Intelligent Information Retrieval Using Fuzzy Association Rule Classifier

Sankaradass Veeramalai, Anna University, India
Arputharaj Kannan, Anna University, India

ABSTRACT

As the use of web applications increases, when users use search engines for finding some information by inputting keywords, the number of web pages that match the information increases at a tremendous rate. It is not easy for a user to retrieve the exact web page which contains information he or she requires. In this paper, an approach to web page retrieval system using the hybrid combination of context based and collaborative filtering method employing the concept of fuzzy association rule classification is introduced and the authors propose an innovative clustering of user profiles in order to reduce the filtering space and achieves sub-linear filtering time. This approach can produce recommended web page links for users based on the information that associates strongly with users’ queries quickly with better efficiency and therefore improve the recall, precision of a search engine.

Keywords: Association Rule, Fuzzy Association Rule, Hybrid Filtering Method, Recommendation System, User Profile Clustering

1. INTRODUCTION

As the Internet usage rate has rapidly increased, the volume of electronic documents that matches the user’s interest can be seen on the web has also substantially increased. The User gives the keywords as input to a search engine and the search engine returns a set of pages that are related to the key word topics or terms. The result page set consists of too much irrelevant information and may the relevant information of the user’s choice cannot be retrieved. In particular, users have to browse and view sites one after another for a long time until they are satisfied to have a good set that is relevant to their interest or more likely, they give up the search.

Earlier recommendation system uses the association concept for classification purpose but Association rule mining does not have a fixed target (Lin, Alvarez, & Ruiz, 2002; Lei, Kang, Lu, & Yan, 2005). That is, any item can appear on any part of a rule. The use of association rules for classification in recommendation system leads to problems where the instances can only belong to a discrete number of classes. The reason is that association rule mining is only possible for attributes that are
subjected to the terms of keywords. However, in general the usage of association rules has its own disadvantages. Every item, which is not present in the entire body of the rule, may occur in the head of the rule.

To cope with this problem, we propose a recommendation engine for supporting information retrieval system based on the classification of web pages by fuzzy association rules and also we have introduced user profile clustering concept to reduce retrieval time of the user profile from the group of user profile. The framework of fuzzy classified association rule mining is presented which incorporates fuzzy set modeling based on the time in an existing association rule mining technique (Kumar & Thambidurai, 2004; Lei, Kang, Lu, & Yan, 2005). Interoperability architecture for building customized search engines is developed by Cecil et al. Through this, the existing search engines are decomposed into self-contained components (Chua, Chiang, & Storey, 2009).

In order to implement efficiently these services, information filtering systems rely upon clustering of user profiles (Greening, 1998; Wang, Xie, & Li, 2006; Mostafa, Mukhopadhyay, Lam, & Palakal, 1997). This approach clusters Web site users into different groups and generates common user profiles so that these profiles can be filtered in the recommendation system. This concept improves the ability to capture the precariousness among Web user’s navigation behaviour. The usage of clustering methodology increases the filtering efficiency with respect to time.

The goal of this proposed recommender Systems is to generate web pages about new items and to predict the substitute of a specific item for a particular user.

2. RELATED WORK

The primary goal of information systems is to retrieve or filter objects and classify them based on the rules described in the system. The filtering system may be classified as context based filtering or collaborative filtering (Zhou, Li, Bruza, Wu, Xuusing, & Lau, 2007; Wang, Xie, & Li, 2006). In information-filtering environments, uncertainties associated with changing interests of the user and the dynamic document stream is handled (Mostafa, Mukhopadhyay, Lam, & Palakal, 1997) by the filtering algorithm. The classification process is based on association rule mining technique where a rule is defined to make distinct classification. Recommendation using association rules is to predict preference for item k when the user preferred item i and j, by adding confidence of the association rules that have k in the result part and i or j in the condition part. Association rules capture relationships among items based on patterns of co-occurrence across transactions (Kumar & Thambidurai, 2004). The hierarchical structure is also used for classifying a large, heterogeneous collection of web content. The hierarchical structure is initially used to train different second-level classifiers. In the hierarchical case, a model is learned to distinguish a second-level category from other categories within the same top level (Dumais & Chen, 2000).

2.1. Recommendation System

The aim of recommendation system is to select information whose content are most relevant to the user’s interest from a greater volume of information available and to present them in suitable way for the user (Fernandez, Calderón, Barrenechea, Bustince, & Herrera, 2010). A recommendation system makes suggestions about web pages to a user. Social recommendation methods collect ratings of web pages from many individuals to make recommendations to a user concerned (Basu, Hirsh, & Cohen 1998). OntoClippy is a tool-supported methodology for the user-friendly design and creation of ontologies. Existing ontology design methodologies and tools are targeted at experts and not suitable for users without a background in formal logic (Dahlem, 2011). Web information recommendation fall into three categories:

1. Recommendation based on the similarity of web pages.
A Computational Model of Mitigating Disease Spread in Spatial Networks
Investigations into Living Systems, Artificial Life, and Real-World Solutions (pp. 106-122).
www.igi-global.com/chapter/computational-model-mitigating-disease-spread/75923?camid=4v1a