

Enhancing Web Search with Semantic Identification of User Preferences

Naglaa Fathy¹, Nagwa Badr¹, Mohamed Hashem¹, Tarek F. Gharib^{1, 2}

¹Faculty of Computer and Information Sciences, Ain-Shams University, Cairo, Egypt

²Faculty of Computing and Information Technology, King Abdul-Aziz University, KSA

Abstract

Personalized web search is able to satisfy individual's information needs by modeling long-term and short-term user interests based on user actions, browsed documents or past queries and incorporate these in the search process. In this paper, we propose a personalized search approach which models the user search preferences in an ontological user profile and semantically compares this model against user current query context to re-rank search results. Our user profile is based on the predefined ontology Open Directory Project (ODP) so that after a user's search, relevant web pages are classified into topics in the ontology using semantic and cosine similarity measures. Moreover, interest scores are assigned to topics based on the user's ongoing behavior. Our experiments show that re-ranking based on the semantic evidence of the updated user profile efficiently satisfies user information needs with the most relevant results being brought on to the top of the returned results.

Keywords Search Personalization, user profile, ODP, re-rank, semantic similarity

1. Introduction

With the massive growth of the information available on the World Wide Web, it becomes more difficult for search engines to provide the desired results due to the ambiguity of user needs. Different users have different goals including research, entertainment, finding new jobs or purchasing items. Moreover, current search engines generally process search queries without considering user interests or contexts in which users submit their queries. This can be explained with the search query "Racetrack" One user may need information about Racetrack game. Another user might be looking for results of racetrack playa, while other users might be searching for the memory device racetrack. Obviously, different users would prefer different answers, However, Users surfing the Web in search of relevant information have less time and patience to formulate queries and filter the results returned. Therefore, most web search engines prefer to provide a large set of search results while the users have to determine what is relevant and what is not.

In order to address this problem, recent researches proposed Personalized search which

aims to provide users with results that are relevant to their interests by considering user's search history in the retrieval process. It is challenging to identify and exploit the user profile so that search quality could be improved. In particular, personalized search observes all web pages visited by the user, together with the user's search behavior to build a user model which is then used to re rank the top search results returned by a non-personalized web search engine.

The easiest way to get more information about the user and incorporate it in the search process is to ask the user explicitly about his interests and preferences and save them in the user profile. The main disadvantage of this method is that users are reluctant to spend time to provide their intensions before each search. Moreover, it is very difficult for users to define their own interests accurately. Another complex method is based on implicitly observing user's browsing activities and adapting the system according to them. Information stored in user profile can be used to disambiguate or to infer user's query context. Studies in personalized search include [2] which provided the searcher with different search topics and monitored clicked search results so as to learn user's current interests and re-order web search result accordingly. Another approach in [1] models user interests as a vector of weighted terms from visited URLs, and apply a snippet scoring method to re rank search results. In [10] the user profile is composed of each submitted query with its clicked URLs and their corresponding topics, then re-rank is achieved by boosting results with similar topics to topics of queries in the profile that are relevant to current query.

In [3][4][5] [6] the user profile is created based on reference ontology to link information extracted from visited web pages, whereas re ranking is based on a numerical estimate of the results' relevance to the user's profile.

In this paper, we propose a personalized ranking approach for web search results in two main modules. The first module includes capturing the user's preferences and interests from past searches in an ontological user profile which is defined by assigning interest scores to existing