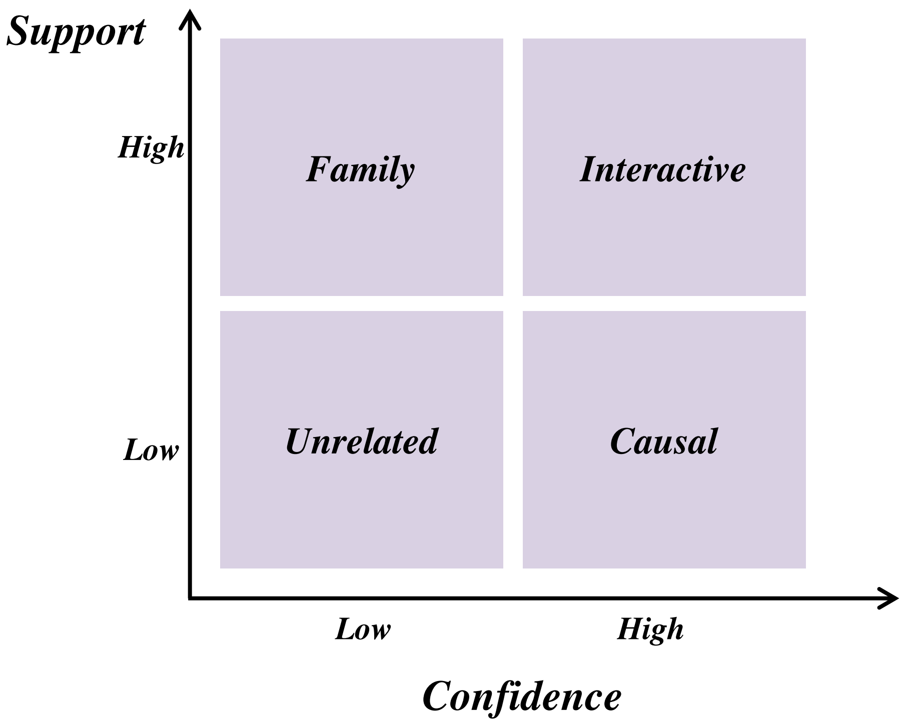

Fig. 3. Keyword portfolio map.

dependency. As easily identified from its name, keyword pairs in this quadrant frequently occur together, despite their low causal relationship. Therefore, those keywords share similar characteristics, provide similar functions, or provide similar customer utility. Therefore, when developing new services, these keywords should be considered as family, considering their co-occurrence.

4) *Unrelated*

The final quadrant is called the unrelated keyword set. These keywords are unrelated in terms of their co-occurrence and dependency. There is thus no need to consider the relationship between these keywords.

3.3.5.2. Keyword relational map. Based on the measurements conducted in the previous stage, the support and confidence of a specific keyword set are represented as a form of graph. Since support represents the affinity of two keywords, i.e. keyword family, this information is represented in the horizontal axis, i.e. within the layer. In contrast, since confidence measures the dependency or causality of two keywords, this is represented in the vertical axis to show the dependency between two layers, i.e. across the layer. The dependency represents how a product (or a service) occurs if a certain service (or product) exists. Types of relationships and corresponding measures are given in Table 4.

Based on this relationship, a keyword relational map is constructed. In this map, nodes represent each keyword and arcs represent the strength of relationship. Fig. 4 shows the generic structure of a keyword relational map.

Based on this graph, a technology roadmap is developed. The keyword relational map is a keyword-level roadmap, i.e. a micro-level roadmap. By aggregating keyword-level information into a product-level or service-level, more practical information can be derived. For example, the relationship between products and services and the relationship between products and technologies are identified. In addition, the keyword-relational map can be extended to the dynamic level roadmap by linking each snapshot for a certain timeline. This dynamic level keyword relational map could represent the dynamic changes of markets, products, services, and technologies.

4. Case study

To illustrate the working of the proposed approach, a simple case study was conducted. In this case study, we employed the Appstore database which is an ample source of current mobile application services. Although many types of data relate to products and services such as reports from service firms, academic papers, books, newspapers, and survey data, we selected the mobile application data from the Web (especially, Apple Appstore) for two reasons. First, it encompasses plentiful and diverse information to represent the current business environment. Since today's business environment is significantly attached to the mobile business, use of mobile application service as a target is a good choice. Secondly, web data can be easily accessed and downloaded using a computer-based technique, which enables convenient methods of analysis.

Table 4
Types of relationships and corresponding measures.

Index	Relations	Meaning	Axis of roadmap
Support	Co-occurrence	Affinity	Horizontal
Confidence	Dependency	Causality	Vertical

4.1. Data and pre-processing

As an initial step, we collected the documents that represented the product information and service information to develop the technology roadmap. This study downloaded the application services in Appstore, from the utility category. For effective analysis, the amount of data should not be too much or too little, and should be adjusted at a certain level. Therefore, we randomly selected 300 service documents from utility categories to provide a moderate level of data in order to achieve an effective result. To collect the data, a computer-based system using JAVA was used to automatically collect data from the web. Since this case was conducted for illustrative purposes, we developed a product layer and a service layer of technology roadmap.

4.2. Identification of keyword list

The second step is to identify the keyword list for each layer. Since two layers – a product layer and a service layer – are developed for the illustrative purpose, keywords representing product features and service features are identified. A number of studies into knowledge discovery in texts have been based on the assumption that a set of keywords in a document represents the topics of the given document [25].

The keyword selection process is conducted via a two-step process — a quantitative process which deals with automatic collection through a text-mining technique, 'TextAnalyst 2.1.' From 300 documents downloaded from the utility category, we collected frequent keywords. Frequent keywords seem to represent the core functionality of each service, showing what kinds of keywords (or concepts) a firm should develop. Following this step, a qualitative filtering process – screening and repetition – was conducted using expert judgments. For this process, experts with more than 5 years of experience in developing new mobile services were invited to adjust the keyword selection process.

The method used to select keywords for each layer is closely related to the application domain. When selecting product keywords, keywords related to the product specification (e.g. wifi), product performance (e.g. battery life), and device characteristics (e.g. speaker) should be included. Similarly, when selecting service keywords, keywords related to the service function and customer utility should be selected. For example, keywords such as e-mail and password are relevant for service function whereas keywords such as interface and notification are related to customer utility. The final keyword list, after the two-step process, is listed in Table 5.

