Expert Systems with Applications 37 (2010) 2882–2890

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Discovering competitive intelligence by mining changes in patent trends

Meng-Jung Shih, Duen-Ren Liu *, Ming-Li Hsu

Institute of Information Management, National Chiao Tung University, Hsinchu, Taiwan

ARTICLE INFO

Keywords:
Change mining
Patent analysis
Competitive intelligence
Knowledge discovery

ABSTRACT

Obtaining sufficient competitive intelligence is a critical factor in helping business managers gain and maintain competitive advantages. Patent data is an important source of competitive intelligence that enterprises can use to gain a strategic advantage. Under existing approaches, to detect changes in patent trends, business managers must rely on patent analysts to compare two patent analysis charts of different time periods. The discovery of change of trends currently still needs laborious human efforts and no efficient computer-based approaches are available for helping this task. In this paper, we propose a patent trend change mining (PTCM) approach that can identify changes in patent trends without the need for specialist knowledge. The proposed approach consists of steps including patent collection, patent indicator calculation, and change detection. In change detection phase the approach firstly excavate rules between two different time periods, comparing them to determine the trend changes. These trend changes are then classified into four categories of change, evaluated with change degree and ranked by their change degree as the output information to be referred by decision makers. We apply the PTCM approach to Taiwan's semiconductor industry to discover changes in four types of patent trends: the R&D activities of a company, the R&D activities of the industry, company activities in the industry and industry activities generally. The proposed approach generates competitive intelligence to help managers develop appropriate business strategies.

1. Introduction

For a competitive organization, competence management is critical to organization development and even to survival issue. Complete competence management generally consists by processes including competence identification, assessment, acquisition and knowledge usage (Berio & Harzallah, 2007). But before the four processes of competence management, a more perspective issue is to determine which competence to obtain. To accomplish this task, competitive organizations need to keep tracing the trends of competence change and find potential elements which may substantially improve the organization competitiveness. Unfortunately, most competences—especially competitive intelligence, are neither structured nor quantifiable. So how to effectively discover the trends of change among these abundant unstructured valuable data like intelligent properties, or more precisely say patents, will be very essential to an organization to "lock on" the target competences to obtain. For instance of patent data, they embody technological novelty and serve as important sources of competitive intelligence with which enterprises gain strategic advantages (Stembridge & Corish, 2004). Patents directly represent the competitive intelligence of an industry. Any variation on patent trends in an industry as a whole will directly influence the research and development strategies of all involved enterprises. It emerges when a novel technique developed or when a revolutionary product (or parts) are invented. To maintain a leading position in the highly competitive business environment, enterprise managers need comprehend key intelligence properties of their own organization, of their competitors, and of the environment in which they operate. By analyzing patent data, managers can evaluate and understand trends in the development of technologies and plan suitable strategies (Stembridge, 2005).

There has been a great deal of researches on patent data analysis, and several applications, such as patent map, patent citation analysis, and patent indicators, have been developed (Breitzman & Mogee, 2002; Brockoff, 1991; Chang, 2005; CHI-Research; Dou, Leveillé, Manullang, & Dou, 2005; Dürsteler, 2007; Kim, Suh, & Park, 2008; Reitzig, 2004; Yang, Akers, Klose, & Yang, 2008). Most of these studies and tools use statistical methods to analyze patent data in a specific period, and represent patent trends by visualization graphs and tables. However, these tools fail to express changes in patent trends over two time periods. A patent map visualization method proposed by Kim et al. (2008) overcomes drawbacks of conventional patent maps; it enables user to understand the

* Corresponding author. Tel.: +886 3 5131245; fax: +886 3 5723792.
E-mail address: dliu@iim.nctu.edu.tw (D.-R. Liu).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.09.001

progresses of technologies, but it cannot provide a clear insight into the changes in patent trends for different periods. In real scenario, experts still have to identify changes in patent trends by comparing charts/tables for different periods. This task is laborious and there still has no corresponding automatic tools to help accomplish this work.

Changes in patent trends represent movements in the direction of technology development. For example, suppose a company X has its patents mainly on field A in 2003 and 2004. If the company's main field of patents in 2005 and 2006 has became field B, we can say the technology development direction of company X has changed from A to B. To capture changes in patent trends in different periods, this study proposes an approach which identifies patent trend changes with absence of specialist knowledge. These changes are ranked with change degree which is introduced in this paper. We combine association rule change mining (Song, Kim, & Kim, 2001) with patent indicators (Brockoff, 1991; CHI-Research) to develop a technique called patent trend change mining (PTCM), which transforms patent documents into a rule format and then identifies the most frequent rules. The frequent rules represent a patent trend in a specific period and thus, we can observe changes in patent trends by comparing the frequent rules of two time periods. The patent trends of four different business levels are discussed in this study: one in enterprise scope and three in industrial scope. We analyze each level of changes revealed by the proposed method, and these changes are classified, evaluated and ranked as the output.

The remainder of this paper is organized as follows. In the next section, we review literature relevant to this research, including association rule mining, change mining, patent analysis, and patent indictors. Section 3 provides an overview of our patent trend change mining (PTCM) technique. In Section 4, we describe the methods for mining changes in patent trends in detail. In Section 5, we investigate changes in patent trends in Taiwan's semiconductor industry. Then, in Section 6, we present our conclusions and directions for future research.

