Available online at www.sciencedirect.com
ScienceDirect
Electronic Commerce Research and Applications 6 (2007) 399–412
Electronic Commerce Research and Applications
www.elsevier.com/locate/ecra

Multi-domain collaborative exploration mechanisms for query expansion in an agent-based filtering framework ☆

Sahin Albayrak, Dragan Milosevic *

DAI-Labor, Technical University Berlin, Salzufer 12, 10587 Berlin, Germany

Received 7 June 2006; received in revised form 19 November 2006; accepted 19 November 2006
Available online 21 December 2006

Abstract

Novice users often do not have enough domain knowledge to create good queries for searching information on-line. To help alleviate the situation, exploration techniques have been used to increase the diversity of the search results so that not only those explicitly asked will be returned, but also those potentially relevant ones will be returned too. Most existing approaches, such as collaborative filtering, do not allow the level of exploration to be controlled. Consequently, the search results can be very different from what is expected. We propose an exploration strategy that performs intelligent query processing by first searching usable old queries, and then utilising them to adapt the current query, with the hope that the adapted query will be more relevant to the user's areas of interest. We applied the proposed strategy to the implementation of a *personal information assistant* (PIA) set up for user evaluation for 3 months. The experimental results showed that the proposed exploration method outperformed collaborative filtering, and mutation and crossover methods by around 25% in terms of the elimination of off-topic results.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Collaborative filtering; Information space exploration; Intelligent query processing; Multi-agent systems; Query adaptation; Recommendation systems

1. Introduction

Without clearly understanding what kind of information is being searched for, the chance that the "right" recommendations will be offered by today's search engines is low. The precise specification of information needs is imperative, but the imperfect nature of human beings unfortunately introduces significant difficulties for ordinary users to create queries that are both compact and sufficiently descriptive. The underlying reasons include the following: (1) Users often need to articulate their thoughts related to their current interests while specifying their queries. (2) They also need to be familiar with query syntax, and this can sometimes be complex when the information sought is very specific. (3) They often need to possess at least some general knowledge about the information to be retrieved. Although there have been a number of online portals that provide information search services, most of these problems hinder users from fully exploring the information space effectively. For example, it makes no sense to search a database of books for an author's family details, as only "bulk" author names can be found there. The lack of exploratory processing of queries, as adopted by most of the today's search engines, is the main reason for queries with too many or too few results returned [2]. Also, the search results returned are often not related to the information needs that users had in mind. Users typically do not consider the fact that not getting the expected results is due to their poorly formed queries. So they may

☆ An earlier version of this paper [1] was presented at the 2005 International Conference on Electronic Commerce in Xi'an, China, and it has been expanded and refined for publication in this special issue of Electronic Commerce Research and Applications. The final version was edited by William Cheung and Rob Kauffman.

* Corresponding author. Tel.: +49 30 31425318; fax: +49 30 31421799.
E-mail addresses: sahin@dai-lab.de (S. Albayrak), dragan@dai-lab.de (D. Milosevic).

1567-4223/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.elerap.2006.11.006

lose confidence in a particular search engine, and may switch to another one that can better satisfy their information needs.

Since the users' habits regarding the creation of queries cannot easily be influenced, intelligent processing techniques delivered by a filtering engine are necessary. One such technique is called *intelligent query answering* [3], where the underlying intent of a query is analysed in order not to provide exactly what is being asked for, but also to explore potentially relevant information areas. The user who is looking for a nice holiday destination might also want to get information about car rental, even though the user has not asked for that. An intelligent query system can deduce that. How? By knowing that this extra information is the most frequently requested at the same time for a particular holiday offer, for example. Then, the provided results will be much more diverse, and have a better chance to satisfy the stated, as well as the implicit needs of the user who created the query.

Even though providing extra information may be useful, providing too much will lower the precision of the results and may have a negative effect. A user can be overwhelmed by irrelevant results and decide not to use the search engine in the future due to the bad experience. Thus, the level of exploration should be carefully controlled. Conventional exploration techniques like collaborative filtering normally ignore the extent of exploration, which causes the delivery of results unrelated to the actual information needs. Although recommending items that are liked by similar users sounds logical, these items will have a low usability if they are too far away from the context of the query. For example, offering weather information about cities could be favoured when sightseeing recommendations are requested but not when we are comparing prices for travel packages. The deployed exploration mechanisms should be able to control the automatic adaptation of the specified information needs so they are within users' preferences.

A promising technique for performing controlled exploration is *automatic transformation* of an actual query based on past queries. Novice users can benefit from the past queries posted by like-minded experts. To do so, one may need to first identify past queries that are related to the currently posted one, and then utilise the extended context formed due to the past queries to achieved controlled exploration. For example, a query about tools for designing multi-agent systems might lack important keywords which might be found in other queries that have been previously posted by domain experts.

Such an application of query adaptation to achieve exploration is fully transparent in the sense that a query is modified without user awareness. In order to integrate exploration techniques in a multi-agent framework that contains different filtering agents, this transparency will be essential. In the remaining sections, after critically reviewing the drawbacks of other attempts to provide exploration, one scenario will be used to illustrate the exploration challenge. Our core contribution is presented in Section 4. It will show how similar past queries can be found. It also will show how they can be used to carefully adapt the current one by both adding new attributes that might be important but missing, as well as adapting the existing attributes so they have more realistic values. We conclude with implementation details and experimental results.

2. Related work

The importance of helping a user to explore hidden information needs has long been recognized and there have been a number of related projects reported in the literature for addressing the issue. To the authors' knowledge, there are four main directions for achieving exploration which are presented in the following subsections.

2.1. Collaborative filtering

An obvious way to achieve effective exploration is to use filtering strategies that are exploratory in nature. One such filtering technique is *collaborative filtering* or *social filtering* [4–8], which are based on the assumption that items that are liked by similar users are good to recommend. A well known problem with collaborative filtering is its poor explanatory power [9]. For example, people will hardly be satisfied with recommendations on holiday destinations when the only explanation for why they have been presented is that the destinations are liked by similar users. Users may be similar or different in other aspects. Another problem with pure collaborative filtering lies in its difficulty to control exploration. Even though a user shows interest in articles from a particular domain, collaborative filtering can find recommendations in completely unrelated domains just because similar users like them. If one only treats collaborative filtering as a "black box", these problems can hardly be solved no matter how the collaborative filtering techniques are applied. One way to overcome these drawbacks is to use collaborative filtering not as single, but only as an auxiliary exploratory engine, as we will show in this paper.

2.2. Skyline

For efficient processing of queries where a user has not specified the importance levels of attributes, a *skyline technique* [10] can be applied. A *skyline* is a collection of items that are not dominated by other items. An item is said to be *dominated* when there is another item with better attribute values. For example, when "hotel price" and "distance from the beach" are the two key attributes, it is obvious that smaller attribute values will be preferred. But a particular hotel at a given room price will not be dominated by others when it is the closest to the beach.

For exploration, a skyline is used so that only items that are members of a skyline should be recommended. The parts of a skyline correspond to the different importance levels of the selected attributes for which the skyline has

متن کامل مقاله ← دریافت فوری

ISIArticles
مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات