Chapter 7

Implementing Background Net with Knowware System for Personalized Keyword Support

Yuan Chen  
Macau University of Science and Technology, China

Liya Ding  
Macau University of Science and Technology, China

Sio-Long Lo  
Macau University of Science and Technology, China

Dickson K. W. Chiu  
Dickson Computer Systems, Hong Kong

ABSTRACT

This article proposes a novel approach that combines user’s instant requirement described in keywords with her or his long-term knowledge background to better serve article selection based on personal preference. The knowledge background is represented as a weighted undirected graph called background net that captures the contextual association of words that appear in the articles recommended by the user through incremental learning. With a background net of user constructed, a keyword from the user is personalized to a fuzzy set that represents contextual association of the given keyword to other words involved in the user’s background net. An article evaluation with personal preference can be achieved by evaluating similarity between personalized keyword set based on user’s background net and a candidate article. The proposed approach makes it possible to construct a search engine optimizer running on the top of search engines to adjust search results, and offer the potential to be integrated with existing search engine techniques to achieve better performance. The target system of personalized article selection can be automatically constructed using Knowware System which is a development tool of KBS for convenient modeling and component reuse.

INTRODUCTION

The rapid increase of information distributed through the Internet (Chiu et al., 2010) has recently raised an important and emerging requirement for a more effective selection of information based on various criteria. Personalized article selection among the topics in this field is currently attracting a lot of attention (Chen & Yang, 2003; Makris et al., 2007; Micarelli et al., 2007; Speretta & Gauch, 2005). Knowledge-based intelligent approaches have been applied in the capturing and learning of relatively stable and long-term criteria for information retrieval and classification, such as fuzzy data retrieval (Takagi & Kawase, 2001) and

DOI: 10.4018/978-1-4666-2470-2.ch007
Implementing Background Net with Knowware System for Personalized Keyword Support
text classification (Lv & Liu, 2005). On the other hand, keyword-based search techniques have shown their obvious advantages in practical applications for their convenience of representing and handling of user’s dynamic desire for article selection. However, keywords can only capture user’s instant desires. Therefore, many existing search engines actually apply a lot of supplementary techniques at backstage in augmenting front keyword processing, e.g., keyword extraction from websites or articles (Bracewell et al., 2008).

In particular, we observe that a same keyword often links to different interpretations for different people. It should be helpful if we can make good use of personal knowledge background for personalizing keywords in order to better capture and fulfill personal preferences in article selection. Such a knowledge background works at the backstage and makes use of keywords as instant triggers to “fire” the application of relevant knowledge for a specific task of selection. With this understanding, we propose in this article a novel approach that takes into account of user instant requirements as well as their long-term knowledge background to improve article selection.

The rest of the article is organized as follows. First, we give a brief explanation of the significance of personal knowledge background for article selection and introduce the overall processing flow of the proposed approach. Next, we propose a novel knowledge representation for capturing user’s long-term interest then we provide the details of algorithm of article selection, and show a system application design of background net for help user for achieving personalized information retrieval. Finally, we conclude our paper with our continuing and future work.

PERSONAL KNOWLEDGE BACKGROUND FOR ARTICLE SELECTION

A classical approach is to use the cosine similarity by frequency of keywords for article selection. However, this is often insufficient, as it may be too rough in decision. One improvement is the Graph Model (Wu et al., 2009) that considers both the frequency and position property of keywords in articles to obtain a more accurate result by calculating the similarity of two articles. Thus, this method can also used for detecting fraudulent or plagiarized articles, which can be hardly done by only using keyword frequency. There are also other methods proposed to improve the quality of article selection, such as the use of ontology in text classification (Yang et al., 2009), interest mode (Zhao et al., 2011; Teevan et al., 2005) and semantic-based methods (Lv & Liu, 2005).

However, those methods are not dealing with text documents with personal preferences. A word may often have different meanings to different people and thus it is reasonable to obtain different results with the same keywords given when personalization is considered. In this sense, a keyword is not purely a symbol but a symbolized representation of a concept that can be personal in terms of its meaning.

As a research topic, personalized article selection has attracted a lot of attention for its possible applications in Search Engine Optimization (SEO), such as Brin et al. (1999) and Leubner and Kießling (2002). However, most of them capture personal preferences only in a limited space by a rather inflexible set of features represented by keywords, and unable to represent a broad range of user long-term interest and knowledge with incremental learning.

In our approach, we propose a novel background net (Chen et al., 2011) model to capture the knowledge of personal interest in a long-term and accumulative manner. Further, we use fuzzy sets (Zadeh, 1965, 1997) to express a personalized keyword as its contextual association with other words in the same background net. With sample articles provided, the background net of an individual user can be automatically constructed through incremental learning. An output of article selection selected, filtered, and possibly also re-ranked from the archive using user’s background net according to specified keywords.