ARTICLE IN PRESS

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

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

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

# Generating novel research ideas using computational intelligence: A case study involving fuel cells and ammonia synthesis

Takaya Ogawa*, Yuya Kajikawa

*Graduate School of Innovation Management, Tokyo Institute of Technology, 3-3-6 Shibaura, Minato-ku, Tokyo 108-0023, Japan*

A R T I C L E I N F O

*Keywords:*
R & D management
Bibliometrics
Keyword similarity
Ammonia synthesis
Fuel cell

A B S T R A C T

We proposed a method to help researchers create novel research ideas using bibliometrics. Different concepts and techniques exist in different research areas, and when the fields are sufficiently similar, a salient combination of two different areas can lead to the development of novel research. We have assumed that two different research areas, sharing a high number of similar keywords, would be excellent candidates for integration. We combined link mining and text mining techniques to elucidate hidden but implicit opportunities among apparent, explicit research clusters. To demonstrate the effectiveness of our approach, we conducted a case study on fuel cells and ammonia synthesis. Fuel cells are a rapidly growing research field, while ammonia synthesis is relatively mature. Our results successfully extracted a plausible and post-mature research idea.

## 1. Introduction

Ideas for new and original products, services, systems, and techniques generally derive from the human ability to imagine, design, and invent. Impressive original works that result in the creation of value are regarded as innovation. Innovation, the key to much human activity and societal advancement, is expected to promote development and to resolve both overt and latent social problems. In addition to the enthusiastic experimentation of practitioners, this method for achieving innovation has been theoretically and empirically studied in academia. However, as we know well, no formula or formal methodology can guarantee innovation. Academic research does not generally support innovation because it focuses on understanding how innovation occurs from a scientific point of view rather than on using engineering and design to innovate. The recent development of computational intelligence offers new ways to support intelligent human invention and innovation. Computational creation and the support of salient and innovative ideas constitute a great challenge; they can help to empower human beings and develop society.

Among myriad methods for creating innovation, one possible approach is to combine different types of knowledge (Schumpeter, 1934). The combination, integration, and fusion of different areas of expertise can generate a seed of original knowledge (Swanson, 1986). Human activity often involves collaboration as people combine knowledge to uncover novel insights (Katz and Martin, 1997). However, the extent of accumulated human knowledge is so huge that no one researcher can make use of it all (Kajikawa et al., 2006). There must therefore be many undetected combinations of knowledge with the potential to produce innovations.

The field of bibliometrics has recently been developed to overcome the above difficulty. In bibliometrics, citation network analysis is being used effectively to identify emerging academic research clusters and to analyze their characteristics without the need to review individual papers. 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 (Kajikawa et al., 2008). Ho et al. have applied citation network analysis to research trends and the development path of fuel cell technology (Ho et al., 2014). Citation network analysis has also been used to identify mutually influential biofuel research topics (Kajikawa and Takeda, 2008). Other researchers have used journal citation data and journal classification data to describe the network of energy-related journals (Dalpe and Anderson, 1995; Tijssen, 1992). In addition to network analysis, text analysis is used to assess multi-word phrase frequencies and phrase proximities, and to extract the taxonomic structure of energy research (Kostoff et al., 2002, 2005).

The existing literature on bibliometrics tends to focus on describing the overall structure and trend of a selected research domain using either text or citation analysis. Recently, the combined use of citation network and text analysis has been developed; this computational method can reveal linkages between two research areas using publication data. Citation network analysis is a powerful tool for illuminating the explicit relationships among papers, while text analysis can be used to extract implicit relationships. After extracting citation clusters based on explicit data (i.e., citations), text analysis can be applied to extract

* Corresponding author.
*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.2017.04.004
Received 7 September 2015; Received in revised form 27 February 2017; Accepted 5 April 2017