2. Background and related work

We begin this section by reviewing the definition of association rule mining used to discover trends in patent documents, and then present an overview of state-of-the-art change mining techniques. The third subsection contains an introduction to patent analysis. Then, in the fourth subsection, we discuss commonly used patent indicators.

2.1. Association rule mining

Data mining techniques have been widely used in various fields of information science (Chang, Lin, & Wang, 2009; Chen & Liu, 2004; Kuo, Lin, & Shih, 2007; Ngai, Xiu, & Chau, 2009; Yen & Lee, 2006). Association rule mining is a data mining technique used in various applications, such as market basket analysis. The technique searches for interesting associations or relationships among items in a large data set (Han & Kamber, 2001). Different association rules express different regularities that exist in a dataset; and two measures, support and confidence, are used to determine whether a mined rule is a regular pattern (Han & Kamber, 2001; Ian & Eibe, 2000). The support measure determines the probability that a transaction contains both the conditional and consequent parts of a rule, while the confidence measure is the conditional probability that a transaction containing the conditional part of a rule also contains the consequent part. The apriori algorithm (Agrawal & Skrikant, 1994) is typically used to find association rules by discovering frequent itemsets (sets of items), which are considered to be frequent if their support exceeds a user-specified minimum support threshold. Association rules that meet a user-specified minimum confidence can then be generated from the frequent itemsets.

In this work, we apply association rule mining to patent data to find patent patterns (rule patterns).

2.2. Change mining

The objective of change mining is to discover changes in two datasets (e.g., about customer behavior) belonging to different time periods. Change mining approaches can be classified as follows:

(a) *Decision Tree Models:* this method constructs decision trees for two datasets, and then identifies the differences by comparing the two decision trees (Liu & Hsu, 1996; Liu, Hsu, Han, & Xia, 2000).
(b) *Association Rules:* this method determines changes by comparing the association rules mined from two datasets (Song et al., 2001; Chen, Chiu & Chang, 2005; Liu, Hsu, & Ma, 2001). Users can decide the type of rule changes according to the similarities and differences between the rules in the datasets. There are several types of change mining patterns (Song et al., 2001; Chen, Chiu & Chang, 2005):
  - *Emerging patterns: The* concept of emerging patterns captures significant changes between datasets. An emerging pattern is a rule pattern whose support increases significantly from one dataset to another.
  - *Unexpected consequent changes: These* changes are found in newly discovered association rules whose consequent parts differ from those of the previous rule patterns.
  - *Unexpected condition changes: These* changes are found in a newly discovered association rules whose conditional parts differ from those of previous rule patterns.
  - *Added rules: These* are new rules that only exist in the present dataset.
  - *Perished rules: These* are rules that only exist in the previous dataset.

Association rule change mining techniques are used to analyze transaction data and discover changes in customer behaviour. In this work, we identify changes in patent trends from patent data.

2.3. Patent analysis

Rapid technological development has made it easier for companies to search and access patent documents. Many patent offices already allow free download of the abstracts and complete texts of their patents [e.g., WIPO (WIPO, 2007), USPTO (USPTO, 2007) and EPO (EPO, 2007)].

Several software tools and services have been developed in the patent field (Breitzman & Mogee, 2002; Dou et al., 2005; Dürsteler, 2007; Huang, Ke, & Yang, 2008). These tools analyze patents by classification, clustering, and statistical methods to find the relationships between patents with similar content/structure. The results of patent analysis are usually presented as graphs or tables, and provided to specialists, researchers, and R&D practitioners to help them plan their strategies.

Patent information can be analyzed either quantitatively or qualitatively (Huang et al., 2003). Quantitative measures are based on statistical processing, and indicate the level of patenting activity of an analytical unit (e.g., the number of patents owned by an assignee). Qualitative measures are calculated according to citation information and used to assess the quality of a patent.

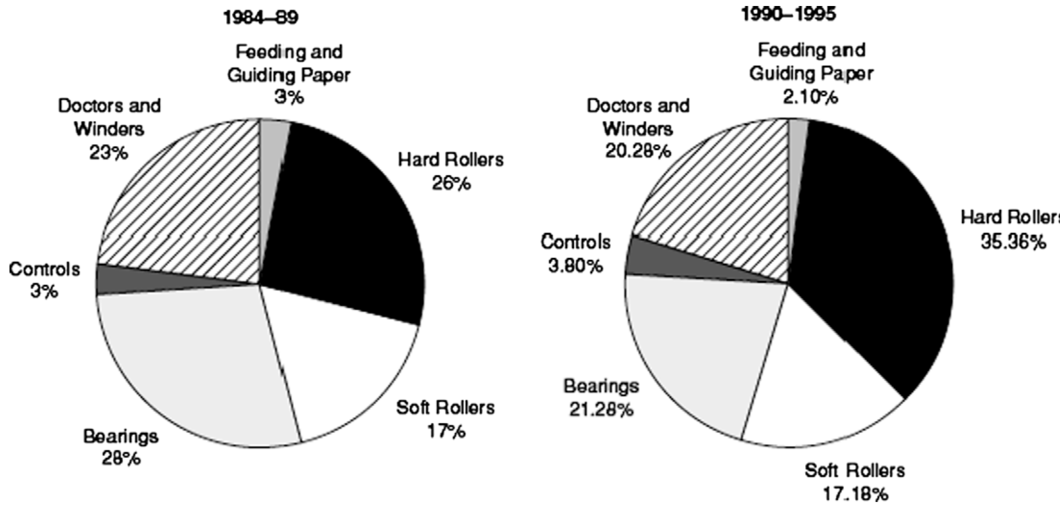

Fig. 1. Distribution of technological fields of paper-making machinery.