4.3. Construction of keyword vector

Based on the keywords, downloaded documents are then transferred to the keyword vector, more specifically; a

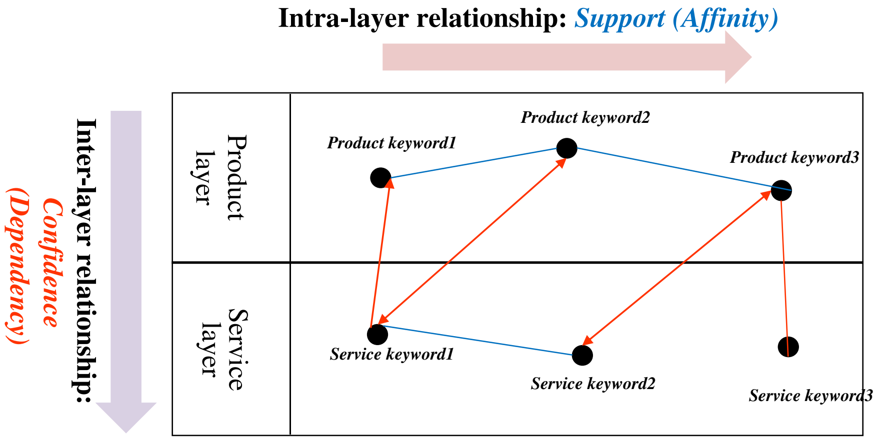

Fig. 4. Keyword relational map.

document–keyword vector in which the row consists of documents and column consists of keywords. This vector is prepared for all documents, in which service documents are converted into keyword vectors, composed of the keyword's existence in the keyword feature list. This is based on assumptions that any given document can be represented by a set of relevant keywords [25]. Since there are 300 documents and 51 keywords (27 product keywords and 24 service keywords), a 300 $*$ 51 matrix is constructed. Each cell is filled with a corresponding value according to whether the keyword appears in a certain document or not.

To calculate the association rule, this document–keyword vector should be converted to a keyword–keyword vector to represent support and confidence between keywords. First, product–product relationship and service–service relationship are calculated based on the support. Second, relationships between product keywords and service keywords are measured via the confidence for measuring dependency.

4.4. Calculate the associations

Based on the keyword vector, support and confidence information is calculated as follows. Suppose P is the product keyword and S is the service keyword, support and confidence are calculated as follows.

$$\mathrm{Support}(P \rightarrow S) = \frac{N(P \cap S)}{N}$$
$$\mathrm{Confidence}(P \rightarrow S) = \frac{P(P \cap S)}{P(P)} = P(S|P) = \frac{N(P \cap S)}{N(P)}.$$

Based on the calculus for support and confidence, associations between keywords are identified. The result of measuring the association rule is shown in Appendix A.

4.5. Develop a keyword-based technology roadmap

4.5.1. Keyword portfolio map

Based on the associations, a keyword portfolio map is developed. For 806 pairs of keywords whose co-occurrence is more than zero, the keyword portfolio map is constructed, as shown in Fig. 5.

As shown in this figure, keyword pairs are distributed in the keyword portfolio map. Table 6 shows the type of keyword pair and its representative pairs. The cutoff value is set as 0.03 for the support value and 0.4 for the confidence value.

4.5.2. Keyword-relational map

Based on the associations, a keyword-relational map is developed. To visualize the relationship between keywords, we employed UCINET, graphical software to represent the relationship between elements. Using affinity as an intra-layer relationship and dependency as the inter-layer relationship, a 2-dimensional map was developed. Fig. 6 shows the results of developing a keyword-relational map. To provide more clear information, we set the cutoff value as 0.4, which means that keyword relationships more than 0.4 are only described in the figure.

As a result, many interesting implications can be identified. Many different kinds of keywords are identified. Firstly, interesting phenomena are found in terms of product influences

Table 5
Keyword list of this case study.

Type	Keyword list
Product keyword	Battery (BT), device (DE), mac (MA), wifi (WF), keyboard (KB), text (TX), camera (CM), flashlight (FL), mms (MM), system (SY), server (SV), audio (AU), dialer (DI), desktop (DT), volume (VO), tv (TV), video (VI), memory (ME), brightness (BR), product (PD), bluetooth (BT), speaker (SP), barcode (BC), GPS (GP), battery life (BL), radar (RD), webcam (WC)
Service keyword	Photo (PT), email (EM), auto (AT), conversion (CV), music (MU), theme (TH), interface (IF), password (PW), browser (BW), converter (CT), location (LC), safari (SF), SMS (SM), calculator (CC), movie (MV), software (SW), iTunes (IT), ringtones (RT), timer (TM), feedback (FB), unit converter (UC), playback (PB), notification (NT), language (LG)

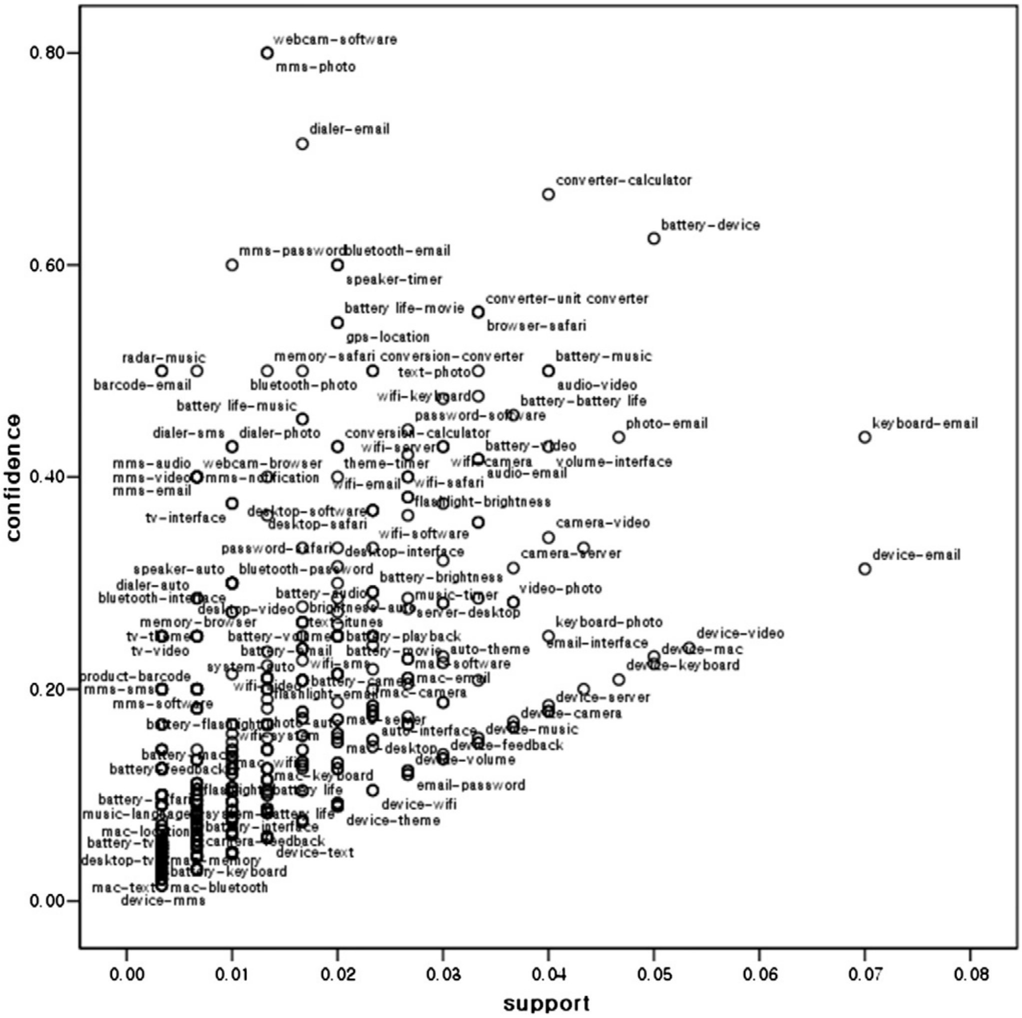

Fig. 5. Keyword portfolio map for utility category.

on service development. Battery life is determined as of critical importance to the development of interactive services, such as music, iTunes, and videos. These services are provided to the customer not only for a certain time, but also for all the time. Therefore, to develop new services which are operated on a long term basis, increase of battery life should be considered an

Table 6
Result of keyword portfolio map.