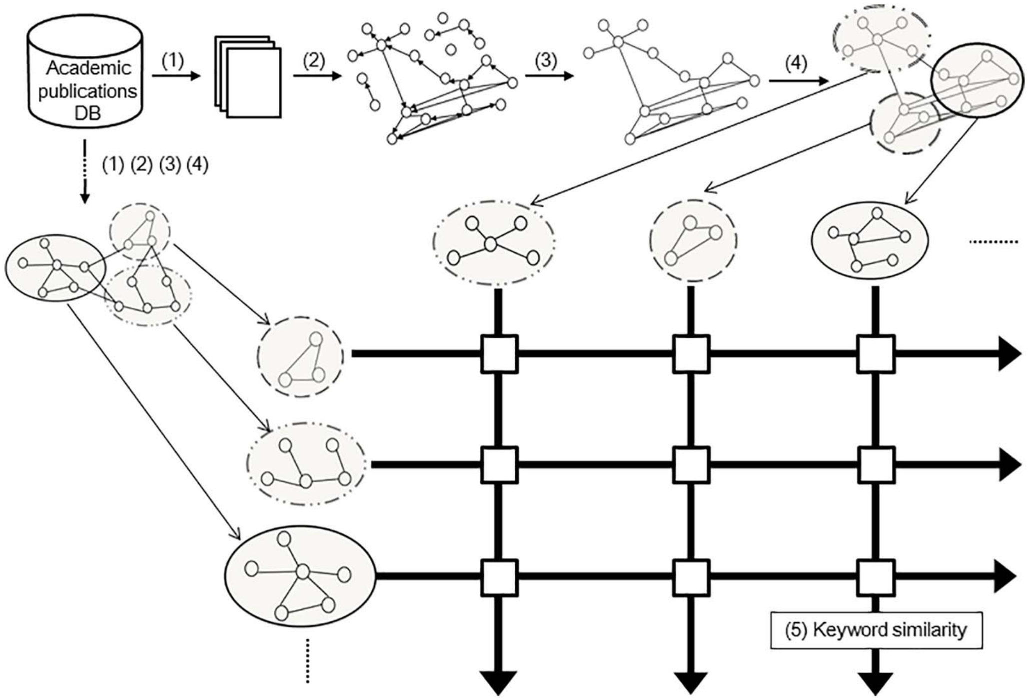


**Fig. 1.** Schematic illustration of clustering and measurement using keyword similarity.

the hidden relationships that do not appear in the citation information. For example, Shibata et al. compared structures related to the citation network of scientific publications with those of patents in solar cell research and discussed whether linkages existed, revealing academic research areas that were not yet commercialized (Naoki Shibata and Sakata, 2010). Ogawa et al. analyzed research elements with commercial potential in areas related to polymer electrolyte fuel cells (Ogawa and Kajikawa, 2015). They used network analysis to divide academic research papers into research clusters, measuring "patent relatedness" by using textual similarities between the research cluster and existing patents. Nakamura et al. highlighted unexplored areas, including water use in the aviation industry, which was detected by combining the citation and text analysis of environmental and aviation industry issues (Nakamura et al., 2014). In the same way, Vitavin et al. succeeded in finding links between robotics and gerontology, leading to the development of new ways to use robots to help elderly people (Ittipanuvat et al., 2014).

In this study, we applied the computational method to create a novel research idea by combining knowledge drawn from two different research areas using network analysis and keyword similarity. We assumed that recent areas of research have a lot of novel technologies and will eventually develop into traditional research areas. For the recent research area, the computational method was applied to the fuel cell, which is a highly energy-efficient device that was developed recently. The traditional research area was ammonia synthesis, an extremely important process that has changed very little during the past 100 years and still relies on the Haber-Bosch process, which requires very high pressure and temperatures. The final goal of this research was to develop a new approach that can revitalize the traditional research area of ammonia synthesis by applying the latest fuel cell techniques. Bibliometric data from both research areas were divided into several clusters (focusing on each individual technology) using network analysis. The divided clusters of one area were compared with clusters from the other research area using keyword similarity so that areas of relatedness could be identified. A high level of keyword similarity between clusters suggested that techniques used in those clusters could be shared. In the next section, we will provide illustrations of the methodology adopted in this paper.

## 2. Data and method

### 2.1. Data

We collected bibliographic data from academic publications on fuel cells and ammonia synthesis. Data from these academic papers, including the title, author, publication year, abstract, address, and references, were retrieved from the Science Citation Index Expanded (SCI-EXPANDED), compiled by the Thomson Reuters Institute for Scientific Information (ISI). We used the query "fuel cell*" to collect data on fuel cells. In the case of ammonia synthesis, we used a rather complex query to identify and recall relevant papers. The query was as follows: "ammonia synthesis*" or "synthesis* of ammonia" or "ammonia formation*" or "formation* of ammonia" or "$NH_3$ formation*" or "formation* of $NH_3$" or "nitrogen protonation*" or "protonation* of nitrogen" or "$N_2$ protonation*" or "protonation* of $N_2$" or "$NH_3$ synthesis*" or "synthesis* of $NH_3$" or (("nitrogen fixation*" or "fixation* of nitrogen") and ("ammonia" or "$NH_3$")) or (("$N_2$ fixation*" or "fixation* of $N_2$") and ("ammonia" or "$NH_3$")) or "synthetic* ammonia" or "synthetic* of ammonia" or "synthetic* $NH_3$" or "synthetic* of $NH_3$" or "ammonia synthetic process*" or "synthetic process* of ammonia" or "$NH_3$ synthetic process*" or "synthetic process* of $NH_3$" or "ammonia production*" or "production* of ammonia" or "$NH_3$ production*" or "production* of $NH_3$." The query terms were determined using the following procedure. First, "ammonia synthesis*" and "synthesis* of ammonia" were used to collect papers. Then, we read the collected papers and added terms such as ammonia formation, $N_2$ fixation, and synthetic ammonia. Papers were collected again using the new query; we read them and searched for other terms. We repeated this process until an additional term did not increase the total number of papers by more than 100 over the number collected without the additional term. Data collection was carried out in December 2012.

### 2.2. Method

Our analytical procedure is schematically shown in Fig. 1. In Step (1), the data from academic papers were downloaded. In Step (2), we constructed citation networks by treating the papers as nodes and the