2012 International Conference on Applied Physics and Industrial Engineering

The Personalized Information Retrieval Model Based on User Interest

Songjie Gong
Zhejiang Business Technology Institute, Ningbo 315012, P. R. China

Abstract

Personalized information retrieval systems can help customers to gain orientation in information overload by determining which items are relevant for their interests. One type of information retrieval is content-based filtering. In content-based filtering, items contain words in natural language. Meanings of words in natural language are often ambiguous. The problem of word meaning disambiguation is often decomposed to determining semantic similarity of words. In this paper, the architecture of personalized information retrieval based on user interest is presented. The architecture includes user interface model, user interest model, detecting interest model and update model. It established a user model for personalized information retrieval based on user interest keyword list on client server, which can supply personalized information retrieval service for user with the communications and collaboration of all modules of the architecture.

© 2011 Published by Elsevier B.V. Selection and/or peer-review under responsibility of ICAPIE Organization Committee. Open access under CC BY-NC-ND license.

Keywords: personalized information service; information retrieval; user interest; content-based filtering; collaborative filtering

1. Introduction

Information retrieval model for consideration has a long history. For decades, many of the search models have been proposed various types have been tested. Model rather than a single integrated search has been proven to be the most effective diversity of approaches and methodologies that have been developed; the field is progressing in two different ways. On the other hand, has been developed theoretical studies of the underlying model, this direction, for example, is represented by various types of logic models and probability models. On the other hand, many experimental studies, including a variant of the model vector space model, have been there. In some cases, the models have been motivated to perform well in theory experience.

Recently, based on language models are applied to the problem of the new approach is successfully retrieved. The basic idea behind the new approach is very simple: the possibility of a query language based on the model, each document's language model estimation, and rank documents found. However,
this new framework, the theoretical statistical much has been expected because of its foundation, speech recognition, work a lot of complementary on language modeling, natural language processing search language model a very simple method is the fact that you have done very well empirically.

And search-based filtering content and specific features of the item the user is built on basic assumptions are required to be able to formulate a query to express the information needs and their interests. However, in some cases, the keywords you can use to accurately describe the benefits, the topic may be difficult to identify the appropriate descriptors and genre. But in other cases, such as electronic commerce, the user simply may be unaware of their own interests at least inattentive. In either case, one is to predict user preferences explicitly formulate a query, without requiring the user would like to recommend items of interest.

Personalized information retrieval systems can help customers to gain orientation in information overload by determining which items are relevant for their interests. One type of information retrieval is content-based filtering. In content-based filtering, items contain words in natural language. Meanings of words in natural language are often ambiguous. The problem of word meaning disambiguation is often decomposed to determining semantic similarity of words. In this paper, the architecture of personalized information retrieval based on user interest is presented. The architecture includes user interface model, user interest model, detecting interest model and update model. It established a user model for personalized information retrieval based on user interest keyword list on client server, which can supply personalized information retrieval service for user with the communications and collaboration of all modules of the architecture.

2. The Personalized Information Retrieval Model Based on User Interest

Personalized based on user interest model and retrieval method of information retrieval in general the biggest difference is received from the user interface, user questions. Analysis will be based on user preferences, while the results matched filter also needs to be based on user preferences. Through the study of mechanisms and reasoning mechanisms, on the one hand to learn the user information, the demand on the preferences, on the other hand, the user needs can be derived, summarized.

The architecture of the personalized information retrieval based on user interest is as Figure 1 shown.

![Figure 1. The architecture of the personalized information retrieval based on user interest](image-url)
Users to personalize information retrieval model, including user interface module, the user feature database, the information within the storage module, and change detection module. Change detection module information of individual users regularly look at the library collections of information in the Internet, whether the document changes, if changes, new results will retrieve the original information on coverage, and the change in the next time the user login to notify the user when the user model. Which user interface module linked together and the user feature database and user characteristic model called the user interface module in turn, learning modules, query and display module, user behavior monitoring module, the user feature database is keyword by the user preference table classification dictionary, the user base and user personalized information library consisting of historical action. The user through the user personalized information retrieval model will be submitted to personalize the query expansion query expression formed by the information collection module submitted to search engines, search engine query expression in the Inter2net by the search and returns the results to the users to personalize information retrieval model. Users to personalize information retrieval model and then filter the results, the filtered results to the user.

2.1. User Interface

Including the query and display module, learning module and user behavior monitoring module, which makes a request to the system and receive system services, it is not just a human-computer interaction interface, is a local program to monitor learning and reasoning function block. When the user after the first successful landing, UIM require users to register, fill in personal background information, select the dictionary from the classification category information of interest to the user preferences as the initial keyword list (user profiles), and can modify the keyword weight or add, delete keywords and their weights to reflect the user's interest in preference to the user preference update is achieved through the learning modules. User behavior monitoring module's main function is to monitor user activities, including keyboard keystrokes and mouse activity, application state analysis and the analysis of local resource users. Learning modules on the user feedback to learn, will learn the results of key words into the user preferences of individual users information retrieval model design and implementation, it also regularly monitoring user behavior and user personalized information module in the content analysis database mining derived user preferences, user preferences and the preferences into the keywords table, complete the list of keywords updated user preferences. While learning modules will be submitted to the derivation of the new search engine preferences, access to the user preference information. Feedback learning modules and adjust the algorithm to determine the semantic query word keyword as the target, the results returned by the user of the document collection action taken, the results were calculated with the user submits the document keyword query semantics between keywords relevance to the user's semantic query word input adjusted accordingly, and to adjust the results of the keyword list into the user preferences, to achieve an accurate representation of user preferences. User model can solve this is to make the query expression to more accurately represent the needs of users, to improve the relevance of the purpose of the user.

2.2. User Characteristics Model

Users to personalize information retrieval model is the client user feature database data platform, consists of four parts: classification dictionary, user preferences keyword list, the user personal information database and user behavior of library history. Users to personalize information retrieval model is the core of the user preferences keyword list, the user preference is to rely on keyword expansion table to support the classification of dictionaries, classification dictionary contains common knowledge in the field of representative keywords, when the user model, a new keyword as a new user preferences user preferences to be extended to the following key words table, the system will follow the
classification of knowledge, knowledge base and structure to the preferences extended to the corresponding tree node, the user in order to achieve preferred extension.

2.3. Change Detecting Model

The working principle of change detection module is built based on the HTTP protocol, the use of HTTP protocol commands, and return the information request header code to be informed of WWW documents on the server where the property values of some documents, such as the modification time of the document, the length of the document such as system needs to be informed of the information stored by the user compare the properties of the document can know whether the changed document information, so you can follow to achieve changes in the document. The structure of change detecting model is as Figure 2 shown.

![Figure 2. The structure of change detecting model](image)

3. Conclusions

Personalized information retrieval systems can help customers to gain orientation in information overload by determining which items are relevant for their interests. One type of information retrieval is content-based filtering. In content-based filtering, items contain words in natural language. Meanings of words in natural language are often ambiguous. The problem of word meaning disambiguation is often decomposed to determining semantic similarity of words. In this paper, we presented the architecture of personalized information retrieval based on user interest. The architecture includes user interface model, user interest model, detecting interest model and update model. It established a user model for personalized information retrieval based on user interest keyword list on client server, which can supply personalized information retrieval service for user with the communications and collaboration of all modules of the architecture.

References