Type	Keyword pair
Interactive	Battery–device, battery–video, battery–battery life, battery–music, wifi–keyboard, keyboard–email, text–photo, audio–video, audio–email, volume–interface, photo–email, browser–safari, converter–calculator, converter–unit converter
Causal	Wifi–camera, wifi–server, mms–photo, mms–password, dialer–photo, dialer–email, dialer–sms, desktop–email, memory–safari, brightness–music, bluetooth–photo, bluetooth–email, speaker–timer, barcode–email, gps–location, battery life–music, battery life–movie, battery life–playback, radar–auto, radar–music, webcam–software, conversion–interface, conversion–converter, conversion–calculator, conversion–unit converter, password–software
Family	Device–mac, device–keyboard, device–camera, device–server, device–audio, device–video, device–email, device–auto, device–music, device–interface, device–software, device–iTunes, keyboard–photo, keyboard–safari, camera–server, camera–video, camera–software, server–email, volume–converter, video–photo, video–email, video–software, email–auto, email–interface, email–safari, mail–software, email–iTunes
Unrelated	Email–music, email–location, email–feedback, device–theme, device–timer, barcode–safari, notification–language, memory–playback, movie–feedback, timer–language

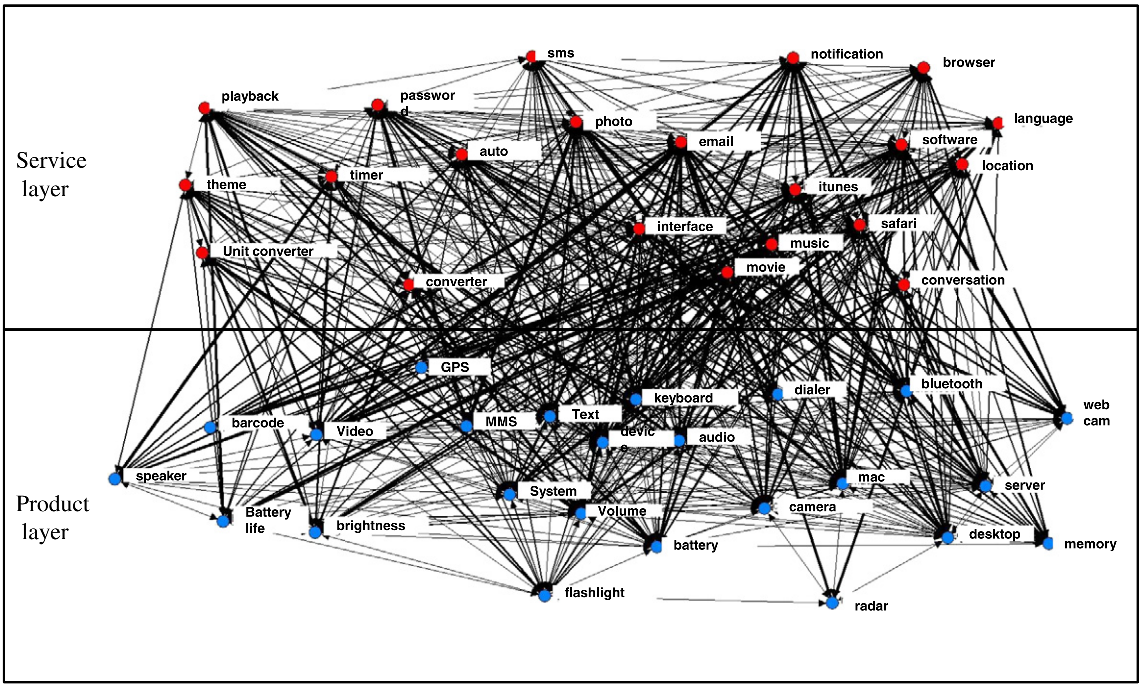

Fig. 6. Keyword-relational map for utility category (cutoff = 0.4).

important issue. A product element, speakers, also significantly affects the timer services or music services. This is quite natural since the timer or music is highly related to the functional performance of audio devices.

Some relationships which seem to be unrelated were also found. For example, a product keyword, mms, is closely related to the service keywords, photo and password. These keywords seem to be unrelated but are strongly linked. mms is generally executed for sending data such as photos or long sentences. The development of data-sending product attributes is thus closely related to the development of data-related services, including security issues. Similarly, a product attribute, video, is also related to e-mail or music, which seems to be unrelated. This can be interpreted in the same way.

From the perspective of new service development, many interesting phenomena can be found. First, notification services should be developed with in-depth consideration of various product elements, such as video, battery, barcode, flashlight, device, keyboard, audio, and camera. This means that development of a new notification service requires these prerequisites for product attributes. In addition, it can be assumed that notification services can be applied to various product elements.

In terms of network centrality, service elements such as e-mail, photos, movie, software, location, and notifications have strong relationships with other elements. When investigating product elements, keywords such as keyboard, audio, speaker, battery, and memory show a strong relationship. In most nodes, mutual interactions between product elements and service elements are identified, which means new products (or new services) are not developed after developing new services (or new products). The development process of a certain product or service is not sequential or dependent, but conducted in parallel, considering both the product attributes and service attributes. This is linked with the recent promising phenomenon of product–service systems in which products and services are integrated to satisfy customers.

In order to show a clearer view, we developed the keyword-relational map using an increased cutoff value, 0.6. The result is shown in Fig. 7.

In the service layer, e-mail is shown to be a critical element and shows a strong interaction with most nodes in a map, not only with the product elements, but also with the service elements. As mentioned earlier, notification services are also closely related to the various product elements, which implies that close relationships with product elements are required to develop new notification services, and vice versa. In the product layer, mms is related to other service elements, which means that it is highly affected by the development of other service elements, and vice versa. Product elements such as video and speaker show a strong relationship between service developments. This denotes changes in product roles. Traditionally, a speaker is used when developing products or the technological elements of a product, such as qualified voice communications. However, now this product attribute is employed for supporting relevant services or facilitating the effective service operations.

