Technological Forecasting & Social Change 90 (2015) 469–475

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Assessing the industrial opportunity of academic research with patent relatedness: A case study on polymer electrolyte fuel cells

Takaya Ogawa [a,*], Yuya Kajikawa [a,b]

[a] Graduate School of Innovation Management, Tokyo Institute of Technology, 3-3-6 Shibaura, Minato-ku, Tokyo 108-0023, Japan
[b] Innovation Policy Research Center, Institute of Engineering Innovation, School of Engineering, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-8656, Japan

ARTICLE INFO

Article history:
Received 25 October 2013
Received in revised form 5 April 2014
Accepted 7 April 2014
Available online 3 May 2014

Keywords:
R&D management
Bibliometrics
Citation network analysis
Patent analysis
Fuel cell

ABSTRACT

The detection of promising academic research is vital for firms in a variety of sectors. Bibliometric tools can be used to detect such research hidden in a pile of papers and patents; however, the relationship between academic research and industrial technology development has not been well documented. In this paper, we introduced patent relatedness, which measures the semantic similarity of papers and patents, and conducted a case study on polymer electrolyte fuel cells (PEFC). The results show that in an academic research area with a small number of papers, recent average publication year, low patent relatedness has a high potential to increase in subsequent years. Research areas are identified by clustering the citation network of academic papers, and their patent relatedness and time series variation were measured and analyzed. Our results showed that potential research areas were characterized by small but emerging features. Using these findings, we identified the potential PEFC research areas and the research capability of each country.

1. Introduction

The demand for innovation to overcome problems such as the dearth of food in developing countries, water shortages, and growing global energy requirements has drawn increasing attention. Academic research is expected to provide fundamental solutions, and its importance is well recognized. A previous study claims that 10% of new products and processes are from academic research [1]. Therefore, the identification of potential academic research areas, which can produce key technologies in industry, has been one of the major challenges in technology and innovation management literature [2–6].

However, this does not mean that all academic research areas will lead to the advancement of industrial technologies. A prior concern of academic research has been to explain unknown phenomena and mechanisms rather than to produce new products and services. The latter are certainly the outcome of academic research, but not its prior aim in most cases. It is also known that although academic research leads to various seeds of innovation, it takes years for this innovation to come to market and be commercialized due to various obstacles. For example, if a new material is found with a promising chemical property, it might also have a low physical property and a low throughput of the manufacturing process. Therefore, the assessment of academic research areas in term of potential for industrial application is a difficult task.

Another practical difficulty arises from the fact that the number of research papers and patents have increased to the point that they cannot all be read manually. It has been stressed that society is currently experiencing an excess of publications [7]. A recent observation reports that review papers and meta-analysis papers have more citation impact than standard articles, which reflects the problem of excessive publications [8]. Although strategic decisions on research and development

* Corresponding author. Tel.: +81 45 924 5793.
E-mail addresses: takaya.o.aa@m.titech.ac.jp (T. Ogawa), kajikawa@mot.titech.ac.jp (Y. Kajikawa).

http://dx.doi.org/10.1016/j.techfore.2014.04.002
0040-1625/

(R&D) should be made after observing and analyzing global R&D trends by reading research papers or patents [2], there is a tendency to miss emerging and potential research areas because of the flood of publications. It brings the risk that R&D will miss potential emerging research areas. Traditional approaches such as reviewing papers and assessing international patent classification trends are still useful in providing a brief overview of a research area. However, these approaches face a fundamental difficulty in keeping up with the rapid pace of development and the moving target of technology.

Bibliometrics has been developed to overcome the above difficulties. In bibliometrics, citation network analysis is an effective way to identify emerging academic research clusters and analyze the characteristics of each cluster without reviewing papers one by one [5,9–12]. For example, citation network analysis has been used to confirm the rapid growth of fuel cell and solar cell technology research in the field of energy research [5]. Ho et al. (2013) applied citation network analysis to research trends and the development path of fuel cell technology [11]. Citation network analysis is also used for mutually influential biofuel research topics [12]. Other researchers have used journal citation data and journal classification data to describe the network of energy-related journals [13,14]. In addition to network analysis, text analysis is used to consider multi-word phrase frequencies and phrase proximities and to extract the taxonomic structure of energy research [15,16].

