J. Eng. Technol. Manage. 39 (2016) 45–64

Contents lists available at ScienceDirect

Journal of Engineering and Technology Management

Patent portfolio analysis of the cloud computing industry

Jia-Yen Huang*

Department of Information Management, National Chin-Yi University of Technology, No.57, Sec. 2, Zhongshan Rd., Taiping Dist., Taichung City 41170, Taiwan, ROC

ARTICLE INFO

Article history:
Received 11 May 2014
Received in revised form 25 January 2016
Accepted 31 January 2016
Available online 11 February 2016

Keywords:
Cloud computing
Patent portfolio
Patent indicators
TOPSIS
MDS
GRA

ABSTRACT

In recent years, cloud computing has become a popular theme in the area of network technology; however, related literature on patent analysis is limited. This study seeks to reveal the technological trends and competition status of cloud computing industry by proposing a hybrid patent portfolio analysis scheme. First, the technique for order preference by similarity to ideal solution (TOPSIS) is adopted to integrate the relevant indicators of patent quality into new indices of patent quality, which are normally evaluated by the cited ratio only. Next, multivariate analysis techniques are employed to provide supplementary information to the R&D decision maker. By conducting factor analysis (FA) on patent class codes under the international patent classification (IPC), this study reveals that there are three mainstream technologies of cloud computing: including virtualization and information retrieval, network system, and commercial data process. This study not only uses multidimensional scaling analysis (MDS) to illustrate the proximity of technologies and firms on a perceptual map, but also applies the grey relational analysis (GRA) method to provide quantitative data for interpreting the perceptual relations. Based on the analysis results, the technological strength and the R&D strategies of several big companies are investigated. The findings of this study can provide valuable references for enterprises that wish to develop technologies and deploy their patent portfolios of cloud computing.

1. Introduction

In recent years, many enterprises have considered cloud computing as a seminal technology. Numerous cloud computing services and platforms have increased dramatically, including notable examples such as Google's File System (GFS), Amazon's Dynamo, and Microsoft's Azure. Owing to rapid market and technological changes, network-related enterprises must monitor the trends of technological development from time to time. A high-tech enterprise needs to make strategic decisions based on the information acquired on the volatility of technology in order to chart its direction in the marketplace, which involve determining the market segment in which it will compete and the competitive position that it will take. To this end, patent portfolio analysis has been employed by many enterprises and proved to be a very usable tool for R&D decision makers (Ernst and Omland, 2011).

Nowadays, the effective use of scarce resources in R&D projects to yield the most profound and sustainable advantages over the increasingly fierce competition are becoming more and more important (Mohr et al., 2010). The technological trend of cloud computing has so far been actively driven by certain enterprises that control most of the market share. Among them,

* Fax: +886 4 23935902.
E-mail address: jygiant@ncut.edu.tw (J.-Y. Huang).

http://dx.doi.org/10.1016/j.jengtecman.2016.01.002
0923-4748/

20 cloud computing enterprises (henceforth called the samples or the targeted enterprises) nominated from specialized marketing reports (magazines such as CIO, Network World, and Computer Reseller News) are targeted for analysis. These include Adobe (Ad.), Amazon (Am.), Apple (Ap.), Cisco (Cis.), Citrix (Cit.), Dell (De.), EMC (EM.), Google (Go.), Huawei (Hu.), Juniper (Ju.), Microsoft (Mi.), Novell (No.), Oracle (Or.), Parallels (Pa.), Red Hat (Re.), Salesforce (Sa.), SAP (SAP), Sun (Su.), Verizon (Ve.), and VMware (VM.). Fast followers who try to imitate the leader's successful business model should make appropriate strategy decisions before entering this market. This motivates us to investigate the patent portfolios and the R&D planning of big companies in the cloud computing industry.

The remainder of this paper is organized as follows: the status of the patent analysis of cloud computing is presented in Section 2. Section 3 provides a description of the proposed compound policy for retrieving the patents of cloud computing. In Section 4, based on the new indexes that are integrated from several relevant indicators of patent quality through the technique for order preference by similarity to ideal solution (TOPSIS), we conduct patent portfolio analysis at both the company and the technological levels. In Section 5, we employ factor analysis (FA), multidimensional scaling analysis (MDS), and grey relational analysis (GRA) to investigate the R&D strategies and competence statuses of all enterprises. Finally, Section 6 contains our concluding remarks and suggestions for future research. One may refer to Fig. 1 for the overall process of the proposed hybrid patent portfolio analysis scheme.

2. Literature review

Cloud computing is a style of computing where scalable and elastic IT-related capabilities are provided as a service to external customers using Internet technologies (Bal, 2012; Madhavaiah et al., 2012). In fact, cloud computing is not a new technology; it is a system that allows data to be located centrally and accessed by businesses through a network. It is similar to the concepts that have been recognized since the 1950s in the work done by AT&T in the area of telephone networking.

The essence of cloud computing is inherited from distributed computing and grid computing (Li et al., 2015). Distributed computing is a field of computer science that studies distributed systems, which consist of multiple autonomous computers that communicate through a computer network, and the computers interact with one another in order to achieve a common task that a single computer would not be able to do. Grid computing is a technology that applies the resources of many computers in a network to a single problem at the same time. One example of grid computing in the public domain is the ongoing Search for Extraterrestrial Intelligence (SETI)@Home project that started in 1999; since then, more than 5 million computers of participants are used to analyze the operation of radio signals with the hope of finding life in outer space.

The National Institute of Standards and Technology (NIST) clearly defined the following three service models related to cloud computing (known as the SPI model): SaaS (software as a service), PaaS (platform as a service), and IaaS (infrastructure as a service). SaaS is a network that provides various softwares for users; PaaS offers a full or partial application development

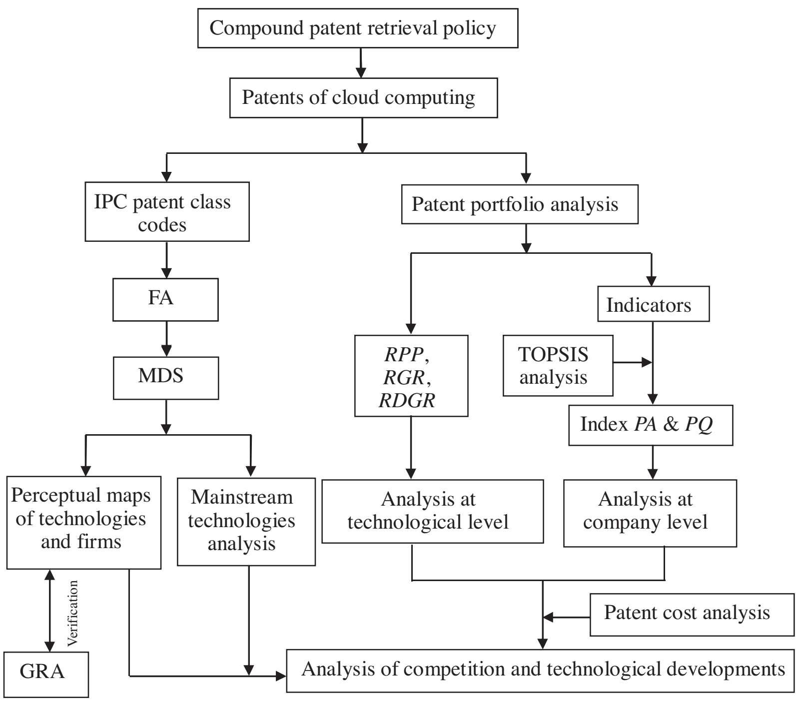

Fig. 1. Overall framework of hybrid patent portfolio analysis.

environment that users can access and utilize online, even in collaboration with others; IaaS provides infrastructure (such as an IT system and a database) to enterprises for renting and data saving.

As pointed out by Sultan (2010), cloud computing may not be suitable for all enterprises. Kaur et al. (2012) addressed seven potential threats for cloud computing services, including abuse and nefarious use of cloud computing, insecure interfaces and APIs, malicious insider, shared technology issues, data loss and leakage, account or service hijacking, and unknown risk profile. However, the market for cloud computing is evolving rapidly (Yeboah-Boateng and Cudjoe-Seshie, 2013). Economy, simplicity, and convenience are the key elements that propel the growth of cloud computing (Erdogmus, 2009). Enterprises are expected to cut operational and capital costs, and allow IT departments to focus on strategic projects rather than keep the datacenters running (Velte et al., 2009).

Apart from being a tool for the protection of intelligent property, patent rights have long been used as a tactic and competitive vehicle for creating higher profits for the enterprise (Chakrabarti, 1991). Patent portfolio is a collection of patents owned by an enterprise. A well-crafted patent portfolio may be used for a variety of business objectives, such as bolstering market position, protecting research and development efforts, generating revenue, and encouraging favorable cross-licensing or settlement agreements.

In his patent portfolio analysis of German machine-tool companies, Ernst (1998) combined the company level with the technological level proposed by Brockhoff (1992) in order to evaluate each company's technological ability and strategic R&D planning. He characterized the patenting strategies at the company level according to two different dimensions: patent activity and patent quality. Whereas patent activity measures the level of R&D activities, and usually uses patent applications as the fundamental indicator, patent quality measures the impact of these activities and is normally evaluated by indicators that include patents granted, patent applications, and patent citations. The patent portfolio at the technological level also has two dimensions: the relative patent position and the attractiveness of technology. The former is derived from the number of patent applications belonging to the firm relative to the number of patent applications from its competitors. The latter is assessed by using the growth rates of the patent applications of each technological field.

Tseng et al. (2011) used a-Si thin-film solar cells as an example to show a hybrid method of patent portfolio analysis. They combined the indicators proposed by Ernst (1998) with those from CHI research Inc. to maximize the number of potential indicators. Then, they adopted the FA method to extract the patent indicators' potential characteristics and condense them into a few factors. Based on these extracted factors, they categorized corporations and compared patent portfolios among the main competitive corporations to identify the leaders of the industry.

Hsieh (2013) integrated the patent value indicators suggested by Harhoff et al. (2003) and Ernst (2003), and proposed a hybrid method to assess patent value and determine a strategy in the early stage of commercialization. According to the benefits and risk factors extracted from the FA method, he or she categorized the patents into four groups. For each group of patents, possible strategies for commercialization were proposed.

Ernst (1995) proposed that patent quality is the sum of the relative measures of indicators that include patents granted, valid patents, patent applications in major foreign countries, and patent citations. However, since these indicators are different in unit, it seems inappropriate to determine the patent quality by adding these indicators directly. This problem exists not only in Ernst's study, but also in the aforementioned literature. Although Tseng et al. (2011) used extracted factors to measure the corporations' overall patent performance instead of the traditional approach of analyzing patent quality and quantity only, they still in fact added these indicators in a direct way, since the relationship between the latent variable (combined indexes) and the observed variables (indicators) can be expressed in a weighted linear combination in the FA method.

Most of the studies on cloud computing have focused on technological developments, security, and commercial applications (see Nimkar and Ghosh, 2013; Cartlidge and Clamp, 2014). Thus far, there has been limited published studies on the patent analysis of cloud computing. Chiou (2010) applied bibliometric analysis to investigate the evolution of the life cycle of cloud computing and showed that the development of cloud computing is still in the growing stage. Huang et al. (2012) employed traditional method proposed by Ernst (1998) to conducted patent portfolio analysis on cloud computing, and they targeted six enterprises of cloud computing only. Both studies considered cloud computing as a technological extension of distributed computing, and thus limited the range of patent retrieval in the international patent classification (IPC) classes of G06 (including electronic data-processing, data handling system or data processing means, and information storage memory) and H04 (telecommunication and communication). Due to these reasons, their investigation results on the technological trends of cloud computing are questionable.

In general, patent analysis is divided into quantitative and qualitative analysis. This study, based on the bibliographical information contained in patent documents, focuses on a quantitative analysis for statistically evaluating the status of technology-based activity in cloud computing industry. Huang (2016) conducted a qualitative analysis on the technical content of cloud computing patents by employing a text mining scheme. He investigated the overall patent structure and identified the key technologies and important patents establishing foundation of cloud technologies by using visual networks and technology centrality indexes. The result is adopted in this study to assist in the extraction of managerial insights, especially in the technological trends.

3. The compound patent retrieval policy

Before conducting patent portfolio analysis, patents related to cloud computing must be retrieved first. Although relevant patents of cloud computing should be classified under IPC G06 and/or H04 (Huang et al., 2012), these IPC classes are vast classes that contain much that cannot be considered as cloud computing. Since G06 and H04 are significantly broad classes, the definitions of which are open to different interpretations, Huang et al.'s study is evidently inappropriate for investigating the R&D strategies of cloud computing. Therefore, this study proposes a compound patent retrieval policy to examine the cloud computing patents.

There are three kinds of strategies generally applied to patent retrieval, including known item search, citation pearl growing, and block building (Trappey et al., 2012). The first method is suitable for analyzers who already hold some bibliometric information. The second method is applicable to identify relevant patents based on the citation relationship with patents on hand. Analyzers may use the third method, the block building, when the retrieving topics containing a number of concepts. In this study, we proposed a two-stage retrieval method which mainly comprises the stages of pearl patents search and block building search, as depicted in Fig. 2. Detailed processes are presented below.

Patent retrieval is the basis of patent analysis, however it is very time-consuming, and the scope of targeted patents is difficult to define. Since cloud computing is inherited from distributed computing and grid computing, the range of patent retrieval is limited to IPC classes G06 and H04. In our study, we retrieve patent data via the Patent Guider, an online searching tool developed by Learningtech Corp., by restricting the title, abstract, and claim with the keyword "cloud computing." Judging from the retrieved 104 patents, one may find that they all belong to the classes of G06 and H04. However, as classes G06 and H04 are too broad for cloud computing, further processes should be conducted to determine the adequate IPC class codes of cloud computing, as depicted in Fig. 2.

The rationale of the two-stage search is to first find a small number of patents (the pearls) which surely belong to cloud computing; then, based on the features of these patents, to acquire more patents from the United States Patent and Trademark Office (USPTO) database via a block-building search. Since the scope of the retrieved patents after extension may be too broad, we employ the constraint of targeted enterprises to screen out the unwanted patents.

This study retrieves patent data by using "cloud computing" as a query string and limits the search domain to title, abstract, and claim parts of the patents. In total, 104 patents were retrieved. Through manual reading of the retrieved patents, one may find that they all belong to the classes of G06 and H04; however, only 64 "pearl" patents out of the 104 patents are identified as cloud patents.

To extend the searching range to include more cloud patents into the database, we extract keywords of each service model from the pearl patents and the cloud ontology. We select the top-50 keywords of each service model according to the value of

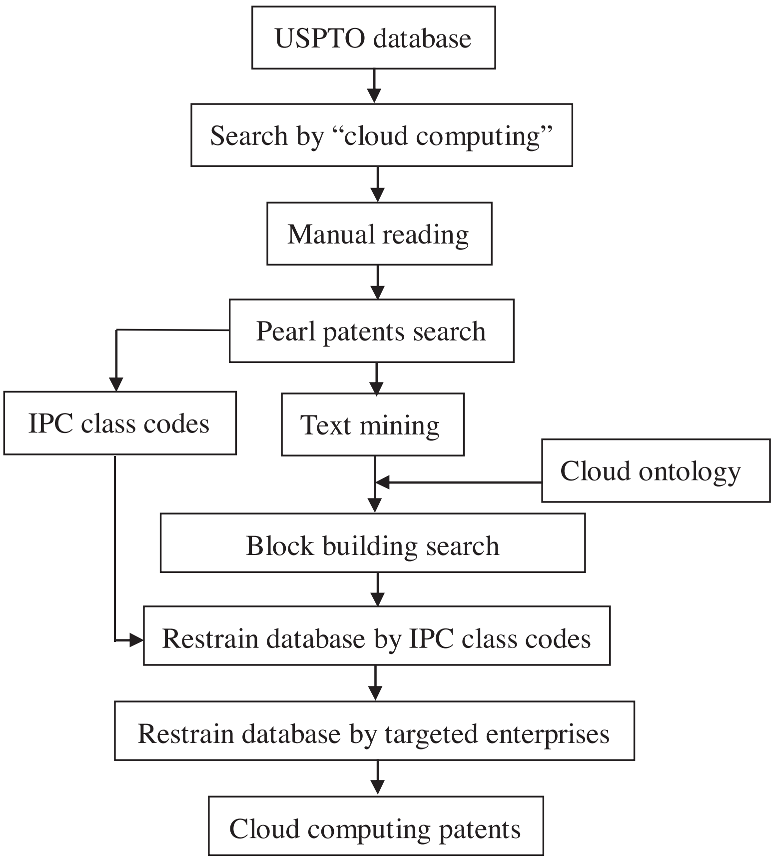

Fig. 2. The compound policy for retrieving patents of cloud computing.

TF-IDF and the opinions of cloud experts (a senior patent engineer and a R&D manager from company in cloud computing industry). The TF-IDF is the product of two terms: term frequency (TF); and inverse document frequency (IDF) (Tseng et al., 2007). Apart from the keywords in common; such as virtual; database; network; computing; device; application; etc.; there are 7 keywords remained in IaaS (network bandwidth; storage space; application servers; operation system; memory; mobile; thermal); 5 in PaaS (data security; platform; worker role; metering; host); and 10 in SaaS (domain name; computer security; graphic; game; communication; access control; wireless; intrusion; remote; software).

The second stage of the patent retrieval process uses the block building search technique to formulate the search statement. The block building approach starts with single concept searches, which usually result in a very large number of hits. By combining single concept searches with appropriate Boolean operators, the complex search task can be simplified and the number of hits can be reduced (Chu, 2003).

Two block buildings are included in this process. In the first block building retrieval process, the study uses “cloud computing” as the search concept, and the searching scope covers all parts of the patents. During the second block building retrieval process, the keywords screened by TF-IDF are used as the search base. By connecting the first block building and the second block building with Boolean operator “AND”, this study retrieved a higher number of possible patents related to cloud computing.

Since many irrelevant patents may have been included during the expansion process, we filter these out by using two restraints: IPC class codes and targeted enterprises. The structure of an IPC classification is made up of a section, class, subclass, main group, and sub-group. Since the classes G06 and H04 are too broad for cloud computing, the subclasses, main group, and sub-group belong to these two classes should be considered. Therefore, we delete the patents that do not belong to the IPC class codes that have been summed up from the pearl. We further remove the patents that do not belong to the targeted enterprises. Finally, the top 20 patent class codes and patent numbers of each targeted enterprise are determined, as listed in Table A1.

To minimize the space used for presenting the results, we use two or three letters to represent the enterprise. We also use variable names IPC1 to IPC20 to represent the patent class codes, as follows: G06F003/048, G06F007/00, G06F007/04, G06F009/455, G06F009/46, G06F011/00, G06F013/00, G06F015/16, G06F015/173, G06F015/177, G06F017/00, G06F017/30, G06F019/00, G06F021/00, G06Q010/00, G06Q030/00, H04J001/16, H04L012/28, H04L012/66, and H04M003/42.

4. Patent portfolio analysis

In this study, patent portfolios at the company and the technological levels of the cloud computing industry are investigated, respectively.

4.1. Analysis at the company level

Henderson et al. (1997) argued that the dispersion of citations made across different patent classes is a measure of the importance of the invention, and that the number of references cited in a patent is evidence of a patent’s value. After examining data on every patent that was issued between 1963 and 1999 as well as every patent lawsuit that terminated during 1999–2000, Allison et al. (2003) concluded that valuable patents are subjected to a more intensive prosecution process involving more claims, more citations of prior art, and longer durations for the issuing of patents. Lanjouw and Schankerman (2000) used the number of different IPC classes into which an invention is placed by the Patent and Trademark Office as evidence of the invention’s value. Similar to the results of Lerner’s (1994) study, Allison et al. (2003) reported that broader patents are more valuable than narrower ones. Harhoff et al. (2003) argued that the more prior art considered, the more thorough the review; for instance, they claimed that the number of references cited in a patent is evidence of a patent’s likely validity, and therefore, its value.

The traditional index of patent activity is evaluated by the patent application; however, the R&D results of an enterprise do not always appear as patent activity. Ernst (1998) mentioned that patent activity should have certain connections with R&D investment, and if the R&D investment information is not available, the number of employees in R&D can replace it. Basically, an enterprise’s R&D activity is proportional to the size of its R&D manpower. In view of the above reasons, we

Table 1
The indicators included in the indexes of patent activity and patent quality.

Indexes		Indicators
Patent activity (PA)		Number of patents granted (N_{PG})
		Number of R&D manpower (N_{MA})
Patent quality (PQ)	PQ_1	Total citation ratio (CR_T)
		Number of patent references (N_{PR})
		Number of IPC patent classifications (N_{PC})
	PQ_2	External citation ratio (CR_E)
		Number of patent references (N_{PR})
		Number of IPC patent classifications (N_{PC})

integrate the number of patents granted and the number of R&D manpower into the index for patent activity. Other than the indicators previously proposed by Ernst (1998), this study proposes several new indicators related to patent quality and patent activity.

In Table 1, CR_T and CR_E stand for the total citation rate (sum of self-citations and external citation ratio) and the external citation ratio, respectively. Considering that the number of citations received by a patent in subsequent patent documents is frequently viewed as a sign of an economically important invention (Albert et al., 1991), citations are treated as one of the influential indicators for assessing patent quality. The citation ratio measures the number of patent citations over the number of patent applications. The values of CR_T and CR_E for enterprise i are calculated as follows:

$$CR_{Ti} = 100\tanh\left[\ln\left(\frac{\sum_j TC_{ij}}{\sum_j P_{ij}} \Big/ \frac{\sum_i\sum_j TC_{ij}}{\sum_i\sum_j P_{ij}}\right)\right] \tag{1}$$

$$CR_{Ei} = 100\tanh\left[\ln\left(\frac{\sum_j EC_{ij}}{\sum_j P_{ij}} \Big/ \frac{\sum_i\sum_j EC_{ij}}{\sum_i\sum_j P_{ij}}\right)\right] \tag{2}$$

where CR_{Ti} is the citation ratio for enterprise i, which measures TC_{ij} (the total number of patent citations) over P_{ij} (the number of patent applications). CR_{Ei} is no more than the RCI (relative citation index), where EC_{ij} is the external citation number of the jth class code of technologies of the ith assignee (enterprise).