What is notable is the close relationship between the product layer and service layer. Compared to the product–product relationship or service–service relationship, inter-layer relationship, i.e., product–service relationship seems to be evident and strong. This denotes the importance of

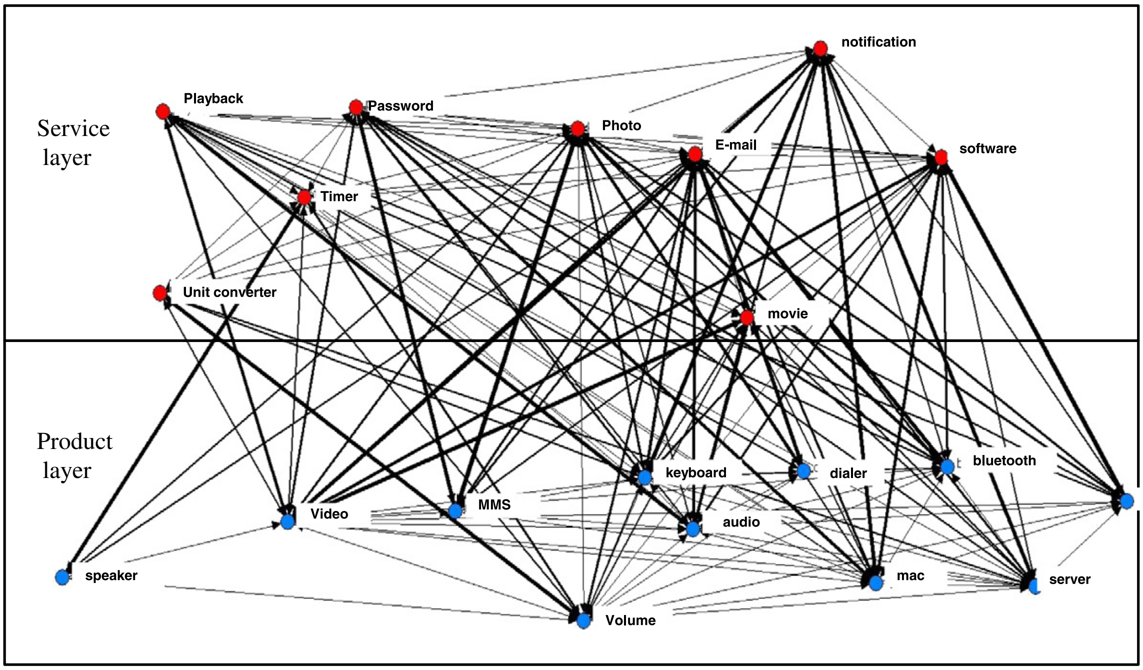

Fig. 7. Keyword-relational map for utility category (cutoff = 0.6).

identifying inter-layer relationships by measuring dependency information.

5. Conclusion

To incorporate business data into the planning process, this paper suggests use of ARM-based technology roadmap to identify the dependency relationship between different layers. Using ARM, two types of associations are measured: support and confidence. Using these measures, two types of keyword-based roadmaps are developed: a keyword portfolio map and a keyword relational map.

This paper contributes to the fields in many ways. First, from a theoretical perspective, this paper provides an ARM-based roadmapping process to effectively visualize the dependency relationship between different layers. Using ARM, this paper addresses the shortcomings of previous research that simply measures the similarities between two elements. Considering the characteristics of inter-layer relationships, we suggest a concept of measuring "dependency" to develop the technology roadmap, and provide a way of measuring dependency between two elements using ARM. In addition, this paper also suggests the use of business data in the proactive planning process, by indicating layer-specific data sources that can be used for systematic planning. From a methodological perspective, this paper combines the ARM and the technology roadmap to support data-driven business planning. Thus, this paper extends the application area of ARM by applying it into the technology roadmapping process.

Despite these contributions to the field, however, this paper is subject to some limitations. First, since the development of technology roadmap mainly focuses on the relationship between two elements, associations are calculated for a keyword pair, i.e. relationship between two keywords. However, using ARM, associations among several keywords can be calculated. Relationships among more than two keywords can provide more detailed but practical information to the firm. Second, the case study of this paper is only conducted for the product layer and the service layer which are considered most important key areas in the technology roadmap. Future research should cover more diversified case examples, including a variety of data sources of technology roadmap. Finally, this study mainly deals with quantifying keyword relationships between roadmap layers. However, what is the most important thing is the identification of functional relationships. As noted in Kostoff and Schaller's [10] work, technology roadmaps are important visualization aids for identifying and characterizing the linkages among different layers. Especially, any changes at a certain node should be reflected to the technology roadmap network through the functional relationships. This denotes the importance of measuring and reflecting functional relationships in the technology roadmap. Therefore, extended from measuring the keyword relationship, how to measure and reflect the functional relationship can be an important future work.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (No. 2011-0030814).

Appendix A. The result of association

Appendix A.1
Result of support (P → P).

