Technological Forecasting & Social Change 94 (2015) 187–201

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Knowledge combination modeling: The measurement of knowledge similarity between different technological domains

Hiroko Nakamura [a,*,1], Shinji Suzuki [a,1], Ichiro Sakata [b], Yuya Kajikawa [c]

[a] Center for Aviation Innovation Research (CAIR), The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, Japan
[b] Innovation Policy Research Center, The University of Tokyo, Tokyo, Japan
[c] Graduate School of Innovation Management, Tokyo Institute of Technology, Tokyo, Japan

ARTICLE INFO

Article history:
Received 17 January 2014
Received in revised form 7 April 2014
Accepted 14 September 2014
Available online 8 October 2014

Keywords:
Patent analysis
Citation analysis
Bibliometrics
Breadth search
Knowledge recombination

ABSTRACT

This paper proposes the DB-Combination model that considers three different knowledge combinations in depth (D) and breadth (B) based on similarities of two technological knowledge domains. We also investigate three methodologies A1, A2 and A3 to highlight the three knowledge combinations. To identify technological knowledge domains, citation analysis on patent information was used for A1 and A2 and pre-existing patent classification analysis was used for A3. And to measure the similarity between identified technological knowledge domains, text similarity measurements, existing intra-industrial citation tracing and IPC share similarity comparison were used for A1, A2 and A3 respectively. The usability of the model and methodologies were demonstrated through a case study on technological knowledge of the automobile industry and the aircraft industry. While these methodologies still need to be improved, it was demonstrated that the three measurements can highlight candidates of the three knowledge combination proposed in DB-Combination model. This research contributes to accelerate breadth knowledge recombination in a complex technology industry.

1. Introduction

It is said that innovation comes from a recombination of knowledge (Dosi, 1982; Nelson and Winter, 1977; Schumpeter, 1934) and that combining one's own knowledge with that of different industry and different technology domains has the possibility of bringing new knowledge creation (Schoenmakers and Duijsters, 2010; Gassmann and Zeschky, 2008; Dosi, 1982). On the other hand, the technological domain of industries with a complex system has a wide range of technological sub-domains (Fleming and Sorenson, 2001; Eriksson, 2000). It is not easy for engineers of an industry to search for a new candidate of knowledge combination in knowledge of another industry with a complex system; firstly they should need to identify the technological sub-domains of the other, and then, to select sub-domains to look further. The information of sub-domains also must be updated frequently (Herrero et al., 2010). The cost of the collection and the integration of different knowledge (Nakamura et al., 2011a; Kajikawa et al., 2006; Tijssen, 1992) and the uncertainty of success (Schilling and Green, 2011; Moorthy and Polley, 2010) are problems that limit practitioners to explore new opportunities.

To support practitioners to have such exploration, namely breadth activities, and to bring innovation, the authors focus on one side of breadth activities; namely, searching the technological knowledge of other industries and integrating it to their own knowledge. We propose a knowledge combination model and discuss methodologies that effectively identify a technological sub-domain that can be combined between the two industries.

We focus on patents as the information source of technological knowledge of focused industries because firstly patent is considered to be the best available indicator for R&D invention related to technology and outcomes of innovation activities (OECD, 1994). And secondly we believed that there is a

* Corresponding author.
E-mail address: techhn@mail.ecc.u-tokyo.ac.jp (H. Nakamura).
[1] Tel.: +81 3 5841 0744.

http://dx.doi.org/10.1016/j.techfore.2014.09.009
0040-1625/

potential need in practitioners for scientometrics that can support breadth search with patents. Practitioners contacted in this study explained that, although a patent is essential to protect practitioners' intellectual products and investigate competitors' strategy, difficulty in searching information from patent data hinders frequent use of patents as a source of knowledge. And computer-based bibliometrics approaches are taken because it can process vast amount of data and it is expected to ease breadth search (Alavi and Leidner, 2001; Smallheiser and Swanson, 1998; Smalheiser, 2012; Herrero et al., 2010; Cantu and Ceballos, 2010; Fleming and Sorenson, 2001; Kostoff, 2008).

This paper is organized as follows: The next section reviews previous literature. The third section proposes a knowledge combination model and three measurements. The fourth section conducts a case study on technological knowledge of automobile and aeronautic industry and shows identified technological domains and highlighted pairs of technological domains. The section also discusses the results with automobile and aeronautic experts. The Discussion section compares the three methodologies. The final section concludes this paper with the findings.