The values of the related indicators of patent activity and patent quality in the samples are shown in Table 2. This study conducts patent analysis based on the data of 20 firms. However, as the data of the samples would have overcrowded the display of some of the following figures, only the results of the top 10 firms, which account for almost 71% of the patent numbers of the 20 firms, are presented in Table 2. The new index of patent activity is obtained by combining the indicators PG_i and N_{MA} through TOPSIS. The new index of patent quality PQ_1 is acquired from combining CR_T,N_{PC}, and N_{PR}. The new index of patent quality PQ_2 is acquired from combining CR_E, N_{PC}, and N_{PR}.

Other than the indicators proposed by Ernst (1998), we propose new indicators and integrate them by TOPSIS. TOPSIS has worked well in solving the multi-response problem in the Taguchi method. The Taguchi method is an efficient method for optimizing a single quality response; however, most products/processes have more than one quality response of interest. TOPSIS has been applied by several scholars to convert a multi-response problem into a single-response one (Liao, 2003). Similarly, based on the same concept, we use TOPSIS to convert multi-indicators into a single index.

Other methods have been used to solve the multi-response problem in the Taguchi method. For example, Al-Refaie and Al-Tahat (2011) solved the multi-response problem in the Taguchi method by utilizing data envelopment analysis (DEA). In their study, each experiment in Taguchi's orthogonal array was treated as a decision-making unit (DMU) with inputs and/or outputs of multiple response sets. However, DEA is not suitable for this study since no input/output data are available.

The biggest difference between TOPSIS and other principle-based decision-making methods is that TOPSIS considers both ideal and non-ideal solutions. A positive ideal solution is the sum of the optimal solutions that maximize the benefits and minimize the costs of each attribute; a negative ideal solution is the sum of the solutions that are farthest from the ideal solution. We rank the order of the alternatives by relative closeness, which is a measure of the relative distance between the alternative and the positive ideal and negative ideal solutions. In this study, the indicators correspond to the attributes, the firms correspond to the alternatives, and the relative closeness is the proxy of the combination of indicators. The bigger the

Table 2
Measures of indicators for evaluating patent portfolio at the company level.

	PA		PQ_1			
				PQ_2		
Firm	N_{PG}	N_{MA}	CR_T	N_{PC}	N_{PR}	CR_E
Amazon	479	673	722	648	19481	4
Apple	556	736	507	682	24924	34
Cisco	2334	3091	2967	3730	55298	45
Google	794	1129	840	1024	26458	99
Microsoft	2611	4505	2889	3383	49976	18
Novell	153	227	162	238	3077	0
Oracle	1266	1921	1542	1778	40938	156
Red Hat	235	147	208	283	6066	0
Verizon	440	640	399	624	14307	7
VMware	170	195	194	250	1953	0
Total	9388	13264	10430	12640	242478	363

Table 3
The values of new indexes created by TOPSIS.

Index / Firm	PA	PQ_1	PQ_1
Amazon	0.1334	0.2189	0.1615
Apple	0.1383	0.2415	0.2563
Cisco	0.7684	1.0000	0.6033
Google	0.2458	0.3081	0.4777
Microsoft	0.9702	0.9184	0.5152
Novell	0.0141	0.0113	0.0100
Oracle	0.4409	0.5399	0.7119
Red Hat	0.0218	0.0428	0.0369
Verizon	0.1150	0.1475	0.1265
VMware	0.0084	0.0070	0.0018

relative closeness, the better the patent activity or patent quality, and vice versa. This concept is similar to that used to solve the multi-response problem in the Taguchi method. Besides the aforementioned reasons, the original reason behind this study's choice of TOPSIS as the tool to unify indicators is that TOPSIS yields several advantages, including (a) a simple, rationally comprehensible concept, (b) good computational efficiency, and (c) the ability to measure the relative performance of each alternative in a simple mathematical form (Yeh, 2002).

Based on the following TOPSIS steps, the order of the preferred solutions can be obtained, and the relative closeness C_i^* is used as the new index.

Step 1: Construct a normalized evaluation matrix.

Step 2: Determine the positive and the negative ideal solutions.

Step 3: Calculate the separation measures. The distance between solution i and the positive ideal solution is denoted as the degree of separation S_i^+, and the distance between solution i and the negative ideal solution is denoted as the degree of separation S_i^-.

Step 4: Calculate the relative closeness of the activity/quality attributes to the ideal solution by$C_i^* = S_i^-/(S_i^+ + S_i^-)$. The index of C_i^*, ranged $0 \leq C_i^* \leq 1$, now acts as a proxy for patent activity/quality.

Before using TOPSIS to calculate the new index of patent activity PA, we transform the measure of the number of patents granted by dividing the number of patents granted with the average number of patents granted throughout all targeted enterprises. This formulation is shown as follows:

$$PG_i = \sum_j N_{PG_{ij}} / \sum_i \sum_j N_{PG_{ij}} \quad (3)$$

where the numerator is the number of patents granted of the ith enterprise in the jth IPC class code, and the denominator is the total number of patents granted throughout all enterprises.

The values of the new indexes of PA, PQ_1, and PQ_2 for each firm are shown in Table 3.

The classification of the patenting strategies of the samples is shown by the values of the activity and quality dimensions on the abscissa and the ordinate. The patenting behavior of the samples can be categorized into four quadrants, as shown in Figs. 3 and 4 for PA vs. PQ_1 and PA vs. PQ_2, respectively.

In general, an enterprise with a higher citation ratio (located within the first and second quadrants) implies that it owns many basic or leading technologies. According to Ernst's study (1998), for enterprises located in the upper, right hand

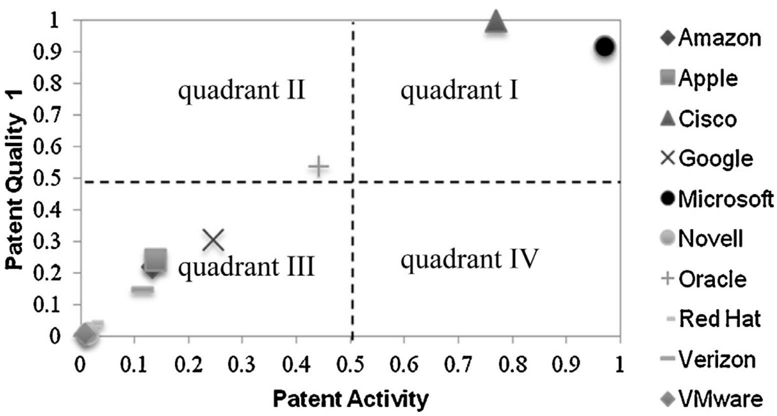

Fig. 3. Identification of patenting strategies with PQ_1 as the index of patent quality.

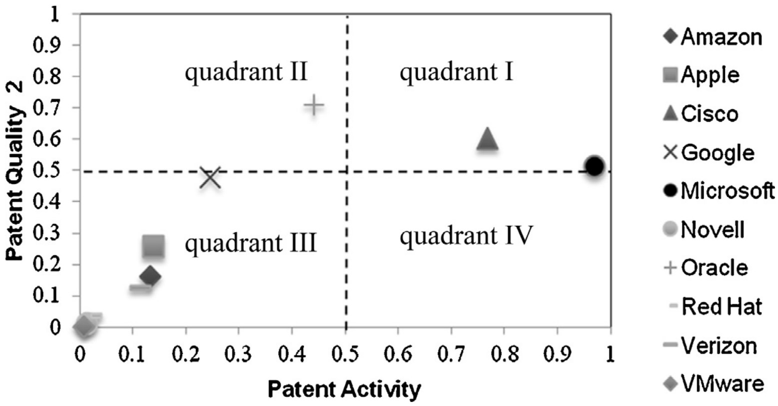

Fig. 4. Identification of patenting strategies with PQ_2 as the index of patent quality.

quadrant of Fig. 3 can be considered as the technological leaders; hence, Microsoft and Cisco appear to be the technological driving forces within this industry. Whereas, for the enterprises positioned in the lower quadrants with a low R&D productivity, it is advised that they should reconsider their R&D activities. An enterprise with a higher self-citation rate, as measured by PQ_1, means that it acts on its own in regards to its R&D activity, and its technologies focus on few specific fields.

On the other hand, an enterprise with a higher number of external citations, as measured by PQ_2, means that its technological quality is high and its application field is wide. PQ_1 and PQ_2 are both combinations of three items, with the difference being the citation ratio, which causes the variation between Fig. 3 and Fig. 4. By comparing these figures, the influence of self-citation ratio on the patenting strategies in patent portfolios can be revealed.

As shown in Fig. 3, the patent quality is proportional to the patent activity. This may be because higher patent activity (implying more patents granted) may induce an increased number of self-citations, and hence produce a higher value of PQ_1 (the sum of self-citations and the number of external citation). On the other hand, as shown in Fig. 4, for PQ_2, the externally cited patent quality derives from the citations by external firms, and the number of citations does not depend on the firm's own patent activity. Therefore, patent quality and patent activity are not correlated.

The top 10 firms shown in Figs. 3 and 4 can be roughly divided into three groups according to their performances in patent activity and patent quality. The first group consists of Microsoft and Cisco; the second group includes Oracle and Google; and the remaining six firms form the third group. Clearly, the first group in Fig. 3 (Microsoft and Cisco) is the most active among the three groups. The second group (Oracle and Google) has less of a patent number than that of the first group and its performance in patent quality PQ_1 is fair. The patent number and patent quality of the third group are both relatively low.

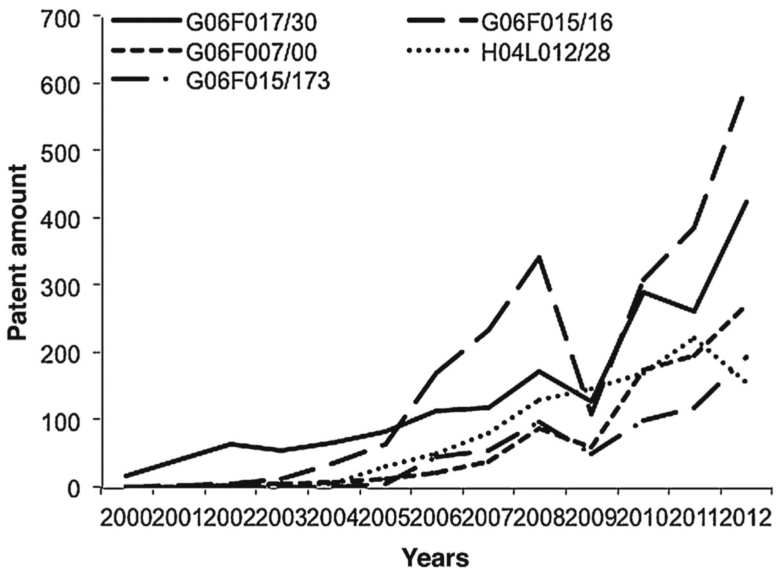

Fig. 5. The technological development trends of the top five patent class codes.