	BT	DE	MA	WF	KB	TX	CM	FL	MM	SY	SV	AU	DI	DT	VO	TV	VI	ME	BR	PD	BT	SP	BC	GP	BL	RD	WC
BT	0.080	0.050	0.010	0.000	0.003	0.003	0.017	0.013	0.000	0.017	0.010	0.023	0.000	0.000	0.020	0.003	0.033	0.010	0.023	0.003	0.000	0.010	0.000	0.000	0.037	0.000	0.000
DE	0.050	0.223	0.050	0.023	0.047	0.013	0.037	0.020	0.003	0.020	0.040	0.040	0.000	0.023	0.027	0.007	0.053	0.013	0.017	0.013	0.007	0.013	0.000	0.007	0.030	0.000	0.003
MA	0.010	0.050	0.127	0.033	0.033	0.020	0.057	0.000	0.000	0.023	0.047	0.013	0.000	0.040	0.007	0.003	0.043	0.007	0.003	0.000	0.010	0.000	0.000	0.000	0.000	0.003	0.017
WF	0.000	0.023	0.033	0.070	0.033	0.000	0.030	0.000	0.000	0.010	0.030	0.003	0.000	0.020	0.007	0.003	0.013	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000
KB	0.003	0.047	0.033	0.033	0.160	0.017	0.027	0.013	0.013	0.013	0.027	0.013	0.013	0.030	0.007	0.000	0.013	0.007	0.000	0.007	0.007	0.000	0.000	0.000	0.003	0.000	0.000
TX	0.003	0.013	0.020	0.000	0.017	0.067	0.013	0.000	0.007	0.003	0.003	0.003	0.000	0.003	0.003	0.003	0.027	0.000	0.003	0.007	0.007	0.007	0.003	0.003	0.000	0.000	0.010
CM	0.017	0.037	0.057	0.030	0.027	0.013	0.117	0.000	0.007	0.013	0.037	0.027	0.000	0.020	0.003	0.003	0.040	0.007	0.000	0.003	0.010	0.000	0.000	0.003	0.010	0.003	0.010
FL	0.013	0.020	0.000	0.000	0.013	0.000	0.000	0.073	0.000	0.003	0.000	0.000	0.000	0.000	0.007	0.000	0.003	0.000	0.027	0.003	0.000	0.000	0.000	0.000	0.007	0.003	0.000
MM	0.000	0.003	0.000	0.000	0.013	0.007	0.007	0.000	0.017	0.000	0.003	0.007	0.000	0.000	0.000	0.000	0.007	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SY	0.017	0.020	0.023	0.010	0.013	0.003	0.013	0.003	0.000	0.080	0.010	0.003	0.000	0.010	0.010	0.000	0.013	0.010	0.010	0.003	0.003	0.010	0.000	0.007	0.007	0.000	0.003
SV	0.010	0.040	0.047	0.030	0.027	0.003	0.037	0.000	0.003	0.010	0.093	0.010	0.000	0.027	0.003	0.000	0.023	0.010	0.003	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.003
AU	0.023	0.040	0.013	0.003	0.013	0.003	0.027	0.000	0.007	0.003	0.010	0.080	0.007	0.010	0.017	0.003	0.040	0.003	0.003	0.000	0.007	0.000	0.000	0.007	0.017	0.000	0.007
DI	0.000	0.000	0.000	0.000	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.007	0.023	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000
DT	0.000	0.023	0.040	0.020	0.030	0.003	0.020	0.000	0.000	0.010	0.027	0.010	0.000	0.063	0.003	0.003	0.017	0.003	0.000	0.000	0.010	0.000	0.000	0.000	0.000	0.003	0.003
VO	0.020	0.027	0.007	0.007	0.007	0.003	0.003	0.007	0.000	0.010	0.003	0.017	0.003	0.003	0.093	0.003	0.010	0.000	0.013	0.003	0.003	0.020	0.000	0.000	0.007	0.000	0.003
TV	0.003	0.007	0.003	0.003	0.000	0.003	0.003	0.000	0.000	0.000	0.000	0.003	0.000	0.003	0.003	0.027	0.007	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000
VI	0.033	0.053	0.043	0.013	0.013	0.027	0.040	0.003	0.007	0.013	0.023	0.040	0.000	0.017	0.010	0.007	0.130	0.000	0.000	0.003	0.017	0.007	0.000	0.000	0.027	0.000	0.013
ME	0.010	0.013	0.007	0.000	0.007	0.000	0.007	0.000	0.000	0.010	0.010	0.003	0.000	0.003	0.000	0.000	0.000	0.027	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BR	0.023	0.017	0.003	0.000	0.000	0.003	0.000	0.027	0.000	0.010	0.003	0.003	0.000	0.000	0.013	0.003	0.000	0.000	0.063	0.003	0.000	0.007	0.000	0.000	0.003	0.000	0.000
PD	0.003	0.013	0.000	0.000	0.007	0.007	0.003	0.003	0.003	0.003	0.000	0.000	0.000	0.000	0.003	0.000	0.003	0.000	0.003	0.033	0.000	0.003	0.007	0.003	0.003	0.000	0.000
BT	0.000	0.007	0.010	0.003	0.007	0.007	0.010	0.000	0.000	0.003	0.003	0.007	0.003	0.010	0.003	0.000	0.017	0.000	0.000	0.000	0.033	0.000	0.000	0.000	0.000	0.000	0.000
SP	0.010	0.013	0.000	0.000	0.000	0.007	0.000	0.000	0.000	0.010	0.000	0.000	0.000	0.000	0.020	0.000	0.007	0.000	0.007	0.003	0.000	0.033	0.000	0.000	0.003	0.000	0.000
BC	0.000	0.000	0.000	0.000	0.000	0.003	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	0.007	0.000	0.000	0.013	0.000	0.000	0.000	0.000
GP	0.000	0.007	0.000	0.000	0.000	0.003	0.003	0.000	0.000	0.007	0.000	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.037	0.000	0.000	0.000
BL	0.037	0.030	0.000	0.000	0.003	0.000	0.010	0.007	0.000	0.007	0.000	0.017	0.000	0.000	0.007	0.003	0.027	0.000	0.003	0.003	0.000	0.003	0.000	0.000	0.037	0.000	0.000
RD	0.000	0.000	0.003	0.000	0.000	0.000	0.003	0.003	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.007	0.000
WC	0.000	0.003	0.017	0.000	0.000	0.010	0.010	0.000	0.000	0.003	0.003	0.007	0.000	0.003	0.003	0.000	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.017

Appendix A2
Result of support (S → S).

	PT	EM	AT	CV	MU	TH	IF	PW	BW	CT	LC	SF	SM	CC	MV	SW	IT	RT	TM	FB	UC	PB	NT	LG
PT	0.107	0.047	0.017	0.000	0.010	0.003	0.030	0.017	0.013	0.007	0.003	0.013	0.013	0.007	0.000	0.020	0.023	0.007	0.007	0.003	0.003	0.003	0.010	0.007
EM	0.047	0.217	0.037	0.010	0.020	0.013	0.050	0.027	0.027	0.010	0.020	0.043	0.030	0.017	0.010	0.033	0.040	0.000	0.010	0.020	0.010	0.010	0.010	0.007
AT	0.017	0.037	0.133	0.000	0.023	0.030	0.023	0.020	0.013	0.000	0.003	0.020	0.013	0.000	0.003	0.023	0.017	0.003	0.023	0.023	0.000	0.010	0.000	0.003
CV	0.000	0.010	0.000	0.047	0.000	0.000	0.023	0.000	0.000	0.023	0.003	0.000	0.000	0.020	0.000	0.007	0.010	0.000	0.000	0.003	0.020	0.000	0.000	0.000
MU	0.010	0.020	0.023	0.000	0.097	0.017	0.010	0.003	0.003	0.000	0.003	0.010	0.000	0.000	0.010	0.010	0.023	0.010	0.027	0.010	0.000	0.017	0.003	0.007
TH	0.003	0.013	0.030	0.000	0.017	0.050	0.017	0.003	0.003	0.003	0.000	0.007	0.000	0.003	0.003	0.007	0.000	0.000	0.020	0.007	0.003	0.003	0.000	0.000
IF	0.030	0.050	0.023	0.023	0.010	0.017	0.153	0.013	0.020	0.027	0.013	0.020	0.010	0.023	0.010	0.013	0.007	0.007	0.013	0.023	0.013	0.007	0.000	0.010
PW	0.017	0.027	0.020	0.000	0.003	0.003	0.013	0.060	0.010	0.003	0.003	0.020	0.003	0.003	0.003	0.027	0.013	0.000	0.007	0.007	0.003	0.000	0.010	0.003
BW	0.013	0.027	0.013	0.000	0.003	0.003	0.020	0.010	0.060	0.000	0.003	0.033	0.007	0.000	0.000	0.017	0.010	0.000	0.000	0.007	0.000	0.000	0.000	0.000
CT	0.007	0.010	0.000	0.023	0.000	0.003	0.027	0.003	0.000	0.060	0.003	0.003	0.000	0.040	0.000	0.003	0.003	0.000	0.003	0.003	0.033	0.000	0.000	0.000
LC	0.003	0.020	0.003	0.003	0.003	0.000	0.013	0.003	0.003	0.003	0.087	0.003	0.003	0.007	0.007	0.013	0.007	0.000	0.000	0.007	0.000	0.000	0.003	0.000
SF	0.013	0.043	0.020	0.000	0.010	0.007	0.020	0.020	0.033	0.003	0.003	0.083	0.007	0.003	0.003	0.010	0.023	0.000	0.000	0.010	0.003	0.000	0.003	0.007
SM	0.013	0.030	0.013	0.000	0.000	0.000	0.010	0.003	0.007	0.000	0.003	0.007	0.057	0.000	0.000	0.013	0.003	0.000	0.000	0.000	0.000	0.003	0.000	0.003
CC	0.007	0.017	0.000	0.020	0.000	0.003	0.023	0.003	0.000	0.040	0.007	0.003	0.000	0.077	0.003	0.003	0.010	0.000	0.003	0.000	0.020	0.000	0.000	0.000
MV	0.000	0.010	0.003	0.000	0.010	0.003	0.010	0.003	0.000	0.000	0.007	0.003	0.000	0.003	0.037	0.007	0.000	0.000	0.000	0.003	0.000	0.013	0.000	0.000
SW	0.020	0.033	0.023	0.007	0.010	0.007	0.013	0.027	0.017	0.003	0.013	0.010	0.013	0.003	0.007	0.107	0.030	0.000	0.013	0.010	0.003	0.007	0.010	0.007
IT	0.023	0.040	0.017	0.010	0.023	0.000	0.007	0.013	0.010	0.003	0.007	0.023	0.003	0.010	0.000	0.030	0.097	0.007	0.007	0.010	0.003	0.003	0.007	0.010
RT	0.007	0.000	0.003	0.000	0.010	0.000	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.007	0.020	0.003	0.000	0.000	0.000	0.000	0.003
TM	0.007	0.010	0.023	0.000	0.027	0.020	0.013	0.007	0.000	0.003	0.000	0.000	0.000	0.003	0.000	0.013	0.007	0.003	0.067	0.013	0.003	0.003	0.000	0.003
FB	0.003	0.020	0.023	0.003	0.010	0.007	0.023	0.007	0.007	0.003	0.007	0.010	0.000	0.000	0.003	0.010	0.010	0.000	0.013	0.090	0.000	0.007	0.000	0.000
UC	0.003	0.010	0.000	0.020	0.000	0.003	0.013	0.003	0.000	0.033	0.000	0.003	0.000	0.020	0.000	0.003	0.003	0.000	0.003	0.000	0.033	0.000	0.000	0.000
PB	0.003	0.010	0.010	0.000	0.017	0.003	0.007	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.013	0.007	0.003	0.000	0.003	0.007	0.000	0.043	0.000	0.003
NT	0.010	0.010	0.000	0.000	0.003	0.000	0.000	0.010	0.000	0.000	0.003	0.003	0.000	0.000	0.000	0.010	0.007	0.000	0.000	0.000	0.000	0.000	0.017	0.003
LG	0.007	0.007	0.003	0.000	0.007	0.000	0.010	0.003	0.000	0.000	0.000	0.007	0.003	0.000	0.000	0.007	0.010	0.003	0.003	0.000	0.000	0.003	0.003	0.030