In the literature, and in practice, several indicators are used to measure patents quantitatively or qualitatively. In the next subsection, we introduce patent indicators.

Although existing patent analysis tools can provide various results, analysts still need to compare the results of two periods to identify changes over time. For example, Fig. 1 shows the distribution of the technological fields of paper-making machinery in two periods, 1984–1989 and 1990–1995 (Breitzman & Mogee, 2002). Patent analysts can discover changes in the technological field by comparing the two distributions. In this case, R&D activities increased for Hard Rollers and Controls, decreased for Bearings, and remained stable for other areas (Breitzman & Mogee, 2002). Making such comparisons requires professional knowledge. Moreover, changes cannot be ranked intuitively; the degree of change must be calculated and ranked by analysts.

The motivation of this study is to discover changes in the patent trends of different time periods without the need for expert knowledge, and report changes to business managers by ranking the degree of change.

2.4. Patent indicators

Patents are one of the major sources of technological and competitive information because such data are easy to access and the content is highly innovative. Since the value of patents is rarely observable, scholars and research organizations have defined a number of patent indicators to determine the value of patents (Brockhoff, 1991; CHI-Research; Reitzig, 2004; Tuomo, Hermans, & Kulvik, 2007).

The common patent indicators are described below (Brockhoff, 1991; CHI-Research; Reitzig, 2004; Tuomo, Hermans, & Kulvik, 2007):

- *Patent age*: The age of a patent (the patent's age is calculated from the date the patent was applied for).
- *Citation made (backward citations)*: The number of patents cited by the target patent.
- *Citation index (forward citations):* The number of citations received by the target patent. It is a measure of the impact of the target patent.
- *Originality*: The originality of a target patent indicates the diversity of cited patents, i.e., the patents cited by the target patent. The measure is based on the distribution (ratio) of cited patents over classes, as expressed in Eq. (1).

$$\begin{aligned} &\mathit{Originality} = 1 - \sum_{j \in S_B} B_j^2 \\ &B_j = \frac{\text{Number of cited patents belonging to Class j}}{\text{Number of cited patents}} \\ &S_B: \text{ the set of classes of cited patents} \end{aligned} \tag{1}$$

- *Generality*: The generality of a target patent indicates the diversity of citing patents, i.e., the patents that cite the target patent. The measure is based on the distribution (ratio) of citing patents over classes, as expressed in Eq. (2).

$$\begin{aligned} &\mathit{Generality} = 1 - \sum_{j \in S_F} F_j^2 \\ &F_j = \frac{\text{Number of citing patents belonging to Class } j}{\text{Number of citing patents}} \\ &S_F: \text{ the set of classes of citing patents} \end{aligned} \tag{2}$$

- *Technology Cycle Time (TCT):* The TCT of a target patent is the median age of the patents cited by the target patent. It is a measure of technological progress.

3. Methods

The proposed patent trend change mining (PTCM) approach comprises four components, as shown in Fig. 2: a patent fetcher, a patent transformer, a patent indicator calculator, and a change detection module. The first three components are described in this section, and we have more detail discussion on change detection process in Section 4.

3.1. Patent fetcher

With the rapid growth of computer and internet technologies, patent documents can now be accessed freely via the Internet. The patent fetcher module uses a keyword search strategy (e.g., Assignee and International Patent Classification Code, IPC) to retrieve patents for analysis. Patent fetcher acquires patent documents (in HTML format) from the patent website and stores them into the patent document pool.

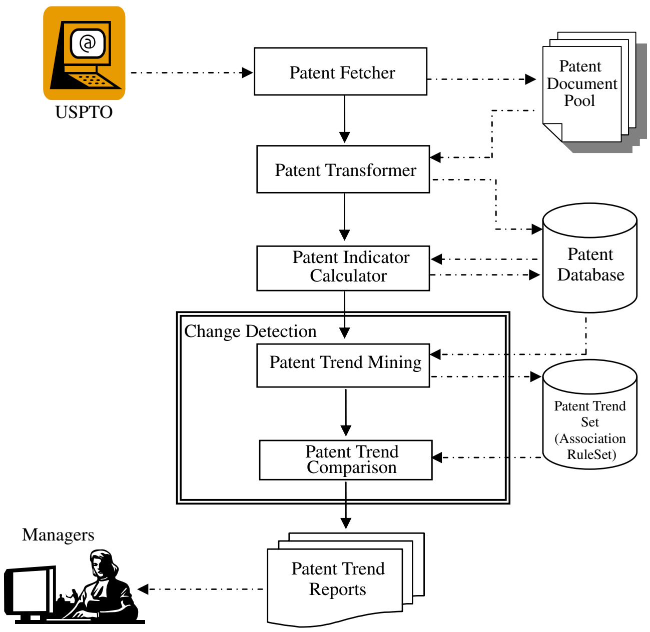

Fig. 2. An overview of the PTCM technique.

3.2. Patent transformer

Initially, a patent document is in a semi-structured HTML format. This module transforms the raw patent document from semi-structured HTML format into a text format, stores it in the database, filters out irrelevant content, and extracts required patent content, including the patent number, International Classification (IPC), Application Date, Assignee Name, and Assignee Country. The extracted content is stored in the database for further processing to compute patent indicators.