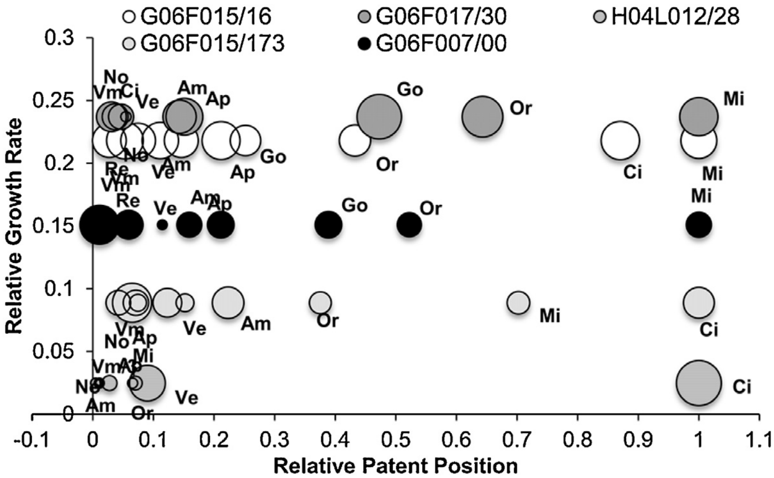

Fig. 6. Patent portfolios at the technological level using RGR as a measure of technological attractiveness.

By comparing the performance changes in patent quality between Fig. 3 and Fig. 4, one may find that the patent qualities of group one, as shown in Fig. 4, are no longer conspicuous. The variation between Fig. 3 and Fig. 4 is mainly attributed to the citation ratio difference between PQ_1 and PQ_2. The reason that the patent quality performance of Microsoft and Cisco drop comparatively in Fig. 4, as compared to Fig. 3, is because the number of external citation is not included. This can be interpreted as affirming that many patents of Microsoft and Cisco are self-cited. In a way, this result shows the patentees' technological independence. We further verify this by the results of MDS that are shown in Fig. 11 in Section 5.

On the other hand, although the patent numbers of group two do not file as many patents as group one, and their patent quality PQ_1 also falls behind, their performance in patent quality PQ_2 are outstanding. Therefore, the technological potential of Oracle and Google ought not to be underestimated. These companies are potential technological competitors against those in group one.

4.2. Analysis at the technological level

Ernst (1998) argued that patent portfolio at the company level contains useful information for the evaluation of overall R&D strategies, but fails to provide information about the technological strengths of specific technological fields. Due to the limited space for illustrating the results (especially in the following Figs. 5–), only five major IPC patent class codes are selected to demonstrate the following analysis.

As depicted in Fig. 5, cloud patents began to accumulate around year 2000 and had a trend of continuous growth, with the exception of the period between 2008 and 2009. This transient phenomenon of decline may be regarded as the outcome of the global financial crisis.

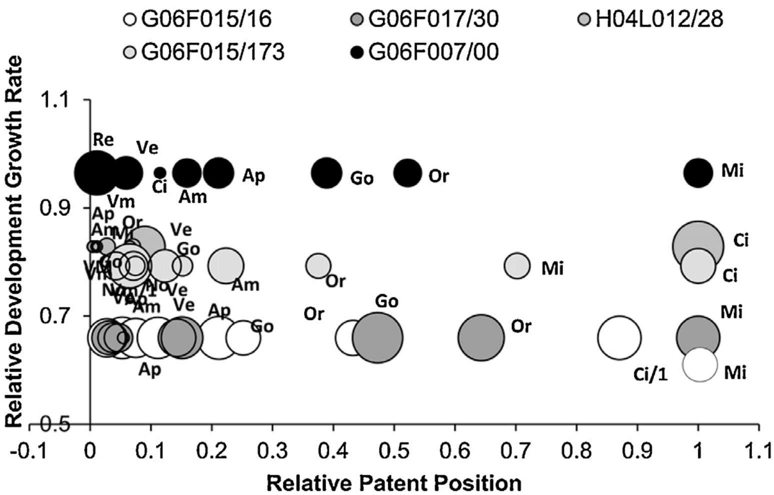

Fig. 7. Patent portfolio using RDGR as a measure of technological attractiveness.

Table 4
The technological contents of the five major patent class codes.

IPC class code	Description of technology
G06F015/16	Combinations of two or more digital computers, each having at least an arithmetic unit, a program unit, and a register, e.g., for a simultaneous processing of several programs.
G06F015/173	Using an interconnection network, e.g., matrix, shuffle, pyramid, star, or snowflake.
G06F017/30	Information retrieval, database structures therefor.
G06F007/00	Methods or arrangements for processing data by operating upon the order or the content of the data handled.
H04L012/28	Characterized by path configuration, e.g., LAN or WAN.

Table 5
The patent numbers and percentages of the top five patent class codes of the top 10 firms.

Firms	Am.	Ap.	Cis.	Go.	Mi.	No.	Or.	Re.	Ve.	VM.
Total patent number of five major patent codes	256	269	1616	556	1569	99	755	158	240	68
Percentage (%)	53	48	69	70	60	65	60	67	55	40

The technological contents of five major codes are listed in Table 4. The total patent numbers of the five major patent class codes owned by the top 10 firms are illustrated in Table 5.

The ratio of the five major patent classes to all related patents of the company is also provided. Although these five class codes are only a quarter of the 20 patent class codes in Table A1, they take up on average about 60% of the total patent number. Among these firms, the two highest percentages are 69% for Cisco and 70% for Google. The data show that most of the patent class codes of Cisco and Google focus on the major technological fields of cloud computing.

It is worth noting that, as illustrated in Fig. 5, the analysis of the five major IPC class codes by annual activities shows that code G06F015/16 includes the highest number of inventions. This code is mainly related to virtualization technology. Based on the study elucidated by Ju et al. (2012), the technological development trend for virtualization are concentrated on the issues of virtual addressing within memory. The life cycle of virtualization technology is in a continually upward developing trend, and the USA has the competitive advantages of virtualization at the international level.

Huang (2016) conducted patent network analysis of cloud computing by using the frequency of keywords' occurrence in patent documents as the input base. With the aid of visual networks and technology centrality indexes, he showed the overall patent structure and identified the influential patents. One of the key patents, US08336049, is titled “Virtual machine utility computing method and system.” This patent, owned by VMware Inc. since 2009, is about an analytics engine that receives real-time statistics from a set of virtual machines supporting a line of business (LOB) application. The results of the qualitative analysis study of Huang (2016) provide further supports to our findings.

Patent portfolio at the technological level contains three elements: relative patent position, technological attractiveness, and technological importance. Patent portfolio at the technological level is constructed by adopting relative patent position (RPP) as the index of the x-axis, and relative growth rate (RGR) or relative development growth rate (RDGR) as the index of the y-axis to assess technological attractiveness.

As proposed by Ernst (1998), index RPP of an enterprise in a particular technological field measures the number of patents owned by the enterprise relative to the number of patents of a benchmark competitor, i.e., the most active competitor or the enterprise with the largest number of patents in this technological field. Thus, the maximum value for RPP in each technological field is 1.

In this study, the attractiveness of each technological field on the ordinate is assessed by calculating the growth of patent applications during the past seven years (2007–2013) relative to the growth in the span of 14 years (2000–2013) to stress the recent changes in patent growth. Index RGR is used to evaluate the attractiveness of a specific technology and to produce the ratio of the average growth of patent applications in a technological field to the average growth of total patent applications in all defined technological fields over the entire period of analysis (2000–2013). Index RDGR is employed to evaluate the developing process and potential of a specific technology and to produce the ratio of the average growth of patent applications in the period 2007–2013 to the average growth of patent applications in the preceding years from 2000 to 2006.

The importance of a technological field shows the distribution of patents over different technological fields and can be viewed as an indication of an enterprise's priorities within its total R&D activities (Ernst, 1998). By calculating the share of

Table 6
The technological attractiveness of the five major patent classes.

IPC	*RGR*	*RDGR*
G06F015/16	0.2181	0.6598
G06F017/30	0.2371	0.6599
H04L012/28	0.0248	0.8286
G06F015/173	0.0888	0.7930
G06F007/00	0.1510	0.9647

Table 7
The relative patent positions of the five major IPC patent class codes in the samples.

	G06F015/16	G06F017/30	H04L012/28	G06F015/173	G06F007/00
Amazon	0.1460	0.1511	0.0107	0.223	0.2823
Apple	0.2117	0.1431	0.0274	0.0744	0.1770
Cisco	0.8704	0.0547	1.000	1.000	0.1292
Google	0.2518	0.4727	0.0107	0.1521	0.7799
Microsoft	1.000	1.000	0.0655	0.7023	1.000
Novell	0.0529	0.0305	0.0046	0.0421	0.0574
Oracle	0.4325	0.6431	0.0701	0.3754	0.000
Red Hat	0.0748	0.0370	0.0046	0.0712	0.0957
Verizon	0.1113	0.0466	0.0899	0.1230	0.0478
VMware	0.0274	0.000	0.0000	0.0647	0.0383

patents in a technological field relative to the overall number of patents owned by the enterprise, we determine the values of technological importance in the technological fields of the leading cloud computing enterprises in the USA.

The results of RPP, technological attractiveness, and technological importance are illustrated in Tables 6–8, respectively.

The patent portfolio of the samples is illustrated in Fig. 6. The strongest patentee in each technological field is located on the right hand side of this figure. Again, we use two letters to represent the enterprise; for example, "Mi" denotes Microsoft. The importance of each technological field is displayed as the circle size in the patent portfolio figure.

As shown on the right hand side of the patent portfolio in Fig. 6, Microsoft owns most of the highest RPP, and is clearly the dominant patentee in most of the technological fields. It appears that technological emphasis of this particular company is put on the technological fields of G06F015/16 and G06F017/30. In other words, the ability to develop virtualization, information retrieval and database structures seems to be its core competence. Summarizing the patent position of Microsoft, it is shown that it holds strong patent positions in all technological fields except G06F015/173 and H04L012/28, which are dominated by Cisco.

Cisco plays an important role in the areas of hardware and Internet peripheral equipment, and the technological field of H04L012/28 has always been the firm's strong point. As shown in Fig. 6, Microsoft and Cisco are equally matched in the growth rate of the technological field of G06F015/16. Based on this observation, we can sense the intense competition between them in cloud computing.

