Research Trends with Cross Tabulation Search Engine

Chengjiu Yin, Research Institute of Information Technology, Kyushu University, Fukuoka, Japan

Sachio Hirokawa, Research Institute of Information Technology, Kyushu University, Fukuoka, Japan

Jane Yin-Kim Yau, Department of Computer Science, Malmö University, Sweden

Kiyota Hashimoto, Osaka Prefecture University, Sakai, Japan

Yoshiyuki Tabata, Research Institute of Information Technology, Kyushu University, Fukuoka, Japan

Tetsuya Nakatoh, Research Institute of Information Technology, Kyushu University, Fukuoka, Japan

ABSTRACT

To help researchers in building a knowledge foundation of their research fields which could be a time-consuming process, the authors have developed a Cross Tabulation Search Engine (CTSE). Its purpose is to assist researchers in 1) conducting research surveys, 2) efficiently and effectively retrieving information (such as important researchers, research groups, keywords), and also 3) providing analytical information relating to past and current research trends in a particular field. Their CTSE system employs data-processing technologies and emphasizes the use of a “Learn by Searching” learning strategy to support students to analyze such research trends. To show the effectiveness of CTSE, a pilot experiment has been conducted, where participants were assigned to do research survey tasks and then answer a questionnaire regarding the effectiveness and usability of the system. The results showed that the system has been helpful to students in conducting research surveys, and the research trend transitions that our system presented were effective for producing research trend surveys. Moreover, the results showed that most students had favorable attitudes toward the usage and usability of the system, and those students were satisfied in gaining more knowledge in a particular research field in a short period.

Keywords: Analysis, Learning by Searching, Research Trends, Search Engine, Text mining

1. INTRODUCTION

When initiating a new research project, researchers are often required to conduct a research survey (also known as a literature review), collect papers of relevance and analyze past and emerging research trends. These activities are integral to the research process for researchers of all types including those situated in governments or professional organizations as well as academic ones. As indicated by Hwang and Tsai (2011), “results analysis could help
policymakers in governments and researchers in professional organisations to allocate the necessary resources and make plans for supporting future research and applications; Doing research survey could be good references for educators and researchers who plan to contribute to the relevant studies.” Being able to predict and detect emerging research topics is often a desired (or essential, even) element of conducting high-quality research. A number of methods have been proposed in the literature review within the area of information-based systems to help researchers identify emerging topics within a research area. Five of these different methods/systems are described below:

1. Bun (2005) proposed an Emerging Topic Tracking System (ETTS) which is an information agent for detecting and tracking emerging topics from a particular information area on the Web. It uses a new TF*PDF (Term Frequency * Proportional Document Frequency) algorithm to detect the changes in the information area of the user’s interest and generate a summary of these changes for the user at set intervals of time. This summary of changes consists of the latest most discussed research issues and may, as a result, reveal an emerging topic or topics.

2. Decker et al. (2007) used a semantic approach that proposes a method for extracting the names of those researchers in their early stages of a research area, indicated by the amount of high-quality publications. It can be effective in retrieving many exact matches of researchers that have major contributions within the research area being explored.

3. The Hierarchical Distributed Dynamic Indexing (HDDI) system mentioned in Bouskila and Pottenger (2000) aims to identify features and methods for improving the automatic detection of emerging trends by generating clusters based on semantic similarity of textual data. The rate of change in the size of clusters and in the frequency and association of features is used as input for applying machine learning techniques to classify topics as emerging or non-emerging.

4. Collaborative Inquiry-based Multimedia E-Learning (CIMEL) is a multi-media framework for constructive and collaborative inquiry-based learning (Blank et al., 2001). The semi-automatic trend detection methodology described in (Roy et al., 2002) has been integrated into the CIMEL system in order to enhance computer science education. Citations traces are used with pruning metrics to generate a document set for an emerging trend. Following this, threshold values are tested to determine the year that the trend was emerged. A multimedia tutorial has been developed to guide students through the process of emerging trend detection.

5. Moreover, some researchers used bibliometric methodologies to analyze the trends and forecasts in different domains, such as e-commerce, supply chain management and knowledge management (Tsai, 2011; Tsai & Chi, 2011; Tsai & Chiang, 2011). Using a bibliometric approach, (Tsai & Yang, 2010) analyzed data mining and CRM research trends from 1989 to 2009 by locating headings “data mining” and “customer relationship management” or “CRM” in topics in the SSCI database. Especially, the approach utilized search categories such as publication year, citation, country/territory, document types to explore the differences in the two fields.

As mentioned above, research trend survey is an essential preliminary step for conducting any academic or non-academic research. In particular, many junior researchers experience difficulties in locating the appropriate keywords and subsequently experience difficulties for conducting a literature review in their research field(s). With the development of ICT technologies such as data-processing, it is possible to design search engines to address such learning difficulties or research needs. Data-processing is a broad category which includes functions such as search engines, data mining, recommendations, and image recognition; and such
Blended Learning Environments
[www.igi-global.com/chapter/blended-learning-environments/12102?camid=4v1a](www.igi-global.com/chapter/blended-learning-environments/12102?camid=4v1a)

Generic E-Assessment Process Development based on Reverse Engineering
[www.igi-global.com/article/generic-e-assessment-process-development-based-on-reverse-engineering/176355?camid=4v1a](www.igi-global.com/article/generic-e-assessment-process-development-based-on-reverse-engineering/176355?camid=4v1a)