3.3. Patent indictor calculator

This module calculates the patent indicators for each patent to determine the patent's value. In this study we use four patent indicators, which are defined in Section 2.4, to analyze patent documents: *Citation Index* (CI) of a patent reflects the technological significance of a patent—the higher the value of a patent's CI, the greater the patent's impact. *Originality* measures the innovation of a patent—the higher the value of a patent's originality, the greater the patent's innovation value. *Generality* measures the scope of cross-field applications on which a patent is applied—the higher the value of a patent's generality, the greater the patent's economic value. A patent is interpreted as having more "generality" if the forward citations are spread over several technological fields. *Technology Cycle Time* (TCT) measures the time between the previous patent and the target patent, which makes improvement on the previous one—shorter TCT means a faster technological progress of patents.

The values of patent indicators are discretized for further patent trend mining. We perform data discretization based on the normalized results derived by SPSS Visual Bander. The values of patent indicators are transformed into linguistic terms as shown in Table 1.

Table 1
Data discretization of patent indicators.

Patent indictor	Linguistic term	Numerical range
CI	Low	⩽0
	Mid	1–4
	High	⩾5
Originality	Low	0–0.39
	Mid	0.40–0.65
	High	0.66–1
Generality	Low	0–0.44
	Mid	0.45–0.65
	High	0.66–1
TCT	Short	0–5
	Mid	6–7
	Long	⩾8

4. Change detection in patent trends

Patents indicate the technological competitiveness as well as the innovation strategy of a company in a given period. Business managers can observe changes in patent trends by comparing the trends of two periods. The process of detecting changes in patent trends is illustrated in Fig. 3.

4.1. Patent trend mining

Before describing the patent trend mining module, we introduce the patent trends analyzed in this study. We define four kinds of patent trends and classify them into two levels for analysis: company-level and industry-level trends.

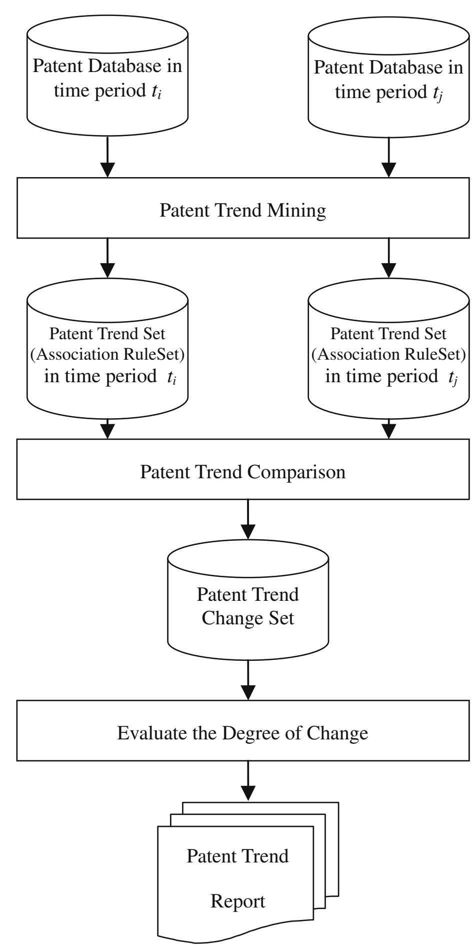

Fig. 3. The process of detecting changes in patent trends.

Table 2
Patent trends and their respective rule formats.

Analyzed level	Patent trend	Rule format		
		Conditional part	→	Consequent part
Company level	R&D activities of a company	IPC	→	CI/Originality/Generality/TCT
Industry level	R&D activities of the specified industry	IPC	→	CI/Originality/Generality/TCT
	Technological competitiveness of companies	Assignee	→	CI/Originality/Generality/TCT
	Technological competitiveness of companies in a specific technological field	Assignee, IPC	→	CI/Originality/Generality/TCT

(a) *Company-level patent trends*: These trends provide information about a company's technological development.

- *Trends in the R&D activities of a company*: Changes in the R&D activities of a company can be determined by comparing the relations between technological fields (IPC) and four patent indicators (the citation index, originality, generality and technology cycle time described in Section 3.3) over two time periods.

(b) *Industry-level patent trends*: These trends provide information about the technological development of an industry.

- *Trends in the R&D activities of an industry*: Changes in the R&D activities of an industry can be determined by comparing the relations between the technological fields (IPC) and four patent indicators over two time periods.
- *Trends in the technological competitiveness of companies*: We identify changes in technology competitiveness of companies by comparing the relations between a patent's assignee (company) and the four patent indicators over two time periods; the patent indicators reflect the technological competitiveness of a company.
- *Trends in the technological competitiveness of companies in a specific technological field*: These changes can be observed by comparing the relations between both a patent's assignee and technological fields (IPC) and four patent indicators over two time periods.

Table 2 shows the four kinds of patent trends and their respective rule formats.

We apply association rule mining to patent data to identify patent trends (frequent association rule patterns). The mined frequent patterns can be regarded as trends extracted from patent documents. For example, if there are sufficient patents belonging to technological field B, whose assignee is X, and the CI value of those patents is high, the frequent association rule pattern "Assignee = X, IPC = B → CI = high" can be identified. The rules identify a patent trend in which the citation index of X's patents in technological field B is relatively high. This information suggests that the quality of X's patents in technological field B is high in the industry. Moreover, we may say that X is a pioneer company in technological field B.

4.2. Patent trend comparison

After the patent trends of different time periods have been discovered, the trends (in rule format) are compared to identify changes. We start with defining the types of change as follows and then discuss the process of trend comparison.