As depicted in Fig. 6, it appears that G06F017/30 is the most attractive technological field since its patent application has grown faster than that of any other technological field. However, this interpretation needs to be adjusted by looking at the other growth variable, i.e., RDGR, as illustrated in Fig. 7. The patent growth of the technologies G06F017/30 and G06F015/16 has been higher over the total period; but has become comparatively lower in recent years. This might be viewed as an indication of technological maturity. On the other hand, the patent growth of G06F007/00 has been lower over the total period, whereas has been increasing in recent years, meaning that new developments in this technological field have recently been given highest priority by the sample enterprises.

The current and potential development of these technological fields are constructed by adopting RGR as the index of the x-axis to assess the attractiveness of special technology, and RDGR as the index of the y-axis to assess the recent changes in the growth trend of patent applications. The variation of the growth rate measuring the technology attractiveness of each technological field can be compared more clearly in Fig. 8.

As displayed in Fig. 8, both values of RGR and RDGR for the technological field of G06F007/00 are high, meaning that development in this field is the major technological trend at present and on the horizon. This field is about search engine technologies, and enterprises that have higher patent numbers in this area are Microsoft and Google (Liang and Sheng, 2013). Since this technological field receives high technology attractiveness, new technological developments in this field would have a positive competitive impact. Thus, competitive activities in this technological field have to be closely monitored and further R&D investment is advisable.

Table 8
The technological importance of the five major IPC patent class codes in the samples.

	G06F015/16	G06F017/30	H04L012/28	G06F015/173	G06F007/00
Amazon	0.2046	0.2404	0.0179	0.1765	0.1509
Apple	0.2555	0.1960	0.0396	0.0507	0.0815
Cisco	0.2602	0.0185	0.3579	0.1686	0.0147
Google	0.1645	0.3504	0.0083	0.0560	0.1943
Microsoft	0.2279	0.2586	0.0179	0.0902	0.0869
Novell	0.2397	0.1570	0.0248	0.1074	0.0992
Oracle	0.1735	0.2928	0.0337	0.0849	0.0000
Red Hat	0.2135	0.1198	0.0156	0.1146	0.1042
Verizon	0.2355	0.1120	0.2278	0.1467	0.0386
VMware	0.2027	0.0000	0.000	0.2703	0.1081

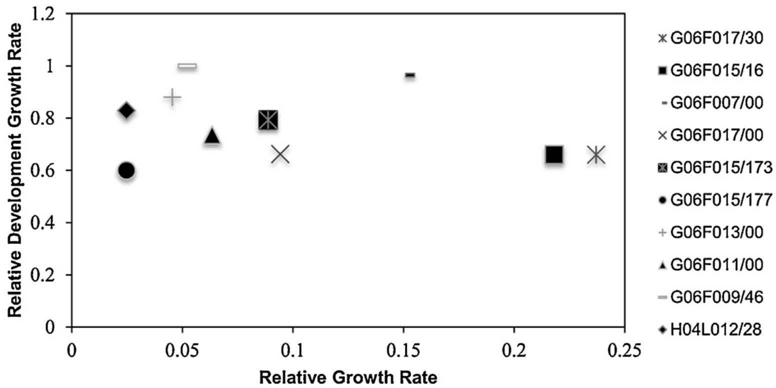

Fig. 8. Distribution of RGR vs. RDGR in 10 technological fields.

On the other hand, the RGR values of fields G06F017/30 and G06F015/16 are high, but those of their potential (RDGR) are low, meaning that either these technological fields tend to mature or their developments have come to a bottleneck.

It is noteworthy that although the RGR value of the technological field of G06F009/46 is under average; however, this field's RDGR value is the highest of all. Technological developments in this field (technology of multiprogramming arrangements) are believed to have the positive competitive impact. Thus, competitive moves in this technological field that improve firm's position have to be closely observed and included in strategic R&D investment decisions. As list in Table A1, enterprises that own higher patent numbers in this technological field are Microsoft, Oracle, Sun, and SAP. On the other hand, both values of RGR and RDGR in the technological field of G06F015/177 are low, which means that this field (initialisation or configuration control) does not have the potential for development; thus, further investment in this technological field may not be justified.

Apart from the patent portfolio analysis at company level and technological level, this study takes the effectiveness of R&D expenditure into consideration to gain more managerial insights. In general, firms that spend more on R&D receive more patents (Ernst, 1998); however, high R&D cost per patent will lead to a competitive disadvantage for the firms. Ernst (1998) mentioned that it is very difficult to get valid R&D figures of competitors because they are either not at all published by companies or, if published, are only available on an aggregate level. Unfortunately, reliable data on the R&D spending in the particular technological field, such as cloud computing, is usually unavailable. Thus, this study calculates the R&D spending per patent of the samples based on the data of the total number of patents acquired via the Patent Guider and the annual R&D expense published in the websites, including CIOZone (http://www.ciozone.com/), wikinvest (http://www.wikinvest.com/) and ychart (http://ycharts.com/).

The R&D spending per patent, i.e., the patent cost, can be regarded as an indicator of the capability for innovative outputs from R&D investment (Palmer et al., 2012). On average, the R&D investment created patent applications with a time-lag of

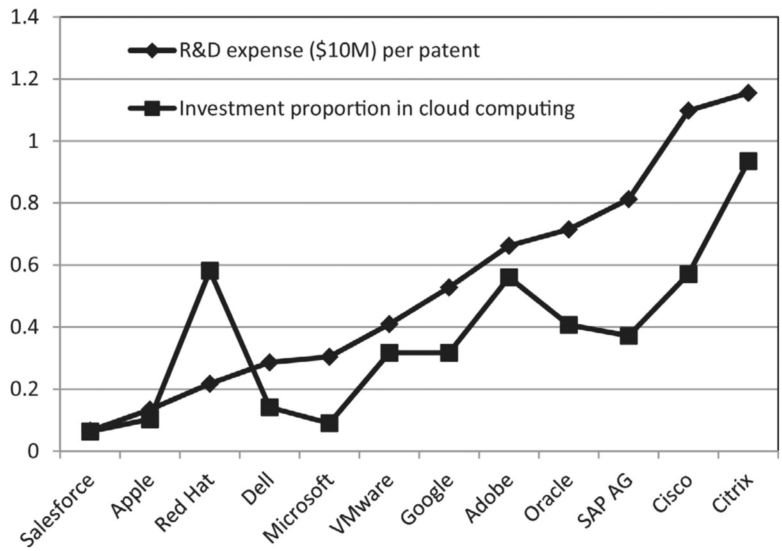

Fig. 9. The R&D spending per patent and the investment proportion in cloud computing of part of the samples.

about a year and a half (Kondo, 1999), and the period between patent application and patent granted takes at least two years (Hall et al., 2005). Therefore, we define the R&D spending (in $10 million) per patent as the averaged R&D spending between year 2007 and 2010 divided by the averaged patent granted between year 2010 and 2013. The R&D spending per patent of part of the samples are shown in Fig. 9. Clearly, the patent cost of Salesforce, Apple, and Red Hat are much lower than that of Cisco and SAP.

The above patent costs are calculated on the basis of the total R&D expenditure and the total patents. To investigate the R&D effect of each company on cloud computing, this study defines the investment proportion in cloud computing for each company as the number of cloud patents listed in Table A1 divided by the total patent granted between year 2000 and 2013. For example, the value of the investment proportion in cloud computing of Citrix is 0.93, which is the proportion of its cloud patents number (201) obtained from Table A1 to its total patents number (215) derived via Patent Guider. As illustrated in Fig. 9, Red Hat, Cisco, and Citrix are companies that have high investment proportion in cloud computing.

It is interesting to see that the measure of the patent cost and the investment proportion in cloud computing show a rough positive correlation. It appears that a company tends to have a higher patent cost if it devotes more R&D spending to cloud computing; so it seems that cloud computing is an expensive investment. It is noteworthy that Microsoft's investment proportion in cloud computing is surprisingly low (0.09). Although, Microsoft owns lots of cloud patents and its patent position in cloud computing is high (as shown in Figs. 6 and 7), the low investment proportion in cloud computing may have adverse effect on its development of cloud computing. In 2011, Microsoft claims to invest 90% of its R&D budget on cloud computing (Eksin, 2011). It seems that Microsoft has noticed the issue of low investment proportion in cloud and has started to adjust its R&D policy.

Among the samples, Red Hat has a high investment proportion in cloud computing but a low patent cost, which means that Red Hat invests most of its resource on cloud computing and it runs efficiently. On the other hand, although Citrix and Cisco have high investment proportion in cloud computing, their patent costs are way above average. It is suggested that they should manage to reduce the R&D cost.

5. The analysis of competition and technological developments

Although Ernst (1998) argued that patent portfolio analysis at the company level is a valuable tool for the R&D decision makers in a company, the information provided from this method seems too simple. In contrast, patent portfolio analysis at the technological level can provide information on the growth of each specific technology and show the technological gaps among firms; however, this type of information would be too fragmented to reveal the overall technological strength of competitors. In short, information on the overall relations among all firms and their technologies is lacking. Therefore, this study employs the methods of FA, MDS, and GRA to provide supplementary information for the R&D decision maker.

The FA method is a method of data reduction. It is used to describe variability among observed, correlated variables in terms of a lower number of latent variables (factors). By using FA to integrate all of the IPC patent class codes owned by the targeted enterprises, we are able to seek the underlying mainstream technologies of cloud computing. MDS is a multivariate technique for revealing the structure of a data set by plotting points in a low-dimensional space. By treating each firm as a combination of patent class codes, as shown in Table A1, we employ MDS to show the differences in patent portfolios arising from the different technical components among firms. This study further uses GRA to verify the results of MDS and provide quantitative data on patent performance and the correlation of firms, as illustrated in the perceptual map.

Table 9
The factor loadings of three mainstream technologies.

IPC class code	Mainstream technology 1	Mainstream technology 2	Mainstream technology 3
G06F009/46	0.937	−0.425	−0.233
G06F017/00	0.904	−0.512	0.12
G06F017/30	0.903	−0.529	−0.034
G06F021/00	0.870	−0.139	−0.085
G06F013/00	0.866	−0.206	−0.207
G06F009/455	0.808	−0.405	−0.488
G06F015/16	0.699	0.154	0.033
G06F015/177	0.686	0.202	−0.089
G06F003/048	0.665	−0.486	0.083
H04J001/16	0.469	0.984	0.142
H04L012/28	0.441	0.9x72	0.118
H04L012/66	0.410	0.941	0.128
H04M003/42	0.659	0.910	0.159
G06F015/173	0.396	0.832	0.085
G06F007/04	0.284	0.704	−0.106
G06F011/00	0.274	0.680	−0.176
G06F019/00	0.134	−0.023	0.961
G06Q030/00	0.211	−0.373	0.907
G06F007/00	−0.405	0.357	0.650
G06Q010/00	−0.542	0.406	0.622

Revealing the mainstream technology of cloud computing, which contains hundreds of patent classifications, may not be easy since each patent may belong to several IPC classifications. Confronted with the masses of qualitative and quantitative variables, many social scientists have turned toward FA to uncover major social patterns. Likewise, in this study, FA is used to simultaneously manage many variables (IPC class codes) and disentangle complex interrelationships into their major and distinct regularities. In other words, we use this statistical approach to analyze the interrelationships among a large number of IPC class codes and then explain these variables in terms of their common underlying factors (mainstream technologies) with a minimum loss of information.

