Technological Forecasting & Social Change xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

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

Exploring technology opportunities by visualizing patent information based on generative topographic mapping and link prediction☆

Byungun Yoon[a,*], Christopher L. Magee[b]

[a] Department of Industrial & Systems Engineering, School of Engineering, Dongguk University, Seoul, Republic of Korea
[b] Institute for Data, Systems and Society, Massachusetts Institute of Technology, Cambridge, MA, USA

ARTICLE INFO

Keywords:
Technology opportunity analysis
Visualization
Patent information
GTM
Link prediction

ABSTRACT

The shortening lifetime of technology requires companies to make intensive efforts to continuously explore new technology. Although many researchers have proposed visualization methods to find technology opportunities, little attention has been paid to present detailed directions of technology development with specified characteristics of technology. Thus, this research aims to suggest a systematic approach to conducting technology opportunity analysis by visualizing patent information, such as patent documents and citation relationships. First, keywords that explain core concepts, functions, and so on are extracted from collected patent documents by text mining. Second, patents are visualized in a two-dimensional space, and vacant cells are identified with their estimated keyword vectors by generative topographic mapping (GTM). Third, since many vacant cells will be potential candidates for developing new technologies, link prediction tools can choose promising vacant cells to connect existing cells with potential, but not yet existent, cells. Finally, the results of prediction are tested by comparing the predicted cells with the actual developed cells. The research reported in this paper is based in three technologies that have emerging, stable, and declining patterns, in order to illustrate the proposed approach, and investigate in which types it is relevant. It is found that the proposed approach provided a good prediction performance in the case of a technology that has a stable pattern. In addition, among link prediction methods, a semantic similarity-based approach showed better prediction results than a machine learning technique due to modest data availability for training. Thus, the results of this research can help R&D managers plan and evaluate R&D projects for technology development.

1. Introduction

In an uncertain economy, the survival of companies has been more challenging than in the past, due to rapidly changing customer needs and technology. In particular, it is pertinent to note the importance of technology, because since the mid-20th century, it has been considered as the key driver of innovation (Dismukes et al., 2005; Schumpeter, 1939). Moreover, disruptive innovation, in restructuring the current industries or markets, can make existing paradigms obsolete. Although many factors can influence the successful management and competition of companies, technology has been regarded as a critical trigger to competitive edges. Emerging technologies have the potential to disrupt the status quo, by changing the patterns of resource utilization, and rearranging value pools (Cheng et al., 2017). Thus, most companies invest a considerable portion of their budget in supporting the activities of technology innovation. Among the subjects of technology innovation management, the identification of technology opportunities is a starting point for managing the subsequent activities, such as technology acquisition and exploitation. New opportunities for disruptive, as well as sustaining technology should be explored to generate considerable profits for companies (Cozzens et al., 2010, 2010; Porter and Newman, 2011).

The concept of technology forecasting has become popular, as it includes systematic activities to predict and comprehend the direction, rate, characteristics, and effects of technological change (Coates et al., 2001). Thus, many researchers have been actively interested in technology opportunities, because the research theme contains academic, as well as practical value (Savioz and Blum, 2002). Since exploring technology opportunities is related to forecasting and must consider the unique characteristics of a technology, both theoretical and methodological approaches need to be applied to address the prediction problems in a substantial way. At the same time, from a practical

☆ It is confirmed that this item has neither been published nor is currently being submitted elsewhere.

* Corresponding author at: Department of Industrial & Systems Engineering, School of Engineering, Dongguk University, 3-26, Pil-dong 3ga, Chung-gu, Seoul 100-715, Republic of Korea.
E-mail address: postman3@dongguk.edu (B. Yoon).

https://doi.org/10.1016/j.techfore.2018.01.019
Received 27 July 2017; Received in revised form 25 December 2017; Accepted 18 January 2018

perspective, technology intelligence that can support decision makers for technology planning by identifying risks and opportunities of the technological environment has been implemented in many global companies (Kerr et al., 2006). However, the existing literature on technology opportunity has some limitations, in terms of analytic methods, performance, and detailed specification. First, most of the existing approaches are based on qualitative processes, and dependent on expert opinion. Thus, reliable outputs are hardly provided for decision making; and moreover, their validation is very difficult, rendering the decisions for technology opportunity still uncertain. Second, while many researchers utilized visualization in order to explore vacant areas for technology development as technology opportunities, the true meaning of a technology vacancy has not been specified. In addition, the results are often weak in designing the details of a technology, because they focus on the provision of the areas of technology opportunities. Third, in even quantitative approaches, the forecasting of technology opportunities is not systematic, because most of the methods concentrate on finding rough vacant areas of technology maps or promising categories, rather than predicting details of technology opportunities. Thus, the results that are derived by existing methodologies are at best weakly connected to practical actions for R&D planning.