4.2.1. Types of change

Based on previous research (Song et al., 2001), four types of change in patent trends are defined:

(1) *Emerging patent trends:* an emerging patent trend is a rule pattern whose support increases significantly from one dataset to another.

(2) *Unexpected changes in patent trends:* unexpected changes in patent trends can be found in newly discovered patent trends whose consequent parts of the rule patterns are different from those of the previous patent trend.

(3) *Added patent trends: An* added patent trend is a new rule, i.e., a rule not found in previous rule patterns.

(4) *Perished patent trends: A* perished patent trend is the opposite of an added rule, as it is only found in previous rule patents.

4.2.2. Rule matching

We use a rule matching method to compare the patent trends of different time periods. The method computes the similarity measures and difference measures of the patent trends $rule_i^t$ and $rule_j^{t+k}$ in time t and time $t+k$, respectively. The modified rule matching method comprises the following four steps (Liu, Shih, Liau, & Lai 2009; Song et al., 2001).

Step 1. Calculate the similarity degree of the conditional/consequent parts of two rules in different time periods.
Step 2. Calculate the similarity measure S_{ij} between two rules. The measure is derived by multiplying the similarity degree of the conditional parts (C_{ij}) of the rules by the similarity degree of the consequent parts (Q_{ij}).
Step 3. Calculate the difference measure ∂_{ij} between two rules. The measure is the similarity degree of the conditional parts minus the similarity degree of the consequent parts.
Step 4. Determine the type of change according to the similarity measures and difference measures.

4.2.3. Identifying the type of change

Table 3 shows the measures used to determine each type of patent change; the measurements are adopted from (Liu et al. 2009; Song et al., 2001). The four types of patent change can be classified according to the two judged factors, i.e., the similarity measure S_{ij} and the difference measure ∂_{ij}, and three predefined thresholds: θ_{em} for emerging patterns, θ_{un} for unexpected changes, and $\theta_{a/p}$ for added and perished rules. Note that $\theta_{em} > \theta_{un} > \theta_{a/p}$. The process of identifying the types of changes follows a pre-determined sequence. First, we identify emerging patterns. If the similarity

Table 3
Measurement for each type of change.

Type of change (r_i^t, r_j^{t+k})	Measurement
Emerging pattern	$S_{ij} \geqslant \theta_{em}$ ($S_{ij} = C_{ij} \times Q_{ij}$) ($C_{ij}$: similarity degree of the conditional parts) (Q_{ij}: similarity degree of the consequent parts)
Unexpected change	$Max(\varsigma_i, \varsigma_j) < \theta_{em},\ \partial_{ij} > \theta_{un}\ (\partial_{ij} = C_{ij} - Q_{ij})$
Added patent trend	$\varsigma_j < \theta_{a/p}\ (\varsigma_j = \max_i S_{ij})$
Perished patent trend	$\varsigma_i < \theta_{a/p}\ (\varsigma_i = \max_j S_{ij})$

Table 4
Measuring the degree of change in patent trends.

Type of change	Degree of change
Emerging patent trends	$\frac{Support^{t+k}(r_j)-Support^t(r_i)}{Support^t(r_i)}$
Unexpected changes in patent trends	$\frac{Support^t(r_i)-Support^{t+k}(r_j)}{Support^t(r_i)} \times Support^{t+k}(r_j)$
Added patent trend	$(1-\varsigma_j) \times Support^{t+k}(r_j)$
Perished patent trend	$(1-\varsigma_i) \times Support^t(r_i)$

measure S_{ij} is greater than or equal to θ_{em}, it means that the two rules are similar and rule r_j^{t+k} can be regarded as an emerging pattern. If the maximum similarity measure $Max(\varsigma_i, \varsigma_j)$ is less than θ_{em} and the difference measure ∂_{ij} is greater than θ_{un}, we regard rule r_j^{t+k} as an unexpected change. Note that $\varsigma_i = \max_j S_{ij}$; $\varsigma_j = \max_i S_{ij}$. Finally, if ς_j is less than $\theta_{a/p}$, rule r_j^{t+k} is identified as an added patent trend; and if ς_i is less than $\theta_{a/p}$, rule r_i^t is identified as a perished patent trend.

4.3. Evaluating the degree of change

As a large number of changes occur in a competitive business environment, managers need to focus on the most important changes. To do this, it is necessary to evaluate the degree of change, and rank the changed rules according to their importance.

Table 4 shows the simple formulations for measuring the degree of change. The formulations, which are adopted from (Liu et al. 2009), measure the degree of change. The notations $support^t(r_i)$ and $support^{t+k}(r_i)$ represent the support value of r_i at time t and and r_j at time $t+k$, respectively; while ς_i and ς_j are the maximum similarity measures of r_i^t and r_j^{t+k}, respectively.

After calculating the degrees of change, the most important changes are reported to business managers, who then analyze the changes in patent trends over different time periods and use the information to understand the changing business environment and plan appropriate strategies.

5. Patent change mining in Taiwan's semiconductor industry

We now apply our proposed PTCM technique to Taiwan's semiconductor industry.

5.1. Data collection

The dataset of semiconductor-related patents was obtained from the USPTO (United States Patent and Trademark Office) patent database. We select Taiwan semiconductor-related patents available online for the period 2001–2004 based on the IPCs belonging to the semiconductor industry, as identified by the Taiwan Intellectual Property Office (see Appendix A). We divided this dataset, which contains 4310 unique patents, into two periods: the first part contains 2352 patent documents for the period 2001–2002, while the second part contains 1958 patent documents for the period 2003–2004.