The factor loadings are the correlation coefficients between the variables (IPC class codes) and the factors. If the value of a loading factor corresponding to a specific IPC class code is high, then it means that this class code is important in its respective mainstream technology. In our case, three factors (i.e., the mainstream technologies) are extracted, as shown in Table 9.

Normally, loadings should be 0.6 or higher to confirm that the independent variables identified a priori are represented by a particular factor. In our study, we discard the IPC class codes with loadings that are less than 0.6 and then name each factor according to its content. As presented in Table 9, the first mainstream technology includes nine components. According to the characteristics of these compositions, we name this mainstream technology “Virtualization and information retrieval,” and name the second mainstream technology “Network system,” which comprises six components. The third mainstream technology includes four components and is named “Commercial data process.”

The purpose of MDS is to transform the hidden structures of data by reducing the variables, to display them in a less dimensional space, and to reveal the similarity among the data points on a perceptual map. Kruskal’s stress is the most commonly used measure for determining a model’s goodness of fit. Stress measure, which indicates the proportion of the variance of the disparities unaccounted for by the MDS model, is minimized when the objects are located in a configuration such that the distances among the objects best match the original distances. Regarding the level of stress to tolerate, the rule of thumb is that anything under 0.1 is excellent and anything over 0.15 is unacceptable. In addition, the R^2 measure is another index of fit, which indicates the proportion of variance of the disparities accounted for by the MDS procedure. By entering the factor scores of the top 20 IPC class codes into the SPSS software package for MDS, we produce the stress measure of 0.032 and the value of R^2 at 0.9978, which show an excellent goodness of fit.

Based on the data of Table A1, perceptual maps of technologies and firms are illustrated in Figs. 10 and 11, respectively. It appears that G06F017/30, G06F015/16, and H04L012/28 are quite different from the other technologies.

By using the results of FA and MDS, the direction of three mainstream technologies can be determined, and the technological performance of each enterprise with regards to the mainstream technologies can be evaluated. By overlapping Figs. 10 and 11, one may have a joint map that shows both technologies and firms, which can help determine how they are related perceptually.

In Fig. 11, there are three vectors pointing in different directions, which mean that these three mainstream technologies are not closely correlated. Firms (points) located in the mainstream direction have higher performance in that technology. One may locate firms on each mainstream by drawing a perpendicular line from the point to each mainstream direction. As

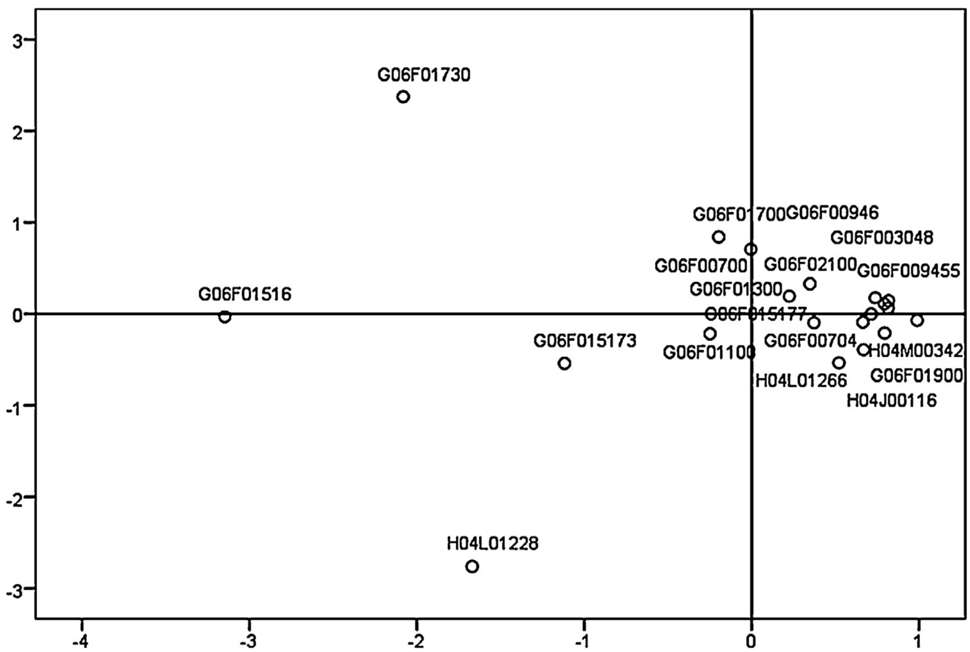

Fig. 10. Perceptual map of technologies.

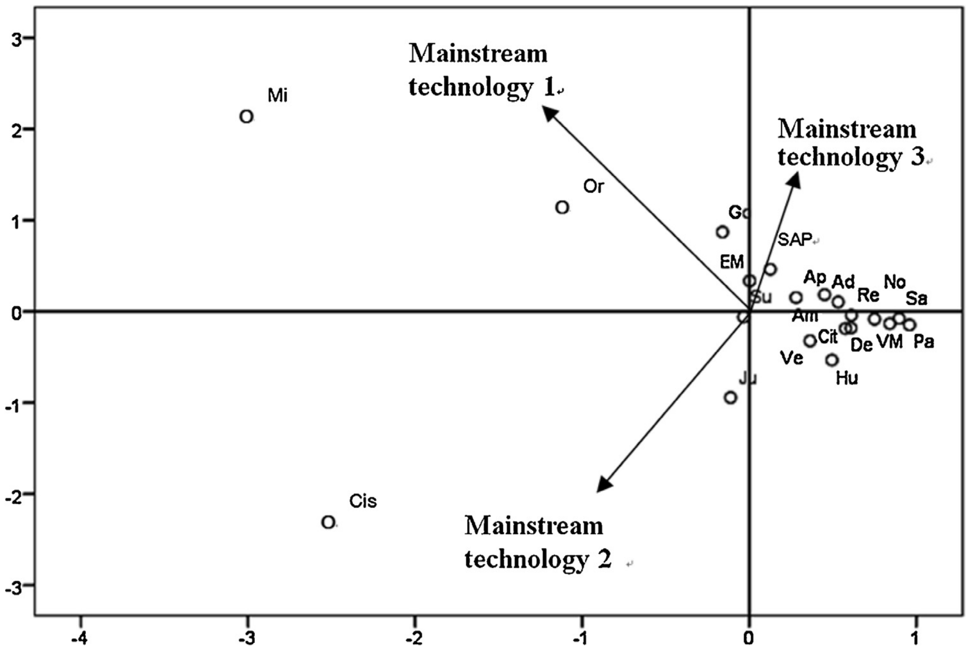

Fig. 11. Perceptual map of firms and the direction of mainstream technologies.

mentioned in Section 4.2, code G06F015/16 is mainly related to virtualization technology. Clearly, Microsoft, Cisco, Oracle, and Google have better performances in the technology of virtualization and information retrieval (mainstream technology 1). Microsoft's application of a large amount of this patent code is evidently related to its Windows Azure Virtual Machines (VMs) service. Ju et al. (2012) used patent analysis to explore the technological developments of virtualization, and found that the major IPC class code consisting virtualization includes G06F015/16, G06F009 and G06F013, etc. Their results are consistent with ours shown in Table 9.

As for the technology of network system (mainstream technology 2), Cisco is in the leading place. This is in line with the market condition, as Cisco has a large market share in the area of network equipment. Regarding the technology of commercial data process (mainstream technology 3), VMware, Google, and SAP are evidently in the dominant position in this area.

The companies with similar patent performances in the same mainstream technology may have a close co-competition relation. Sometimes they may face severe commercial competition with one another, but sometimes they may cooperate. Microsoft and Oracle, which have matched performances in mainstream technology 1, may be taken as an example. When it comes to the importance of multi-tenancy support to their respective public cloud platforms, Microsoft and Oracle have always been on opposite sides (Gong et al., 2010). However, in 2013, after years of competition, Oracle teamed up with Microsoft to bring its database and software to Microsoft's Windows Azure platform (Padhy et al., 2012).

It is worth noting that the results of patent portfolio analysis at the company level (in Fig. 3) show that Microsoft and Cisco are prominent in both patent activity and patent quality; therefore, they are considered to be in severe competition with one another when it comes to innovative products in the cloud computing market. When comparing their technological performances (provided from the patent portfolio analysis at the technological level) one by one, we found that their technological strategies are actually different, and the difference can be easily revealed from the perceptual map of MDS. As shown in Fig. 11, Microsoft and Cisco have technological strategies that are quite different from those of other firms. Moreover, there is an apparent gap between them, which means that even if they are potential competitors in certain technological fields, they are not potential competitors in holistic technological fields, and they each have their own technological development strategies. For example, seeing that the unified communication market has been growing rapidly, Cisco acquired WebEx in 2007; since then, it has become a significant threat to Microsoft (Kaplan, 2007). Two originally irrelevant giants have now reached a direct confrontation in this market battle. Since the business expansion of Cisco, its co-competition relation with Microsoft has grown more and more intense.

Through MDS, the degree of similarity in the R&D strategies and the technological developments of all enterprises can be displayed by transforming the similarities among technologies into distances represented in multidimensional space. However, excluding Microsoft and Cisco, distinguishing the remaining enterprises in Fig. 11 from one another simply via the perceptual map is not easy because they are clustered together. To assist the interpretation of the co-competition relationships among the samples, this study provides another set of quantitative data by using GRA. Grey theory is an effective mathematical means to identify major correlations among the factors of a system with a relatively small amount of data (Li et al., 1997). Currently, the grey theory is widely applied in fields such as economics, agriculture, medicine, geography, seismology, industry, and R&D decision-making (Tzeng and Hu, 1996). In the process of system development, if the degree of synchronous changes between two factors is high, meaning that the change trend is consistent, then the factors will have a higher correlation degree (Liu et al., 2014). As shown in Table A1, each firm can be regarded as a comparability sequence composed of the patent numbers of 20 IPC patent class codes. According to this concept, we may investigate the R&D strategies and technological correlations among firms.