Appendix A3
Result of confidence (P → S).

	BT	DE	MA	WF	KB	TX	CM	FL	MM	SY	SV	AU	DI	DT	VO	TV	VI	ME	BR	PD	BT	SP	BC	GP	BL	RD	WC
PT	0.042	0.119	0.158	0.048	0.250	0.500	0.200	0.045	0.800	0.083	0.143	0.125	0.429	0.211	0.071	0.000	0.282	0.125	0.053	0.200	0.500	0.100	0.000	0.000	0.000	0.000	0.400
EM	0.250	0.313	0.368	0.381	0.438	0.400	0.229	0.182	0.400	0.250	0.357	0.417	0.714	0.474	0.321	0.125	0.282	0.125	0.211	0.300	0.600	0.100	0.500	0.273	0.182	0.500	0.400
AT	0.167	0.149	0.211	0.095	0.125	0.300	0.057	0.227	0.000	0.208	0.214	0.167	0.286	0.211	0.179	0.125	0.128	0.125	0.263	0.000	0.200	0.300	0.000	0.091	0.000	0.000	0.400
CV	0.000	0.060	0.053	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.179	0.000	0.000	0.000	0.250	0.000	0.051	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MU	0.500	0.149	0.105	0.095	0.021	0.050	0.086	0.136	0.000	0.250	0.143	0.167	0.000	0.000	0.250	0.000	0.179	0.375	0.421	0.200	0.100	0.400	0.250	0.000	0.455	0.000	0.000
TH	0.250	0.090	0.000	0.000	0.063	0.050	0.000	0.273	0.000	0.083	0.000	0.042	0.143	0.053	0.143	0.250	0.026	0.000	0.368	0.000	0.000	0.200	0.000	0.000	0.091	0.000	0.000
IF	0.083	0.164	0.132	0.238	0.188	0.250	0.114	0.000	0.000	0.125	0.107	0.208	0.286	0.316	0.429	0.375	0.154	0.250	0.000	0.100	0.300	0.300	0.000	0.000	0.000	0.000	0.000
PW	0.042	0.060	0.211	0.143	0.146	0.200	0.114	0.045	0.600	0.167	0.143	0.083	0.000	0.211	0.000	0.000	0.179	0.000	0.053	0.200	0.300	0.000	0.000	0.000	0.091	0.000	0.200
BW	0.000	0.090	0.184	0.143	0.125	0.200	0.114	0.000	0.000	0.000	0.036	0.042	0.000	0.211	0.036	0.000	0.077	0.250	0.000	0.100	0.200	0.000	0.250	0.000	0.000	0.000	0.400
CT	0.042	0.015	0.000	0.000	0.042	0.000	0.000	0.136	0.000	0.083	0.036	0.000	0.000	0.000	0.357	0.000	0.026	0.125	0.053	0.000	0.000	0.000	0.000	0.000	0.091	0.000	0.000
LC	0.000	0.045	0.105	0.000	0.021	0.100	0.143	0.045	0.000	0.125	0.036	0.167	0.000	0.105	0.071	0.000	0.051	0.000	0.000	0.200	0.100	0.100	0.250	0.545	0.000	0.000	0.200
SF	0.083	0.134	0.237	0.381	0.208	0.100	0.114	0.045	0.000	0.208	0.214	0.000	0.000	0.368	0.000	0.125	0.077	0.500	0.053	0.100	0.200	0.000	0.250	0.000	0.091	0.000	0.000
SM	0.000	0.045	0.026	0.238	0.167	0.100	0.057	0.000	0.200	0.042	0.036	0.042	0.429	0.000	0.036	0.000	0.026	0.000	0.000	0.100	0.100	0.000	0.000	0.000	0.000	0.000	0.000
CC	0.042	0.045	0.026	0.000	0.063	0.000	0.057	0.136	0.000	0.125	0.071	0.000	0.000	0.000	0.214	0.000	0.051	0.125	0.053	0.000	0.000	0.000	0.000	0.000	0.091	0.000	0.000
MV	0.250	0.075	0.053	0.000	0.042	0.050	0.114	0.000	0.000	0.000	0.071	0.208	0.000	0.053	0.071	0.125	0.179	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.545	0.000	0.000
SW	0.167	0.179	0.395	0.333	0.167	0.200	0.286	0.136	0.200	0.167	0.214	0.292	0.000	0.368	0.071	0.125	0.333	0.125	0.105	0.000	0.300	0.000	0.000	0.000	0.182	0.000	0.800
IT	0.042	0.164	0.316	0.238	0.125	0.250	0.171	0.045	0.000	0.167	0.250	0.167	0.000	0.263	0.071	0.125	0.231	0.125	0.053	0.100	0.300	0.000	0.250	0.182	0.000	0.000	0.200
RT	0.000	0.000	0.000	0.000	0.000	0.050	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	0.100	0.000	0.000	0.000	0.000	0.000
TM	0.250	0.090	0.026	0.000	0.000	0.050	0.000	0.091	0.000	0.167	0.000	0.042	0.000	0.053	0.250	0.125	0.077	0.000	0.263	0.100	0.000	0.600	0.000	0.000	0.091	0.000	0.000
FB	0.125	0.134	0.053	0.048	0.104	0.050	0.057	0.045	0.000	0.083	0.036	0.208	0.000	0.105	0.071	0.000	0.077	0.250	0.105	0.000	0.000	0.000	0.000	0.273	0.091	0.000	0.000
UC	0.042	0.015	0.000	0.000	0.021	0.000	0.000	0.136	0.000	0.042	0.036	0.000	0.000	0.000	0.214	0.000	0.026	0.000	0.053	0.000	0.000	0.000	0.000	0.000	0.091	0.000	0.000
PB	0.250	0.060	0.026	0.000	0.021	0.050	0.143	0.000	0.000	0.042	0.036	0.375	0.143	0.000	0.107	0.000	0.179	0.125	0.105	0.100	0.100	0.000	0.000	0.000	0.455	0.000	0.200
NT	0.000	0.030	0.079	0.048	0.063	0.000	0.057	0.000	0.400	0.083	0.107	0.083	0.000	0.105	0.036	0.000	0.077	0.000	0.000	0.000	0.100	0.000	0.000	0.000	0.000	0.000	0.200
LG	0.000	0.015	0.053	0.095	0.021	0.000	0.029	0.045	0.000	0.125	0.036	0.042	0.000	0.000	0.071	0.000	0.026	0.000	0.053	0.000	0.100	0.000	0.000	0.000	0.000	0.500	0.000

