Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Information Processing and Management 41 (2005) 313–330
INFORMATION PROCESSING & MANAGEMENT
www.elsevier.com/locate/infoproman

Using the patent co-citation approach to establish a new patent classification system

Kuei-Kuei Lai [a,*], Shiao-Jun Wu [b,c]

[a] *Department of Business Administration, National Yunlin University of Science and Technology, Touliu, Yunlin 640, Taiwan, ROC*
[b] *Department of Information Management, Kun Shan University of Technology, Taiwan, ROC*
[c] *Graduate School of Management, National Yunlin University of Science and Technology, Taiwan, ROC*

Received 25 August 2003; accepted 19 November 2003
Available online 23 December 2003

Abstract

The paper proposes a new approach to create a patent classification system to replace the IPC or UPC system for conducting patent analysis and management. The new approach is based on co-citation analysis of bibliometrics. The traditional approach for management of patents, which is based on either the IPC or UPC, is too general to meet the needs of specific industries. In addition, some patents are placed in incorrect categories, making it difficult for enterprises to carry out R&D planning, technology positioning, patent strategy-making and technology forecasting. Therefore, it is essential to develop a patent classification system that is adaptive to the characteristics of a specific industry. The analysis of this approach is divided into three phases. Phase I selects appropriate databases to conduct patent searches according to the subject and objective of this study and then select basic patents. Phase II uses the co-cited frequency of the basic patent pairs to assess their similarity. Phase III uses factor analysis to establish a classification system and assess the efficiency of the proposed approach. The main contribution of this approach is to develop a patent classification system based on patent similarities to assist patent manager in understanding the basic patents for a specific industry, the relationships among categories of technologies and the evolution of a technology category.

Keywords: Patent management; Patent classification system; Co-citation; Bibliometrics

* Corresponding author. Tel.: +886-5-53-42-601x5228; fax: +886-5-53-12-074.
E-mail address: laikk@yuntech.edu.tw (K.-K. Lai).

0306-4573/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ipm.2003.11.004

1. Introduction

A patent is a contract between an inventor and the government, whereby in return for full public disclosure of an invention, the government grants the inventor the right to exclude others for a limited time from making, using and selling the invention (Hufker & Alpert, 1994). With the abundant profits brought by the market monopoly and the strategic use of intellectual property rights (IPRs), patent management has played an important role in the effective operation of enterprises. For instance, the royalty income of IBM has topped a billon US dollars every year, which is approximately at least one ninth of its annual gross profit before tax. The success of Texas Instrument's patent in court enabled the firm to earn higher royalty payments from other firms in the semiconductor industry (Rivitte & Kline, 2000). Thus, IPRs have become important company assets, and patent management has played a pivotal role in corporate management and performance.

The current studies on patent management apply the International Patent Code (IPC) or the United States Patent Code (UPC) to identify patents. Ernst (1997) defines the patents with IPC classification code G05B019/00-G05B019/417 as CNC-technology in the machine tool industry to forecast the diffusion of CNC-technology. Narin, Noma, and Perry (1987) use 15 categories of UPC codes, e.g. code 260, 424, etc., to examine the links between corporate patents and indicators of pharmaceutical corporate performance. However, in terms of patent management, the IPC or the UPC system is too general to satisfy the needs for technological forecasting, research planning, technological positioning or strategy making (Archibugi & Pianta, 1996). The result of the analysis by the above two systems is insufficient to reflect the technological niche of a company and mis-categorization results in further difficulties for patent management. In addition, both the UPC and the IPC system are static systems, which mean they do not evolve with the development of technologies. Thus, the main objective of this study is to propose a patent classification system tailored to the needs of a specific industry for the management of its patents.

This study applies citation approach in bibliometrics to create a patent classification system based on the following grounds: First, dissertations and patents are both instruments that record the results of research. In addition, published dissertations and patent specifications are required to identify their cited documents and patents. Secondly, in bibliometrics, the use of citation approach for the assessment of similarity for the classification of documents is a mature methodology. Given the above, it is feasible to apply citation analysis of bibliometrics to patent classification.

In bibliometrics, there have been extensive studies on the assessment of document similarities. For instance, Kessler (1963) proposed the approach of bibliographic coupling, and Small (1973) proposed the co-citation approach. For bibliographic coupling, citing documents are the subject of the analysis. The degree of bibliographic coupling for documents A and B is reflected in the frequency of the documents that are co-cited by both A and B. The focus of the co-citation analysis is on the documents cited, by calculating the frequency of A and B that are co-cited by specific documents. They assess the similarity of A and B based respectively on the number of co-citing or co-cited documents, as illustrated in Fig. 1. In bibliometrics, the perception of being co-cited is applied to evaluating the similarities in documents because the number of documents co-cited by A and B is limited to that of the reference documents but the number of A and B being