The main procedure of GRA consists of four steps: grey relational generation, reference sequence definition, grey relational coefficient calculation, and grey relational grades (GRGs) calculation. At the grey relational generation step, the performances of alternatives, expressed as $Y_i = (y_{i1}, y_{i2}, \ldots, y_{in})$, are normalized into comparability sequences of $X_i = (x_{i1}, x_{i2}, \ldots, x_{in})$. In this study, y_{ij} stands for the number of each IPC class code j of firm i. The grey relational coefficient $\zeta_i(k)$ can be expressed as follows:

$$\varsigma_{ij} = \frac{\Delta_{\min} + \varsigma \Delta_{\max}}{\Delta_{ij} + \varsigma \Delta_{\max}} \tag{4}$$

where Δ_{ij} is the deviation sequence of the reference sequence (X_0) and the comparability sequence (X_i), i.e., $\Delta_{ij} = ||x_{oj} - x_{ij}||$. The distinguishing coefficient ζ is set to 0.3 in this study. $\Delta_{\max}$ is the largest value of Δ_{ij}, and $\Delta_{\min}$ is the smallest value of Δ_{ij}. The GRG $\Gamma(x_{0j}, x_{ij})$ is computed by averaging the grey relational coefficient corresponding to each quality characteristic (Huang and Liao, 2012), and is defined as follows:

$$\Gamma_j(x_{0j}, x_{ij}) = \frac{1}{n}\sum_{i}^{m} \varsigma_{ij}, \text{ for } j = 1, 2, \ldots, n \tag{5}$$

The GRG indicates the degree of similarity between the comparability sequence and the reference sequence. If we define the reference sequence consists of the maximum out of each of the patent class codes of all firms, we may rank the patent performance of sequences (firms) according to the value of GRG.

Unsurprisingly, as shown in Table 10, Microsoft is ranked as number one. Moreover, one may find that the GRGs of Microsoft and Cisco have a significant gap with those of others, which is consistent with the results shown in Fig. 11.

An analysis of each enterprise's patent portfolios can give a competitive landscape of the cloud computing technological innovations, which is very important for future research planning. Enterprises with similar patent portfolios can be interpreted as having similar innovation activities, including observation of technological progress trends and identification of focused research areas for a new R&D strategy, competition status, and direction for a new potential market development, etc.

Though the similarity of patent portfolios among enterprises can be observed in a perceptual map through MDS, the degree of similarity among the enterprises' patent portfolios can be quantified and investigated with the help of GRA. By selecting a specific firm as the reference sequence, one may determine the firms with the most similar patent portfolios as that of the specific firm. For example, by using Google's IPC sequence as the reference sequence, the first two firms with the highest GRG values are Adobe (0.7182) and Apple (0.6956), which implies that their patent portfolios are the most similar to that of Google. On the other hand, when using Cisco's IPC sequence as the reference sequence, the GRGs of most of the enterprises are low. Among these, the GRG of Microsoft is the lowest (0.4334). This means that although the patent performance of both Microsoft and Cisco are high, their patent portfolios are quite different. Similarly, if we use Microsoft's IPC sequence as the reference sequence, we find that it is Oracle that owns the highest GRG (0.5051), and the GRG of Cisco is low (0.4379). These results are consistent with that shown in Fig. 11.

The analysis results of GRA can reflect the co-competition situations among firms in the market. For example, in some cases a few years ago, Apple and Google created technologies that could have benefitted from one another; however, nowadays, the two giants are in direct confrontation with one another in the market battle of cloud computing. Google is occupying Apple's iOS territory with its Android smart phones; its Chromebook is also directly competing against Apple's MacBook. Nevertheless, more often than not, the two companies have worked together to gain higher profits than their competitors in other areas. On the other hand, a similar conflict between Google and Adobe is occurring. Competing against

Table 10
The GRGs of the targeted enterprises.

Rank	Enterprise	GRG	Rank	Enterprise	GRG
1	Microsoft	0.7432	11	Juniper	0.2740
2	Cisco	0.5850	12	Verizon	0.2738
3	SAP	0.3871	13	DELL	0.2691
4	Oracle	0.3761	14	Huawei	0.2671
5	Apple	0.3238	15	Adobe	0.2635
6	Amazon	0.3171	16	Red Hat	0.2593
7	EMC	0.3161	17	Citrix	0.2588
8	Google	0.3093	18	Novell	0.2586
9	Sun	0.2880	19	Parallels	0.2547
10	VMware	0.2798	20	Salesforce	0.2519

each other, Google acquired companies Picasa and Writely in 2007, while Adobe acquired Omniture in 2009, an online marketing and web analytics business unit (Cohen, 2010; Ouyang, 2011).

The annual R&D expenditure of Google is generally higher than that of Apple; for example, in 2010, the annual R&D expenditure of Google and Apple is $B3.76 and $B1.96, respectively. As shown in Fig. 9, the investment proportion in cloud computing of Google is higher than that of Apple. Also, Google's patent quality is higher than Apple, as illustrated in Fig. 3 and 4. Moreover, as shown in Fig. 7, no matter we compare the most attractive technological field (G06F017/30) or the technological field with the most potential for the future development in cloud computing (G06F007/00), Google owns a higher patent position than Apple. Based on these data, Google seems to have a more favorable competitive position than Apple; however, this interpretation may be adjusted if we take the R&D expenditure per patent into consideration. As shown in Fig. 9, the patent cost of Google is much higher than that of Apple, and which bring light to the competitive weakness of Google. In order to reduce the negative impact of this factor on competition, Google should consider reducing patent costs as one of the key improvement plans in the future.

6. Conclusion

In general, patent portfolio analysis is a useful tool for the R&D decision maker; however, it also has some deficits. For example, information provided by patent portfolio analysis at the company level appears to be insufficient for comparing the R&D strategy planning between firms. Similarly, although patent portfolio analysis at the technical level can provide information on the growth of each specific technology and show the technological gaps among firms, the information yielded is too fragmented to reveal the difference of the overall technological strength among competitors. Considering the inadequacies of these methods, this study has proposed a hybrid patent portfolio scheme by combining the traditional patent portfolio analysis with a multivariate method (FA and MDS) and a multi-attribute decision-making method (GRA), in order to facilitate the findings of firm level strategy and technological trends of the cloud computing industry.

The major contributions of this study are five-fold. First, this study has proposed a compound policy to retrieve cloud computing patents. This database has not only been essential for the present study, but may function as the basic data for future research in this area. Obviously, our data presented much detailed IPC information about cloud computing than the studies which simply retrieved the cloud patents by the class codes of G06 and H04.

Second, as it is inappropriate to determine patent quality by adding indicators directly, this study has made improvements by adopting TOPSIS to integrate the relevant indicators of patent quality into new indexes of patent activity and quality.

Third, the technological trends of cloud computing have been revealed. The results of patent portfolio analysis at the technological level have shown that G06F015/16, G06F017/30, G06F015/173, H04L012/28, and G06F007/00 are five technologies in which most of the targeted enterprises have invested their resources. Among these, technologies related to IPC class codes G06F017/30 and G06F015/16 can be regarded as the current core competences for cloud computing enterprises. In the study of Yeboah-Boateng and Cudjoe-Seshie (2013), they argued that the strongest growth area in cloud computing currently appears to lie in infrastructure virtualization. Based on the analysis of RDGR, we have shown that new technological developments related to G06F009/46 are of particular interest to the cloud computing industry. The findings have revealed that technological fields related to G06F009/46 and G06007/00 are in a growing period, and those related to G06F017/30 and G06F015/16 are major current technology trends but in a slowing down growth rates.

Fourth, we have shown that integrating the method of patent portfolio analysis with that of multivariate analysis is necessary for producing clearer and more accurate information on the technological strengths of competitors and the R&D strategies of cloud computing. Traditional patent portfolio analysis at the technical level can reveal the growth rate of each technology, distinguish the ones that are more attractive, and compare the patent performances of each firm; however, it fails to provide an easy method to monitor the overall relations of technological development among firms. With the help of FA and MDS, this study has been able to integrate technologies into mainstream technologies and compare the R&D strategies of firms. The three mainstream technologies (Virtualization and information retrieval, Network system, and Commercial data process) can be regarded as the major development directions of cloud computing in the present and the future.

Fifth, this study has used the multi-attribute decision-making method of GRA to provide quantitative data for interpreting the perceptual relations among the samples that are illustrated by MDS. GRA is suitable for analyzing sequential-type data, and this study has been the first to use this efficient method to compare patent performances and the similarity of the overall development strategies of firms.

This study focuses on a quantitative analysis for statistically evaluating the company strategies and the status of technology-based activity in cloud computing industry. Among the three mainstream technologies revealed, the first mainstream technology is about virtualization and information retrieval. Major IPC class codes related to these technology fields are G06F015/16, G06F009, G06F013, and G06F017/30, etc. In particular, the patent numbers of G06F015/16 and G06F017/30 are high over the past ten years; however, their growth rates have become comparatively lower in recent years. Microsoft, Cisco, Oracle and Google are companies that have better performances in these fields. As for the technology of network system (mainstream technology 2), Cisco is in the leading place. Regarding the technology of commercial data process (mainstream technology 3), VMware, Google, and SAP are evidently in the dominant position in this area. According to their performances in patent activity and patent quality, this study has divided top 10 firms shown in Fig. 3 into three groups. The results of patent portfolio analysis at the company level have shown that, in group one, Microsoft and Cisco are in the lead in both areas of patent quality and patent activity, and a certain degree of co-competition exists between these two enterprises. However, when comparing their technological performances from the perceptual map of MDS, one may found that even if they are potential competitors in certain technological fields, they are not competitors in holistic technological fields, and they each have their own technological development strategies. On the other hand, although the patent numbers of group two (Oracle and Google) do not file as many patents as group one (Microsoft and Cisco), their technological potential ought not to be underestimated, and they are potential technological competitors against those in group one.

On the whole, both Microsoft and Cisco pursue a holistic R&D strategy by carrying out its R&D activity in all of the relevant technological fields in the cloud computing industry, and they own most of the highest relative patent position values; however, each company has its own leading technological fields. Microsoft holds its strong position in G06F015/16 and G06F017/30, and Cisco in H04L012/28 and G06F015/173. Based on the analysis of the attractiveness of technological fields, it is suggested that Microsoft should shift part of the investment from G06F017/30 and G06F015/16 to G06F007/00. Moreover, Microsoft should manage to increase the proportion of its R&D resource to the technological field of cloud computing. As for Cisco, it is advised that it should manage to reduce the R&D cost or it could be unfavorable to the ability of competition.

In this study, the patent performances of several big companies in cloud computing have been assessed from both the output aspects (such as those listed in Table 2) and input aspects (such as R&D costs and R&D efficiency). As the R&D efficiency and innovation are of great concern to the development of firms, and they are gaining more and more attention in recent years, we aim to invest more research on these subjects in our future research.

Appendices

See Table A1.

Table A1
The patent numbers of the top 20 patent class codes from the samples on cloud computing.