In order to address these limitations, this research aims to propose a systematic approach to the exploration of technology opportunities. Basically, patent information, such as patent documents and citations, is utilized to discover promising technology opportunity in a quantitative manner. Since the visualization of patent information can provide an effective locus to find technology opportunity, a process to visualize a large amount of patent information is employed in the form of patent maps. For this, text mining techniques are utilized to extract valuable keywords from unstructured texts that are recognized as a basis for visualization. In addition, and more importantly, the meaning of vacant areas in a patent map needs to be clarified, in order to support a decision making process to plan technology development. Thus, generative topology mapping (GTM) is utilized to identify the meaning of technology opportunities, by estimating their keywords in the map. Since there are many vacant areas in a map that have no existing patent, the promising vacant areas should be forecasted by applying a systematic approach. In this research, a patent network is additionally visualized on a patent map with patent citation information, and link prediction methods are then employed to forecast the future links, among pairs of cells in the map that previously had no links. Thus, this research combines the patent map and patent network as visualization methods, in order to explore vacant areas, and predict missing links in a patent map.

This paper is structured as follows: Section 2 reviews the theoretical background on the main research ideas and methodologies, including technology opportunity analysis, visualization of patent information, and link prediction. Section 3 explains the basic concept and overall process of the proposed approach for exploring technology opportunities. Section 4 presents real technologies for illustration, while Section 5 discusses the managerial and methodological implications. Finally, Section 6 reviews the limitations, and suggests future research following up on this paper.

2. Theoretical background

2.1. Technology opportunity analysis

Technological opportunity can be defined as a set of possibilities for technology advance to enhance either the functions of products, or their production (Olsson, 2005). Many researchers have conducted academic research to identify innovative opportunities to forecast emerging technologies (Savioz and Blum, 2002). Technology opportunity analysis (TOA) that explores and evaluates the risks, as well as the opportunities of technology development, is a type of technology forecasting, because it aims to predict the future scenarios of technology evolution. However, several experts must spend a lot of time to understand, evaluate and specify technology opportunities in practical R&D planning (Lee et al., 2017). Thus, qualitative and quantitative approaches can be applied to TOA, like technology forecasting. In terms of qualitative methodologies, the Delphi approach is effective to synthesize the various opinions of domain experts. In contrast, data analysis can quantitatively support the process to explore promising technology opportunities, by utilizing systematic techniques, such as text mining, data mining, and machine learning.

In terms of data sources for TOA, technological information can be classified into four categories, including patents, scientific and technical publications, people, and products and processes (Granstand, 1999). Since patents comprise a large amount of data as direct outputs of R&D activities, and are evaluated and generated to an international standard, they are often utilized to analyze technology opportunities. In particular, patent analysis has been regarded as a core analytical tool for technology opportunity analysis by employing computerized tools, such as text mining and bibliometric analysis (Lee et al., 2011). Most patent analyses utilize visual forms, such as charts, graphs and networks, to facilitate the exploration of technology opportunities. Thus, various datamining techniques have been integrated with patent analysis, adding more systematic processes to the conventional statistical approaches. Among these, text mining is intensively employed in TOA, because it can analyze the contents of documents that are very critical for understanding the characteristics of technology in technological texts. The text mining techniques are frequently combined with other techniques to investigate technological trends. For example, text mining was used with the Theory of Inventive Problem Solving (TRIZ) to analyze technological evolutionary trends (Wang et al., 2010). Elsewhere, network analysis and citation analysis were combined with text mining to visualize technological and business opportunities (Lee et al., 2009).

Although patent analysis is a representative method for TOA, the combination of scientific and technological knowledge nourishes the process of exploring new technology opportunities (Wang et al., 2015). Scientific knowledge provides the foundation for technological knowledge, and feedback stimulus from technological knowledge can catalyze continuous search of scientific knowledge (Glänzel and Meyer, 2003). It is controversial as to which type of knowledge precedes the other, because patents are often issued before scientific papers on the topic are published. In general, most people think that scientific knowledge is applied to develop patents, because research based on science is close to basic research, and technological knowledge can be obtained by applied research. Since the two types of knowledge are complementary, the data on scientific discovery can improve technological development and commercialization (Hellmann, 2007). Although many researchers have used patent data to concentrate on the utilization of technological knowledge, some studies have tried to integrate the two main data sources, of academic papers and patents, in order to explore technology opportunities. Shibata et al. (2010) applied social network analysis to develop citation networks of science and technology based on citation information in the references of papers, as well as patents. Technology opportunities can then be identified by determining the gaps between clusters of patents and papers.

2.2. Visualization of patent information

Information visualization can help analysts gain an overview of objects in a two-dimensional or three-dimensional space (Kwakkel et al., 2014). The main advantage of information visualization is to extract a critical insight from a large amount of information by reducing the dimensions of data. In the recent patent analysis, various visualization tools, such as maps and networks, have been utilized to investigate patent information. Since the patent database has a very large number of patent documents, as well as potential patent infringement