Appendix A4
Result of confidence ($S \rightarrow P$).

	PT	EM	AT	CV	MU	TH	IF	PW	BW	CT	LC	SF	SM	CC	MV	SW	IT	RT	TM	FB	UC	PB	NT	LG
BT	0.031	0.092	0.100	0.000	0.414	0.400	0.043	0.056	0.000	0.056	0.000	0.080	0.000	0.043	0.545	0.125	0.034	0.000	0.300	0.111	0.100	0.462	0.000	0.000
DE	0.250	0.323	0.250	0.286	0.345	0.400	0.239	0.222	0.333	0.056	0.115	0.360	0.176	0.130	0.455	0.375	0.379	0.000	0.300	0.333	0.100	0.308	0.400	0.111
MA	0.188	0.215	0.200	0.143	0.138	0.000	0.109	0.444	0.389	0.000	0.154	0.360	0.059	0.043	0.182	0.469	0.414	0.000	0.050	0.074	0.000	0.077	0.600	0.222
WF	0.031	0.123	0.050	0.000	0.069	0.000	0.109	0.167	0.167	0.000	0.000	0.320	0.294	0.000	0.000	0.219	0.172	0.000	0.000	0.037	0.000	0.000	0.200	0.222
KB	0.375	0.323	0.150	0.000	0.034	0.200	0.196	0.389	0.333	0.111	0.038	0.400	0.471	0.130	0.182	0.250	0.207	0.000	0.000	0.185	0.100	0.077	0.600	0.111
TX	0.313	0.123	0.150	0.000	0.034	0.067	0.109	0.222	0.222	0.000	0.077	0.080	0.118	0.000	0.091	0.125	0.172	0.167	0.050	0.037	0.000	0.077	0.000	0.000
CM	0.219	0.123	0.050	0.000	0.103	0.000	0.087	0.222	0.222	0.000	0.192	0.160	0.118	0.087	0.364	0.313	0.207	0.000	0.000	0.074	0.000	0.385	0.400	0.111
FL	0.031	0.062	0.125	0.000	0.103	0.400	0.000	0.056	0.000	0.167	0.038	0.040	0.000	0.130	0.000	0.094	0.034	0.000	0.100	0.037	0.300	0.000	0.000	0.111
MM	0.125	0.031	0.000	0.000	0.000	0.000	0.000	0.167	0.000	0.000	0.000	0.000	0.059	0.000	0.000	0.031	0.000	0.000	0.000	0.000	0.000	0.000	0.400	0.000
SY	0.063	0.092	0.125	0.000	0.207	0.133	0.065	0.222	0.000	0.111	0.115	0.200	0.059	0.130	0.000	0.125	0.138	0.000	0.200	0.074	0.100	0.077	0.400	0.333
SV	0.125	0.154	0.150	0.357	0.138	0.000	0.065	0.222	0.056	0.056	0.038	0.240	0.059	0.087	0.182	0.188	0.241	0.000	0.000	0.037	0.100	0.077	0.600	0.111
AU	0.094	0.154	0.100	0.000	0.138	0.067	0.109	0.111	0.056	0.000	0.154	0.000	0.059	0.000	0.455	0.219	0.138	0.000	0.050	0.185	0.000	0.692	0.400	0.111
DI	0.094	0.077	0.050	0.000	0.000	0.067	0.043	0.000	0.000	0.000	0.000	0.000	0.176	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.077	0.000	0.000
DT	0.125	0.138	0.100	0.000	0.000	0.067	0.130	0.222	0.222	0.000	0.077	0.280	0.000	0.000	0.091	0.219	0.172	0.000	0.050	0.074	0.000	0.000	0.400	0.000
VO	0.063	0.138	0.125	0.500	0.241	0.267	0.261	0.000	0.056	0.556	0.077	0.000	0.059	0.261	0.182	0.063	0.069	0.000	0.350	0.074	0.600	0.231	0.200	0.222
TV	0.000	0.015	0.025	0.000	0.000	0.133	0.065	0.000	0.000	0.000	0.000	0.040	0.000	0.000	0.091	0.031	0.034	0.000	0.050	0.000	0.000	0.000	0.000	0.000
VI	0.344	0.169	0.125	0.143	0.241	0.067	0.130	0.389	0.167	0.056	0.077	0.120	0.059	0.087	0.636	0.406	0.310	0.000	0.150	0.111	0.100	0.538	0.600	0.111
ME	0.031	0.015	0.025	0.000	0.103	0.000	0.043	0.000	0.111	0.056	0.000	0.160	0.000	0.043	0.000	0.031	0.034	0.000	0.000	0.074	0.000	0.077	0.000	0.000
BR	0.031	0.062	0.125	0.000	0.276	0.467	0.000	0.056	0.000	0.056	0.000	0.040	0.000	0.043	0.000	0.063	0.034	0.000	0.250	0.074	0.100	0.154	0.000	0.111
PD	0.063	0.046	0.000	0.000	0.069	0.000	0.022	0.111	0.056	0.000	0.077	0.040	0.059	0.000	0.000	0.000	0.034	0.000	0.050	0.000	0.000	0.077	0.000	0.000
BT	0.156	0.092	0.050	0.000	0.034	0.000	0.065	0.167	0.111	0.000	0.038	0.080	0.059	0.000	0.000	0.094	0.103	0.000	0.000	0.000	0.000	0.077	0.200	0.111
SP	0.031	0.015	0.075	0.000	0.138	0.133	0.065	0.000	0.000	0.000	0.038	0.000	0.000	0.000	0.000	0.000	0.000	0.167	0.300	0.000	0.000	0.000	0.000	0.000
BC	0.000	0.031	0.000	0.000	0.034	0.000	0.000	0.000	0.056	0.000	0.038	0.040	0.000	0.000	0.000	0.000	0.034	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GP	0.000	0.046	0.025	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.231	0.000	0.000	0.000	0.000	0.000	0.069	0.000	0.000	0.111	0.000	0.000	0.000	0.000
BL	0.000	0.031	0.000	0.000	0.172	0.067	0.000	0.056	0.000	0.056	0.000	0.040	0.000	0.043	0.545	0.063	0.000	0.000	0.050	0.037	0.100	0.385	0.000	0.000
RD	0.000	0.015	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	0.000	0.000	0.000	0.111
WC	0.063	0.031	0.050	0.000	0.000	0.000	0.000	0.056	0.111	0.000	0.038	0.000	0.000	0.000	0.000	0.125	0.034	0.000	0.000	0.000	0.000	0.077	0.200	0.000