Company	1	2	3	4	5	6	7	8	9	10
IPC	Adobe	Amazon	Apple	Cisco	Citrix	Dell	EMC	Google	Huawei	Juniper
G06F003/048	27	9	64	2	2	4	0	19	1	0
G06F007/00	30	49	35	29	6	7	73	140	3	15
G06F007/04	6	0	14	49	7	0	14	11	2	7
G06F009/455	7	1	10	5	6	9	6	4	0	2
G06F009/46	7	9	18	18	5	5	20	7	1	5
G06F011/00	4	6	12	161	10	60	71	11	12	47
G06F013/00	6	8	52	59	1	33	80	18	5	10
G06F015/16	76	66	110	513	95	47	97	118	60	102
G06F015/173	17	57	22	332	23	27	84	40	16	94
G06F015/177	3	9	26	83	6	44	20	4	6	19
G06F017/00	59	31	48	27	0	4	54	85	0	0
G06F017/30	49	78	85	37	9	16	147	252	0	17
G06F019/00	0	2	4	3	0	16	0	0	0	0
G06F021/00	7	5	10	27	0	7	14	9	0	7
G06Q010/00	4	29	3	1	0	7	2	9	1	0
G06Q030/00	0	113	17	8	0	0	0	55	0	0
H04J001/16	3	1	3	95	0	0	2	2	20	23
H04L012/28	0	6	17	705	19	24	20	6	112	246
H04L012/66	1	1	3	128	9	5	2	0	29	24
H04M003/42	1	0	4	52	3	0	2	5	21	0
Total	310	479	556	2334	201	316	708	794	289	616

Company	11	12	13	14	15	16	17	18	19	20
IPC	Microsoft	Novell	Oracle	Parallels	Red Hat	Salesforce	SAP	Sun	Verizon	VMware
G06F003/048	62	0	16	4	0	1	19	3	7	0
G06F007/00	200	16	0	5	29	6	98	28	12	13
G06F007/04	43	14	19	2	0	2	0	8	11	0
G06F009/455	35	8	35	18	6	0	7	9	0	52
G06F009/46	131	7	80	0	9	0	34	65	4	16
G06F011/00	141	10	65	5	25	2	24	122	33	0
G06F013/00	123	4	56	0	0	2	30	58	10	19
G06F015/16	525	38	224	0	59	9	108	146	74	24
G06F015/173	208	17	110	7	32	3	45	52	46	31
G06F015/177	80	4	33	5	10	1	15	36	10	9
G06F017/00	253	0	128	0	9	0	69	12	12	0
G06F017/30	596	25	378	2	33	22	142	73	35	0
G06F019/00	0	0	8	0	0	0	13	5	0	0
G06F021/00	49	4	21	2	16	0	9	5	7	6

Table A1 (*Continued*)

Company	11	12	13	14	15	16	17	18	19	20
IPC	Microsoft	Novell	Oracle	Parallels	Red Hat	Salesforce	SAP	Sun	Verizon	VMware
G06Q010/00	49	0	18	0	0	0	37	4	7	0
G06Q030/00	35	1	12	0	3	1	26	3	6	0
H04J001/16	9	0	7	0	0	0	0	1	21	0
H04L012/28	41	4	43	0	4	0	0	53	72	0
H04L012/66	11	0	13	0	0	0	0	0	68	0
H04M003/42	21	1	0	0	0	0	2	3	4	0
Total	2611	153	1266	51	235	50	679	685	440	170

Acknowledgement

The authors would like to acknowledge the support from the Ministry of Science and Technology in Taiwan under Grant MOST 103-2410-H-167-008-MY2.

References

Al-Refaie, A., Al-Tahat, M.D., 2011. Solving the multi-response problem in Taguchi method by benevolent formulation in DEA. J. Intell. Manuf. 22 (4), 505–521.
Albert, M.B., Avery, D., Narin, F., McAllister, P., 1991. Direct validation of citation counts as indicators of industrially important patents. Res. Policy 20 (3), 251–259.
Allison, J.R., Lemley, M.A., Moore, K.A., Trunkey, R.D., 2003. Valuable patents. Georgetown Law J. 92, 435–479.
Bal, S.N., 2012. Clouds for different services. IJCSI Int. J. Comp. Sci. Issues 9 (4), 273–277.
Brockhoff, K., 1992. Instruments for patent data analyses in business firms. Technovation 12 (1), 41–58.
Cartlidge, J., Clamp, P., 2014. Correcting a financial brokerage model for cloud computing: closing the window of opportunity for commercialization. J. Cloud Comp. Adv. Syst. Appl. 3 (2) doi:http://dx.doi.org/10.1186/2192-113X-3-2.
Chakrabarti, A., 1991. Competition in high technology: analysis of patents of the United States. Japan, the United Kingdom. France, West Germany and Canada. IEEE Trans. Eng. Manage. 38, 78–84.
Chiou, F.C., 2010. Technology Life of Cloud Computing Using Literature Analysis, Master Dissertation. National Dong-hwa University, Hualien, Taiwan.
Cohen, L.M., 2010. Physical assets in the M&A mix: a strategic option. J. Bus. Strategy 31 (6), 28–36.
Chu, H., 2003. Information Representation and Retrieval in the Digital Age. Medford, New Jersey.
Eksin, A., 2011. Microsoft Says to Spend 90% of R&D on Cloud Strategy. Bloomberg Business. April 7, 2011. http://www.bloomberg.com/news/.
Erdogmus, H., 2009. Cloud computing: does Nirvana hide behind the Nebula? IEEE Software 26 (2), 4–6.
Ernst, H., 1995. Patenting strategies in the German mechanical engineering industry and their relationship to company performance. Technovation 15 (4), 225–240.
Ernst, H., 1998. Patent portfolios for strategic R&D planning. J. Eng. Technol. Manage. 15, 279–308.
Ernst, H., 2003. Patent information for strategic technology management. World Patent Inf. 25 (3), 233–242.
Ernst, H., Omland, N., 2011. The patent asset index – a new approach to benchmark patent portfolios. World Patent Inf. 33 (1), 34–41.
Gong, C., Liu, J., Zhang, Q., Chen, H., Gong, Z., 2010. The characteristics of cloud computing. In Parallel Processing Workshops (ICPPW), 39th International Conference on IEEE (September, 2010), pp. 275–279.
Hall, B., Jaffe, A., Trajtenberg, M., 2005. Market value and patent citations. Rand J. Econ. 36, 16–38.
Harhoff D., Scherer F.M., Vopel K., 2003. Citations family size opposition and the value of patent rights, Res. Policy 32(8), 1343–1363.
Henderson, R.M., Jaffe, A., Trajtenberg, M., 1997. University versus corporate patents: a window on the basicness of invention. Econ. Innovation New Technol. 5 (1), 19–50.
Hsieh, C.H., 2013. Patent value assessment and commercialization strategy. Technol. Forecast. Social Change 80, 307–319.
Huang, J.Y., 2016. Patent Network Analysis of Cloud Computing by Text Mining, J. Technol. 31 (2).
Huang, J.Y., Liao, J.C., 2012. A study on the performance evaluation of computer aided picking system. J. Qual. 19 (2), 41–60.
Huang, J.Y., Yang, H.W., Liu, C.L., 2012. Patent portfolio analysis for cloud computing. 10th Managerial Academy conference, Taichung, Taiwan.
Ju, S.P., Lai, M.F., Fan, C.Y., 2012. Using patent analysis to analyze the technological developments of virtualization. Procedia—Social Behav. Sci. 57, 146–154.
Kaplan, J.M., 2007. Saas: Friend or foe? Bus. Commun. Rev. 37 (6), 48.
Kaur, A., Kaur, A., Bir, S., 2012. Major cloud computing threats and their possible solutions, Int. J. Eng. Res. Technol. 1(9), ISSN: 2278-0181, www.ijert.org (open access).
Kondo, M., 1999. R&D dynamics of creating patents in the Japanese industry. Res. Policy 28 (6), 587–600.
Lanjouw, J.O., Schankerman, M.A., 2000. Characteristics of patent litigation: a window on competition. RAND J. Econ. 32 (1), 129–151.
Lerner, J., 1994. The importance of patent scope: an empirical analysis. RAND J. Econ. 25 (2), 319–333.
Li, P., Tan, T.C., Lee, J.Y., 1997. Grey relational analysis of amine statistics inhibition of mild steel corrosion in acids. Corrosion 53, 186–194.
Li, Y., He, C., Fan, X., Huang, X., Cai, Y., 2015. HCloud, a healthcare-oriented cloud system with improved efficiency in biomedical data processing. In: Terzo, O., Mossucca, L. (Eds.), Cloud Computing with E-science Applications. CRC Press/Taylor & Francis, pp. 163.
Liang, X.P., Sheng, X.T., 2013. Technological development situations of search engine based on patentometrics and visualization. Inf. Sci. 31 (3), 117–120.
Liao, H.C., 2003. Using PCR-TOPSIS to optimise Taguchi's multi-response problem. Int. J. Adv. Manuf. Technol. 22 (9–10), 649–655.
Computer, Intelligent Computing and Education Technology. In: Liu, H.C., Sung, W.P., Yao, W. (Eds.), CRC Press.
Madhavaiah, C., Bashir, I., Shafi, S.I., 2012. Defining cloud computing in business perspective: a review of research. Vis. J. Bus. Perspect. 16 (3), 163–173.
Mohr, J., Sengupta, S., Slater, S., 2010. Marketing of High-Technology Products and Innovation. 3/E, Prentice Hall.
Nimkar, A.V., Ghosh, S.K., 2013. Towards full network virtualization in horizontal IaaS federation: security issues. J. Cloud Comp. Adv. Syst. Appl. 2 (19) doi: http://dx.doi.org/10.1186/2192-113X-2-19.
Ouyang, J.R., 2011. Society for Information Technology & Teacher Education International ConferenceUseful Digital Tools for Teachers: Picasa, Audacity and PhotoStory3, 1. , pp. 1178–1181.
Padhy, R.P., Patra, M.R., Satapathy, S.C., 2012. Windows azure paas cloud: an overview. Int. J. Comp. Appl. 2, 109–123.
Palmer, R., Meek, R., Parkinson, L., Meek, H., 2012. CIM Coursebook 06/07 Managing Marketing Performance. Routledge.
Sultan, N., 2010. Cloud computing for education: a new dawn? Int. J. Inf. Manage. 30, 109–116.

Trappey, A.J.C., Trappey, C.V., Wu, C.Y., Feng, C.Y., 2012. Patent Analysis and Intellectual Property Management–Using Information and Knowledge Management Techniques. Hwa-Tai Publishing Co., Taipei, Taiwan in Chinese.
Tseng, Y.H., Lin, C.J., Lin, Y.I., 2007. Text mining techniques for patent analysis. Inf. Process. Manage. 43 (5), 1216–1247.
Tseng, F.M., Hsieh, C.H., Peng, Y.N., Chu, Y.W., 2011. Using patent data to analyze trends and the technological strategies of the amorphous silicon thin-film solar cell industry. Technol. Forecast. Social Change 78 (2), 332–345.
Tzeng, G.H., Hu, Y.C., 1996. The section of bus system operation and service performance indicators—application of grey relation analysis. J. Chin. Fuzzy Syst. Assoc. 2, 73–82.
Velte, A.T., Velte, T.J., Elsenpeter, R., 2009. Cloud Computing: A Practical Approach. McGraw-Hill.
Yeboah-Boateng, E.O., Cudjoe-Seshie, S., 2013. Cloud Computing: The emergence of application service providers (ASPs) in developing economies. Int. J. Emerg. Technol. Adv. Eng. 3 (5), 703–712.
Yeh, C.H., 2002. A problem-based selection of multi-attribute decision-making methods. Int. Trans. Oper. Res. 9 (2), 169–181.