However, the detection of emerging clusters in academic research is not enough to utilize the results for R&D strategic planning, because academic research does not always aim to develop industrial technology as mentioned before. On the other hand, patent analysis can directly identify present situation of industries, which complement the analysis of academic papers. Daim et al. suggested a forecasting method using bibliometrics and patent analysis [17]. As for energy-related technology, Woon et al. utilized bibliometrics and distributed generation to highlight key trends in related technological developments [18]. Verspagen explained the development trajectory of fuel cell technologies by patent citation networks [19]. However, emerging technology clusters identified from patent analysis seem to be obtained too late to adopt as R&D targets for firms because the technologies included in patent analysis have already been patented.

An alternative approach is to integrate paper analysis with patent analysis. A previous paper reported that the R&D stages of basic research, applied research, development, application, and social impact can be represented by different bibliometric data including papers, patents, newspapers, business, and popular press, respectively [20,21]. In respect to these R&D stages, Martino stated that by "observing a technological innovation at an early stage in this sequence, it may be possible to anticipate when it will reach later stages in the sequence, or at least provide warning that further developments may follow" [21]. Recently, Shibata et al. proposed and demonstrated a bibliometric approach using paper and patent analysis to detect the commercialization gap between academic papers and patents [22]. This approach can detect the academic research area, which has not been patented yet and is the chance for firms to participate in. Later, Leydesdorff et al. renamed this approach as "innovation opportunities explorer" [23].

These methods are beneficial for R&D management to determine the fundamental direction of strategies, but are not sufficient to evaluate specific and potential research areas. Although previous studies have shown the possibility of identifying promising candidates for academic research for future industry applications, we still cannot determine the exact focus of such research with bibliometrics.

In this paper, we introduce a parameter, "patent relatedness," which measures keyword relatedness between patents and academic research clusters. Patent relatedness can be measured in each research cluster, focusing on a specific technology. Hence, we can evaluate the potential of a specific technology using patent relatedness. A case study is done on polymer electrolyte fuel cells (PEFC), which is a major research field in fuel cells. Patent relatedness was applied to assess whether the focal technology in a PEFC research cluster is near commercialization or not. Low patent relatedness of a research cluster means that technologies in the research area have not yet been frequently patented. Additionally, a research cluster with low patent relatedness has many opportunities and allows companies to obtain patents of that technology. Of course, academic research clusters, which will not be commercialized in the long term, are meaningless for industry. Therefore, we analyze the features of research clusters which have low patent relatedness in the analyzed period but will obtain high patent relatedness hereafter. In this paper, we define such clusters as potential research clusters. We determine the features of potential research clusters by determining the parameters relevant to patent relatedness by analyzing several research clusters in time series about PEFC technology. Based on the findings, the future potential PEFC research clusters are identified and reported.

2. Data and methodology

2.1. Data

PEFC is an important technology for reducing energy consumption because of its high energy conversion efficiency. We collected bibliographic data from academic publications on fuel cells. Data for academic papers, including the title, author, publication year, abstract, address, and reference were retrieved from the Science Citation Index Expanded (SCI-EXPANDED), compiled by the Institute for Scientific Information (ISI), Thomson Reuters. The bibliographic data of patents were collected from Derwent enhanced database, Derwent World Patents Index (DWPI), through Thomson Innovation by Thomson Reuters. All patents in the DWPI (data since 1962 as publication year) were collected. We use the same query, "fuel cell*," to collect data from papers and patents. Data collection was done in October 2012. We used Derwent enhanced data with translations to retrieve patents. Patent family was merged and regarded as one node in the network. Application year is used to differentiate the year of the patents. The papers and patents related to PEFC were identified by fuel cell clustering, which is explained in the following section. This is necessary because we collect papers via a simple query, "fuel cell*", and our collected dataset includes papers on other types of fuel cell technologies. When these different types are included in the analysis, interpreting the results becomes complex. Therefore, in this paper, we focus only on PEFC. One plausible approach to collecting only PEFC papers would be to replace the query,