5.2. Changes in the R&D activities of TSMC (Taiwan Semiconductor Manufacturing Co. Ltd)

Changes in a company's R&D activities are identified by comparing the relations between the technological field (IPC) of the target company and the citation index, originality, generality, and technology cycle time over two time periods. We chose TSMC as the target company, and divided its patents into two parts: 2001–2002 and 2003–2004. Table 5 lists some changes in the R&D activities of TSMC between 2001 and 2004.

From patent trend (1), we observe the rapid growth (57%) of the company in terms of high originality in H01L29/788. This information shows that, during the period under study, TSMC exhibited a high degree of inventiveness in the technological field H01L29/788.

Meanwhile, patent trend (3) shows that the citation index of H01L27/108 decreased between 2001 and 2004. A reduction in the CI often indicates a decline in quality, although it can mean that the patent is fairly new. The added patent trends (5) and (6) in Table 6 indicate that H01L21/336 and G01R31/26 are new technological fields that TSMC invested in. The number of citations of these patents is relatively low. Finally, from perished patent trends (7) and (8), we observe that the innovativeness of TSMC declined gradually in terms of H01L21/336 and H01L21/44 in the period under study.

5.3. R&D activities of Taiwan's semiconductor industry

Changes in the R&D activities of an industry are identified by comparing the relations between the technological fields (IPC) of the target industry and the citation index, originality, generality and technology cycle time over two time periods. Table 6 lists

Table 5
Some changes in the R&D activities of TSMC.

Patent trend		Change degree
Emerging patent trends		
(1) IPC = H01L29/788 → Originality = High		0.57
(2) IPC = H01L21/00 → TCT = Short		0.21
Unexpected changes in patent trends		
2001–2002	**2003–2004**	
(3) IPC = H01L27/108 → CI = Mid	IPC = H01L27/108 → CI = Low	0.02
(4) IPC = H01L21/311 → TCT = Short	IPC = H01L21/311 → TCT = Long	0.02
Added patent trends		
(5) IPC = H01L23/62 → CI = Low		0.03
(6) IPC = G01R31/26 → CI = Low		0.02
Perished patent trends		
(7) IPC = H01L21/336 → CI = High		0.05
(8) IPC = H01L21/44 → Generality = High		0.03

Table 6
Some changes in the R&D activities of Taiwan's semiconductor industry.

Patent trend		Change degree
Emerging patent trends		
(1) IPC = H01L29/76 → CI = Low		1.31
(2) IPC = H01L21/00 → CI = Low		1.07
Unexpected changes of patent trends		
2001–2002	**2003–2004**	
(3) IPC = H01L29/40 → TCT = Short	IPC = H01L29/40 → TCT = Mid	0.02
(4) IPC = H01L21/48 → TCT = Short	IPC = H01L21/48 → TCT = Mid	0.01
Added patent trends		
(5) IPC = H01L29/788 → CI = Low		0.03
(6) IPC = G11C16/04 → CI = Low		0.02

Table 7
Some changes in the technological competitiveness of companies in Taiwan's semiconductor industry.

Patent trend		Change degree
Emerging patent trends		
(1) Assignee = Macronix International Co. Ltd → CI = Low		2.63
(2) Assignee = Taiwan Semiconductor Manufacturing Co. Ltd → Originality = High		0.01
Unexpected changes in patent trends		
2001–2002	**2003–2004**	
(3) Assignee = Advanced Semiconductor Engineering, Inc. → CI = High	Assignee = Advanced Semiconductor Engineering, Inc. → CI = Low	0.32
(4) Assignee = Siliconware Precision Industries Co., Ltd → Originality = High	Assignee = Siliconware Precision Industries Co., Ltd → Originality = Low	0.03
Added patent trends		
(5) Assignee = Au Optronics Corp. → CI = Low		0.04
(6) Assignee = Nan Ya Technology → CI = Low		0.03
Perished patent trends		
(7) Assignee = Taiwan Semiconductor Manufacturing Co. Ltd → CI = High		0.07
(8) Assignee = Taiwan Semiconductor Manufacturing Co. Ltd → Generality = Mid		0.07

some changes in the R&D activities of Taiwan's semiconductor industry between 2001 and 2004.

In Table 6, the emerging patent trends (1) and (2) show that companies in the industry invested in H01L29/76 and H01L21/00 consistently throughout the period under study. The high growth rates (131% and 107%, respectively) indicate that companies focused their R&D activities on the two technological fields. However, the low CI indicates that the companies lacked pioneer patents and basic patents in these technological fields.

Patent trends (3) and (4) in Table 6 indicate that the TCT of H01L29/40 and H01L21/48 changed from a short-cycle time to a medium-cycle time, which implies that the speed of innovation in these technological fields slowed down. The added patent trends (5) and (6) indicate that H01L29/788 and G11C16/04 were new technological fields that Taiwan's semiconductor companies invested during 2003–2004.

5.4. Technological competitiveness of companies in Taiwan's semiconductor industry

Changes in the technological competitiveness of companies in an industry are identified by comparing the relations between the assignee of the target industry and the citation index, originality, generality, and technology cycle time over two time periods. Table 7 lists some changes in the technological competitiveness of companies in Taiwan's semiconductor industry between 2001 and 2004.