References

[1] R. Agrawal, T. Imielinski, A. Swami, Mining association between sets of items in large database, Proceedings of the ACM-SIGMOD International Conference on Management of Data, Washington, DC, 1993.

[2] O.H. Bray, M.L. Garcia, Technology roadmapping: the integration of strategic and technology planning for competitiveness, Proceedings of the Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, 1997.

[3] B. Brown, M. Chui, J. Manyika, Are you ready for the era of 'big data'? McKinsey Q. 4 (2011) 24–35.

[4] J. Bughin, J. Livingston, S. Marwaha, Seizing the potential of 'big data', McKinsey Q. 4 (2011) 103–109.

[5] E. Dumbill, Making sense of big data, Big Data 1 (1) (2012) 1–2.

[6] P. Groenveld, Roadmapping integrates business and technology, Res. Technol. Manag. 40 (5) (1997) 48–55.

[7] J. Han, M. Kamber, Data Mining: Concepts and Techniques, Academic Press, San Diego, United States, 2001.

[8] C. Holmes, M. Ferrill, The application of operation and technology roadmapping to aid Singaporean SMEs identify and select emerging technologies, Technol. Forecast. Soc. Chang. 72 (3) (2005) 349–357.

[9] Z. Huang, X. Lu, H. Duan, Mining association rules to support resource allocation in business process management, Expert Syst. Appl. 38 (8) (2011) 9483–9490.

[10] R.N. Kostoff, R.R. Schaller, Science and technology roadmaps, IEEE Trans. Eng. Manag. 48 (2) (2001) 132–143.

[11] R.N. Kostoff, R. Boylan, G.R. Simons, Disruptive technology roadmaps, Technol. Forecast. Soc. Chang. 71 (1) (2004) 141–159.

[12] S. Kotsiantis, D. Kanellopoulos, Association rules mining: a recent overview, GESTS Int. Trans. Comput. Sci. Eng. 32 (1) (2006) 71–82.

[13] S. Lavalle, E. Lesser, R. Shockley, Big data, analytics and the path from insights to value, MIT Sloan Manag. Rev. 52 (2) (2011) 21–34.

[14] S. Lee, Y. Park, Customization of technology roadmaps according to roadmapping purposes: overall process and detailed modules, Technol. Forecast. Soc. Chang. 72 (5) (2005) 567–583.

[15] S. Lee, S. Lee, H. Seol, Y. Park, Using patent information for designing new product and technology: keyword based technology roadmapping, R&D Manag. 38 (2) (2008) 169–188.

[16] S. Lee, B. Yoon, C. Lee, J. Park, Business planning based on technological capabilities: patent analysis for technology-driven roadmapping, Technol. Forecast. Soc. Chang. 76 (6) (2009) 769–786.

[17] G.S. Linoff, M.J.A. Berry, Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management, 3rd edition Wiley Publishing, Indianapolis, Indiana, 2011.

[18] D.R. Liu, Y.Y. Shih, Integrating AHP and data mining for product recommendation based on customer lifetime value, Inf. Manag. 42 (3) (2005) 387–400.

[19] C. Lynch, Big data: how do your data grow? Nature 455 (7209) (2008) 28–29.

[20] J. Manyika, M. Chi, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, A.H. Byers, Big Data: The Next Frontier for Innovation, Competition, and Productivity, McKinsey Global Institute, Jun. 2011.

[21] R. Phaal, C.J.P. Farrukh, D.R. Probert, T-Plan—The Fast-start to Technology Roadmapping: Planning Your Route to Success, Institute for Manufacturing, University of Cambridge, Cambridge, UK, 2001.

[22] R. Phaal, C.J.P. Farrukh, D.R. Probert, Technology roadmapping—a planning framework for evolution and revolution, Technol. Forecast. Soc. Chang. 71 (1) (2004) 5–26.

[23] D. Snchez, M.A. Vila, L. Cerda, J.M. Serrano, Association rules applied to credit card fraud detection, Expert Syst. Appl. 36 (2) (2009) 3630–3640.

[24] J.H. Suh, S.C. Park, Service-oriented TechnologyRoadmap (SoTRM) using patent map for R&D strategy of service industry, Expert Syst. Appl. 36 (3) (2009) 6754–6772.

[25] B. Yoon, Y. Park, A systematic approach for identifying technology opportunities: keyword-based morphology analysis, Technol. Forecast. Soc. Chang. 72 (2) (2005) 145–160.

[26] B. Yoon, R. Phaal, D. Probert, Morphology analysis for technology roadmapping: application of text mining, R&D Manag. 38 (1) (2008) 51–68.

[27] B. Yoon, R. Phaal, D. Probert, Structuring technological information for technology roadmapping: data mining approach, Proceedings of the 7th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases, World Scientific and Engineering Academy and Society (WSEAS), 2008, pp. 417–422

Youngjung Geum received her Ph.D. on Industrial Engineering at Seoul National University (SNU). She is currently an assistant professor of Industrial and Systems Engineering at Seoul National University of Science and Technology, Seoul, Korea. Her main research interests include technology planning, product–service integration, and new product/service development.

Hyeonjeong Lee is a research engineer in LG Electronics. She received the M. S. on industrial engineering from Seoul National University (SNU). Her main job is structural and nonstructural data mining for retrieving customer insights.

Youngjo Lee is a professor of statistics at Seoul National University, Seoul, Korea. He received his M.S. and Ph.D. degrees in Statistics from the Iowa State University, in 1981 and 1983, respectively. His research interests include random effect models and hierarchical generalized linear models. He has authored many articles published in statistics.

Yongtae Park received his M.S. and Ph.D. degrees in Operations Management from the University of Wisconsin, Madison, in 1987 and 1991, respectively. He is currently a professor of industrial engineering at Seoul National University, Seoul, Korea. His research interests include service engineering, new service development, and knowledge network analysis. He has authored many articles published in several journals.