Information Sciences 314 (2015) 55–76

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

On personalizing Web search using social network analysis

Omair Shafiq, Reda Alhajj *, John G. Rokne

Department of Computer Science, University of Calgary, Calgary, Alberta, Canada

ARTICLE INFO

Article history:
Received 23 July 2014
Received in revised form 9 February 2015
Accepted 14 February 2015
Available online 31 March 2015

Keywords:
Personalization
Social networks
Search engine
Optimization
Ranking

ABSTRACT

Most of the existing Web search solutions are built for satisfying broad set of users regardless whether naïve or professionals. Further, with the emergence of high speed internet applications and advanced Web 2.0 based Rich Internet Applications (i.e. blogs, wikis, etc.), it has become much easier for users to publish data over the Web. This brings a challenge for Web search solutions to let individual users find the right information as per their preferences. Different users of the Web may have different preferences. Search results for the same query from different users may differ in priority for individual users. In this paper, we describe our approach of enabling personalized Web search for users based on their preferences. It is a challenge in itself to have the preferences of the users known to and considered by search engines. We have designed and developed our unique approach of finding the preferences of users from the relevant parts of their social networks and communities. We believe that the information related to the queries posed by users may have strong correlation with relevant information in their social networks. In order to find out the personal interests and the social-contexts, we find out (1) activities of users in their social-networks, and (2) relevant information from user's social networks, based on our proposed trust and relevance matrices. We have developed a mechanism that extracts information from a user's social network and uses it to re-rank the results from a search engine. We have also discussed the implementation and evaluation of our proposed solution by emphasizing how it improves the Web search.

1. Introduction

It has become easier for internet users to publish data over the Web by using advanced Web 2.0 based interactive applications (e.g., wikis, blogging, etc.). This leads to a rapid increase in the amount data exchanged over the Web. Such information and data flooding makes it harder for a Web user to find out the target information. Web search engines (e.g., Google, Yahoo, etc.) return thousands of search results for a particular search query. Mostly, they act as "one size fits all" by providing same search results for the same query by different users. However, in real life, different users may have preferences. Therefore, we find out a need for the personalization of the search results reported by Web search engines for different users based on their preferences.

Web search is a facility that allows users to look for target information over the Web. However, with the given huge amount of information available over the Web, it is almost impossible for a user to go through all the content of all the search results manually, and to find out the required information. Some of the search engines attempt to provide manual

* Corresponding author.
E-mail address: alhajj@ucalgary.ca (R. Alhajj).

http://dx.doi.org/10.1016/j.ins.2015.02.029
0020-0255/

mechanisms to rank the search results, higher or lower. However, for each and every search result, it is almost impossible for a user to adjust the ranking manually. This calls for having automated mechanisms to allow re-ranking search results based on the preferences of the users. However, it is a challenge in itself to decide on how to automatically find the preferences of users. Different mechanisms have been investigated and built to collect the preferences of the users. Further, there is considerable effort to enable community and social network aware [1], user-terminology-aware [2] and group-aware [23,5] search engines that may help in reporting results closer to user's interest. Search engines like Google seems to work out social network-aware search mechanisms [24,13]. We argue that this may become the trend in the near future and hence our proposed solution of social-network aware personalized Web search for users is an opportunity to timely contribute to this emerging field of research that may impact the search process.

Our proposed solution acts as a user-oriented front-end to Web search engines; it re-ranks Web search results for the users based on their preferences. There are various ways to find out preferences of the users, e.g., location. But this information is most of the time not enough or limited to return the required search results for the right user at the right time. Different users searching over the Web for the same type of information may have different preferences. For example, an executive going for a business meeting would probably be interested in a hotel closer to the meeting place; however a back-packer searching for accommodation in the same city might want to search for a cheaper hotel, regardless of the location. We already have achieved some progress by investigating the related work in Section 2, and our initial results have been published in [26].

In order to find out preferences of the users, we propose to exploit users' social networks, activities of the users and their communities, i.e., friends in their social networks. This information may greatly help in bringing word-of-mouth recommendations for search results, from trusted friends of the users. Every user has certain set of activities within the social network, e.g., postings, public profile information, tagging, blogging, etc. Most of the times, friends of a user in a social network have to some extent similar interests and therefore could help in finding out the required information that the user may need. The information about the related search query from friends in a social network, or activities of a user in a social network could help in finding and prioritizing the relevant information over the Web. Such kind of community-aware personalized Web search utility could further be used in many different ways and in many applications involving Web intelligence and business intelligence, including summarizing and caching important search results, as well as collaboration techniques.

We have the following hypothesis: How to use relevant information from user's community to personalize search results? This includes where to find the community information of the users and how to model it? How to trust a social network to find relevant information? Finally, how to re-rank the search results based on the information extracted from the user's community? The first objective is to find the important information from a user's community. This will involve developing mechanisms to find activities of users within a social network as well as their friends/communities. This information could be found in different possible ways, e.g., user's blogging, posting information and interacting with other users on publicly available portals. Second, there might be cases where community information might become irrelevant to the user. For example, a user may be interested in sports (particularly football), however, there may be couple of friends in the user's social network relevant to football as well as cricket. In order to find out the relevant information, it is important to include the information obtained from friends that are relevant to football only, rather than cricket.

There may be different kinds of friends and communities in a social network based on whom we can find the relevant information. Same as in real-life, a user may have different trust metrics for different friends in a social network. This means, information collected from a more trusted friend is normally of more importance than that related to friends of lesser trust. Third, we will find how to use this information to re-rank or prioritize the relevant search results based on the preferences collected from the trusted and relevant parts of the social network. This way, we will be able to utilize social networks of the users to find out relevant information over the Web. Details about our proposed solution as well as evaluation results are given in following sub-sections to discuss the degree of improvement in relevance in Web search for users.

The rest of the paper is organized as follows. Section 2 covers the related work and discusses different pros and cons. Section 3 describes our proposed solution of community-aware personalized Web search and discusses how to model the information based on user's individual preferences, trusted as well as relevant friends of the users in their social network. Section 3 describes the social-network aware strategies that are used to collect the information from the trusted and relevant parts of a social network. Algorithms for re-ranking of Web search results based on the information obtained from social networks are also presented. Section 4 covers implementation and evaluation details followed by conclusions.

2. Related work

Personalized Web search has received considerable attention and a number of approaches have been developed for personalized Web search based on contextual information, i.e., interests and preferences of users [30,27,17,22,31]. A number of approaches are based on modeling and collecting contextual information. For example, context-aware search methods [25] use a special ontology which is constructed from the table of contents of books to help formulating queries for efficiency and to refine query search results that are relevant to user's interest. Location-awareness is another form of context-awareness; it is used to filter out search results based on the current location of the user.

There are different ways in which clients access services over the Internet. With the advancements of requirements and the complexity in the consumer applications, consumers expect services to be aware of their current environment and