Patent trends (1) and (2) in Table 7 show the consistent innovative power of TSMC and MIC. Specifically, the marked increase in MIC's patents (263%) indicates the innovativeness of MIC and the direction of its R&D activities. However, the low CI indicates that MIC was a technological follower between 2001 and 2004. Patent trend (4) in Table 8 shows a decrease in the Originality of SPIC. The added patent trends (5) and (6) in the table show several new assignees of semiconductor patents, which means that new companies (AOC and NYT) entered the semiconductor industry during 2003–2004.

From the perished patent trends (7) and (8), we observe that the high value of CI and the Generality of TSMC's patents decreased between 2003 and 2004. This implies that the quality of TSMC's R&D may have declined during 2003–2004, although the phenomenon may be due to new patents.

5.5. Technological competitiveness of companies in specific technological fields

Changes in the technological competitiveness of companies in specific technological fields are derived by comparing the relations between both the patent's assignee and the technological field (IPC) of the target industry with the citation index, originality, generality, and technology cycle time over two time periods. Table 8 lists some changes in Taiwan's semiconductor industry between 2001 and 2004.

The frequent appearance of TSMC in emerging patent trends shows that the company played a leading role in Taiwan's semiconductor industry throughout the period under study. The perished patent trends (3) and (4) in Table 8 show that UMC's technological competitiveness with medium CI and low Originality in H01L21/336 declined, which may imply a change in UMC's innovative activities.

Table 8
Some changes in the activities of Taiwan's semiconductor industry.

Patent trend	Change degree
Emerging patent trends	
(1) IPC = H01L21/302, Assignee = Taiwan Semiconductor Manufacturing Co. Ltd → CI = Low	1.4
(2) IPC = H01L21/44, Assignee = Taiwan Semiconductor Manufacturing Co. Ltd → CI = Low	0.78
Perished patent trends	
(3) IPC = H01L21/336, Assignee = United Microelectronics Corp. → CI = Mid	0.02
(4) IPC = H01L21/336, Assignee = United Microelectronics Corp. → Originality = Low	0.02

6. Conclusions

In this study we proposed a patent trend change mining (PTCM) technique that captures changes in patent trends without the need for specialist knowledge and reports changes to business managers by ranking the degrees of change. Competitive intelligence of business is derived by an automatic change mining approach that business managers can modify and develop appropriate strategies according to their findings. The proposed approach mines changes in patent trends by analyzing the metadata in patent documents. We applied the proposed PTCM to Taiwan's semiconductor industry for the period 2001–2004 to discover changes in four types of patent trends: the R&D activities of a company, the R&D activities of the industry, the technological competitiveness of companies and the technological competitiveness of companies in a specific technological field. The results obtained by the proposed approach can be used as an important reference for decision makers to make more accurate strategies on research and development.

There remain several extended researches to do based on this study. The primary part of most patent document is textual content which contains rich information to utilize (e.g., abstracts and claims). Through analyzing the textual part we can surely improve the quality of change detection and provide more comprehensive results. Therefore the next research will be a patent trend change mining approach which utilizes text mining techniques.

Acknowledgement

This research was supported in part by the National Science Council of the Taiwan under the grant NSC 96-2416-H-009-007-MY3.

Appendix A

IPCs belonging to the semiconductor industry identified by the Taiwan Intellectual Property Office.

IPC	Description
C23C	Coating metallic material; coating material with metallic material; surface treatment of metallic material by diffusion into the surface, by chemical conversion or substitution; coating by vacuum evaporation, by sputtering, by ion implantation or by chemical vapor deposition
016/00	Chemical coating by decomposition of gaseous compounds, without leaving reaction products of surface material in the coating, i.e. chemical vapor deposition (CVD) processes
G01R	Measuring electric variables; measuring magnetic variables
031/02	General constructional details
G03F	Photomechanical production of textured or patterned surfaces, e.g., for printing, for processing of semiconductor devices
007/00	Photomechanical, e.g., photolithographic, production of textured or patterned surfaces, e.g., printed surfaces
009/00	Registration or positioning of originals, masks, frames, photographic sheets, or textured or patterned surfaces
G05F	Systems for regulating electric or magnetic variables
001/10	Regulating voltage or current
G11C	Static stores
007/00	Arrangements for writing information into, or reading information from, a digital store
016/04	Using variable threshold transistors, e.g., FAMOS
H01L	Semiconductor devices; electronic solid state devices
021/00	Processes or apparatus specially adapted for the manufacture or treatment of semiconductor or solid state devices or parts thereof
023/34	Arrangements for cooling, heating, ventilating or temperature compensation

Appendix A (*continued*)

IPC	Description
023/48	Arrangements for conducting electric current to or from the solid state body in operation, e.g., leads, terminal arrangements
023/495	Lead-frames
023/52	Arrangements for conducting electric current within the device in operation from one component to another
023/58	Structural electrical arrangements for semiconductor devices
023/62	Protection against over-current or overload, e.g., fuses
027/108	Dynamic random access memory structures
029/00	Semiconductor devices specially adapted for rectifying, amplifying, oscillating or switching and having at least one potential-jump barrier or surface barrier; capacitors or resistors with at least one potential-jump barrier or surface barrier, e.g. PN-junction depletion layer or carrier concentration layer; details of semiconductor bodies
029/40	Electrodes
029/76	Unipolar devices
029/788	With floating gate
029/94	Metal–insulator–semiconductors, e.g., MOS
031/062	The potential barriers being only of the metal–insulator–semiconductor type
031/113	Being of the conductor–insulator–semiconductor type, e.g., metal–insulator–semiconductor field-effect transistor
031/119	Characterized by field-effect operation, e.g., MIS type detectors

References

Agrawal, R., & Srikant, R. (1994). Fast algorithm for mining association rules. In *Proc. of the int. conf. on very large data bases*, Santiago de Chile (pp. 487–499). San Francisco, CA: Morgan Kaufmann.

Berio, G., & Harzallah, M. (2007). Towards an integrating architecture for competence management. *Computers in Industry, 58*(2), 199–209.

Breitzman, A. F., & Mogee, M. E. (2002). The many applications of patent analysis. *Journal of Information Science, 28*(3), 187–205.

Brockhoff, K. K. (1991). Indicators of firm patent activities. In *Technology management: the new international language* (pp. 476–481).

Chang, M. C. (2005). Quantum computation patent mapping – A strategic view for the information technique of tomorrow. In *Proc. of the Int. conf. on service systems and service management* (pp. 1177–1181).

Chang, C.-W., Lin, C.-T., & Wang, L.-O. (2009). Mining the text information to optimizing the customer relationship management. *Expert Systems with Applications, 36*(2), 1433–1443.

Chen, M. C., Chiu, A. L., & Chang, H. H. (2005). Mining changes in customer behavior in retail marketing. *Expert Systems with Applications, 28*(4), 773–781.

Chen, S. Y., & Liu, X. (2004). The contribution of data mining to information science. *Journal of Information Science, 30*(6), 550–558.

CHI-Research. Available from http://www.chiresearch.com.

Dou, H., Leveillé, V., Manullang, S., & Dou, J. J. (2005). Patent analysis for competitive technical intelligence and innovative thinking. *Data Science Journal, 4*, 209–237.

Dürsteler, J. C. (2007). Patent analysis. Available from http://www.infovis.net/, visited 2007/5.

EPO (2007). European Patent Office. Available from http://www.european-patent-office.org/index.en.php.

Han, J., & Kamber, M. (2001). *Mining association rules in large databases. Data mining-concepts and techniques*. San Francisco, CA: Morgan Kaufmann.

Huang, Z., Chen, H., Yip, A., Ng, G., Guo, F., Chen, Z. K., et al. (2003). Longitudinal patent analysis for nanoscale science and engineering: country, institution and technology field. *Journal of Nanoparticle Research, 5*(3–4), 333–363.

Huang, S.-H., Ke, H.-R., & Yang, W.-P. (2008). Structure clustering for Chinese patent documents. *Expert Systems with Applications, 34*(4), 2290–2297.

Ian, H. W., & Eibe, F. (2000). *Output: knowledge representation. Data mining*. San Francisco, CA: Morgan Kaufmann.

Kim, Y. G., Suh, J. H., & Park, S. C. (2008). Visualization of patent analysis for emerging technology. *Expert Systems with Applications, 34*(3), 1804–1812.

Kuo, R. J., Lin, S. Y., & Shih, C. W. (2007). Mining association rules through integration of clustering analysis and ant colony system for health insurance database in Taiwan. *Expert Systems with Applications, 33*(3), 794–808.

Liu, B., Hsu, W., & Ma, Y. (2001). Discovering the set of fundamental rule changes. In *Proc. of the 7th ACM int. conf. on knowledge discovery and data mining* (pp. 335–340), San Francisco, CA.

Liu, B., & Hsu, W. (1996). Post-analysis of learned rules. In *Proc. of 13th natl. conf. on artificial intelligence* (pp. 828–834), Menlo Park, CA.

Liu, B., Hsu, W., Han, H. S., & Xia, Y. (2000). Mining changes for real-life applications. In *Proc. of the 2nd int. conf. on data warehousing and know discovery* (pp. 337–346), London.

Liu, D. R., Shih, M. J., Liau, C. J., & Lai, C. H. (2009). Mining the change of event trends for decision support in environmental scanning. *Expert Systems with Applications, 36*(2), 972–984.

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: a literature review and classification. *Expert Systems with Applications, 36*(2), 2592–2602.

Reitzig, M. (2004). Improving patent valuations for management purposes – Validating new indicators by analyzing application rationales. *Research Policy, 33*(6–7), 939–957.

Song, H. S., Kim, J. K., & Kim, S. H. (2001). Mining the change of customer behavior in an internet shopping mall. *Expert Systems with Applications, 21*(3), 157–168.

Stembridge, B. (2005). Sorting the wheat from the chaff- the use of patent analysis in evaluating portfolios. Available from http://www.scientific.thomson.com/newsletter.

Stembridge, B., & Corish, B. (2004). Patent data mining and effective patent portfolio management. *Intellectual Asset Management*(Oct./Nov.), 30–35.

Tuomo, N., Hermans, R., & Kulvik, M. (2007). Patent citations indicating present value of the biotechnology business. Available from http://www.etla.fi/.

USPTO (2007). United States Patent & Trademark Office. Available from http://www.uspto.gov/.

WIPO (2007). World Intellectual Property Organization. Available from http://www.wipo.int/portal/index.html.en.

Yang, Y. Y., Akers, L., Klose, T., & Yang, C. B. (2008). Text mining and visualization tools – Impressions of emerging capabilities. *World Patent Information, 30*(4), 280–293.

Yen, S. J., & Lee, Y. S. (2006). An efficient data mining approach for discovering interesting knowledge from customer transactions. *Expert Systems with Applications, 30*(4), 650–657.