2. Literature

Patent is often used in the innovation literatures. For example, patent is used as the indicators of technological knowledge of focused industries in the literatures mapping technological portfolios of a company or an industry. Leydesdorff et al. (2012) discussed methodologies to map the technological portfolios and the relation between the identified technology using International Patent Classification (IPC) and patent citation analysis approach. Schoen et al. (2012) discussed methodologies to map technological domains of major R&D companies and the dynamics of knowledge over firms and regions. Patent is also used as the indicators of technological knowledge of the focused company in the literatures investigating the impact of technological knowledge breadth and depth to company performance (ex. Schoenmakers and Duijsters, 2010; Leiponen and Helfat, 2010; Moorthy and Polley, 2010; Fleming and Sorenson, 2001; Nesta and Saviotti, 2005). Patent is also used as the indicator of the efficiency of innovation policies of the focused countries, domains, or sectors in the literatures analyzing economic development, technological change, speed and change of industrial structures (ex. Criscuolo, 2006; Soete and Wyatt, 1983). Patent is used in these literatures because it is objective and their standard changes slowly (Griliches, 1990).

Our approach seems to be similar to the first group of literatures, such as those of Leydesdorff et al. (2012) and Schoen et al. (2012). However we use patent not as the indicator to obtain the overview of knowledge existing in a domain but as a source of technological knowledge to produce new knowledge combination. The aim of this paper is neither to describe technological trend and portfolio by utilizing patent as indicator nor to evaluate performance of firms and industries by utilizing patent as evidence. The aim of this paper is to explore opportunities for business development by utilizing patents as information resource.

Citation analysis that we use in this paper has been developed because of the need for scientific information retrieval and has established itself as one of the most effective approaches in identifying technological domains in academic documents and in creating the overview because a citation can reflect the self-organizing dynamics of scholars' communication (Leydesdorff, 2008; Kajikawa et al., 2008, 2006; Nicolaisen, 2007; Cronin, 2001). It is widely known that dissemination and exchange of communication are important for the development of science (Everett and Pecotich, 1993). Using citation information to evaluate scientific activities, nonetheless, brings various concerns such as risk of skewed analysis results due to the existence of citations for critics, self-citations, an English-bias, and availability of literatures (Martin, 2012; Everett and Pecotich, 1993; Garfield, 1979). Despite these concerns, citation analysis is utilized among researchers and decision-makers related to science and innovation management as quantifiable and objective approaches that can compensate and validate the experts' judgments (Nerur et al., 2008), and can be used in administration fields (see Garfield, 1979 for example), and its further development and use are expected by practitioners.

Because patent citations are made with the consideration of legal and economic matters, there is still criticism about applying citation analysis to patents because citation behavior is different between academic journals and patents and also citation analysis inclines too much toward documents with links (Leydesdorff, 2008; Michel and Bettels, 2001; Meyer, 2000; Kostoff, 1998). However, there are many attempts to apply such methods to patent analysis because the pre-existing classification system of patents is based on technological and functional characteristics and is often difficult to understand the overview of the system the patents are related from the classification (Griliches, 1990). Furthermore, when technology rapidly progresses and changes, it is not easy for existing classification schema to capture such a change. Narin (1994) discussed the high similarity between the analysis of scientific papers and analysis of patents because both are suited especially to the analysis of national productivity, inventor productivity, referencing cycles, and citation impact.

Considering the criticism and support regarding patent citation analysis, we conducted two different approaches to identify technological domains from patent data; namely, the citation analysis approach and another approach with an international classification standard of patents as we can see in the next section.

3. Methodology

3.1. Depth and breadth knowledge combination model

The methodology proposed in this paper aims to support the following knowledge recombination process of practitioners, that is, identifying the technological sub-domains of other industries, selecting sub-domains to combine and researching for bringing new knowledge. And to do that, firstly we propose a knowledge combination model between two technological domains, named the DB-Combination model (Fig. 1). We assume that, limited to the technological knowledge and the combination between different industries, knowledge recombination in depth and breadth discussed in the previous innovation literatures such as those ofAlavi and Leidner (2001), Gassmann and Zeschky (2008), Schoenmakers and Duijsters (2010), Dosi (1982), can be modeled depending on the "similarity" between the domains as shown in Fig. 1. Fig. 1