

Implementation of Personalized mobile search engine based on Ontology

Ms.Namrata G Kharate

(ME-II, Computer)

Prof.S.A.Bhavsar(Guide)

Department of Computer Engineering,

Matoshri College of Engineering and Research Center, Nashik

University of Pune

Email: namratakharate1@gmail.com

Abstract— Mobile search engine is a program that searches for and identifies items in a database that correspond to keywords or characters specified by the user, used especially for finding particular sites on the World Wide Web. A major problem in mobile search is that the interactions between the users and search engines are limited by the small form factors of the mobile devices. As a result, mobile users tend to submit shorter, hence, more ambiguous queries compared to their web search counterparts. In order to return highly relevant results to the users, mobile search engines must be able to profile the users' interests and personalize the search results according to the users' profiles. A personalized mobile search engine (PMSE) that

captures the users' preferences in the form of concepts by mining their click-through data. Due to the importance of location information in mobile search, PMSE classifies these concepts into content concepts and location concepts. The user preferences are organized in an ontology-based, multifaceted user profile, which are used to adapt a personalized ranking function for rank adaptation of future search results. To characterize the diversity of the concepts associated with a query and their relevance to the users' need, four entropies are introduced to balance the weights between the content and location facets. Based on the client-server model, a detailed architecture and design for implementation of PMSE is also presented. In the implemented system, the client collects and stores locally the click-through data to protect privacy, whereas heavy tasks such as concept extraction, training, and reranking are performed at the PMSE server. User gets the information depending upon query and nearby location.

Keyword: Clickthrough data, concept, location search, mobile search engine, ontology, personalization, user profiling.

I. INTRODUCTION

Nowadays people use internet widely for many personal uses. Search engines like Google, Yahoo, etc. can not be personalized. Leung et al. developed a search engine personalization method based on users' concept preferences and showed that it is

more effective than methods that are based on page preferences [2]. In mobile search, the interaction between users and mobile devices is constrained by the small form factors of the mobile devices. To reduce the amount of user's interactions with the search interface, an important requirement for mobile search engine is to be able to understand the users' needs, and deliver highly relevant information to the users, to overcome the drawback of existing systems. Personalized mobile search engine (PMSE) came into existence which classifies into two concepts: content concept and location concept. By capturing the users' interests in user profiles, a personalized search middleware is able to adapt the search results obtained from general search engines to the users' preferences through personalized

reranking of the search results. In the personalization process, user profiles play a key role in reranking search results and thus need to be trained constantly based on the user's search activities. Several personalization techniques have been proposed to model users' content preferences via analysis of users' clicking and browsing behaviors [5], [9], [12], [14].

In this work, we recognize the importance of location information in mobile search and propose to incorporate the user's location preferences in addition to content preferences in user profiles. Search engines are developed to help the users to search for what they are looking for. The information that the users are looking for is found in the web data repository. Search engines work with the help of predefined automated software programs (spiders or bots and are used for crawling,) which is going through the entire website (all the pages) and then records down the content in the form of databases, which is known as indexes.

In most commercial search engines, the user gets the same result returned for the same query regardless of their real interest. Since queries given by the user are short and ambiguous, search engines are unable to express the user's precise needs.

To the best of our knowledge, our paper is the first to

propose a personalization framework that utilizes a user's content preferences and location preferences as well as the GPS locations in personalizing search results.

The main contributions of this paper are as follows:

- . This paper studies the unique characteristics of content and location concepts

- . The proposed personalized mobile search engine is an innovative approach for personalizing web search results.

- By mining content and location concepts for user profiling.

- . PMSE incorporates a user's physical locations in the personalization process.

- . Privacy preservation is a challenging issue in PMSE, where users send their user profiles along with queries to the PMSE server to obtain personalized search results.

II .DETAILED PROBLEM DEFINITION

To design a system as a PMSE for retrieving relevant information depending upon users supplied query. In this project, approach to personalizing web search in a mobile environment. As a case study, Android phone as the mobile platform to implement work. Main goal is to identify user s interests based on the web pages he visits, and deliver personalized web search results by utilizing the identified user interests. We learn and maintain implicitly an ontological profile of user s interests through passive observation of the user s click stream. The user s interest profile is stored locally on his mobile device and updated with every web page visit. Personalization is achieved by re-ranking standard web search results using the user s interest profile. Separate concepts namely location concepts and content concepts is used to recognize the importance of location information in mobile search. Location preferences others the mobile search engine an additional dimension for capturing a users interest and an opportunity to enhance search quality for users. The implemented system is capable of combining a users GPS locations and location preferences into the personalization process.

III. RELATED WORK

Most of the previous work assumed that all concepts are of the same type. Observing the need for different types of concepts, we present in this paper a personalized mobile search engine (PMSE) which represents different types of concepts in different ontologies. In particular, recognizing the importance of location information in mobile search, we separate concepts into location concepts and content concepts. To incorporate context information revealed by user mobility, we also take into account the visited physical locations of users in the PMSE. Since this information can be conveniently obtained by GPS devices, it is hence referred to as GPS locations. GPS locations play an important role in mobile web search. Many existing personalized web search systems [2], [3], [6], [7] are based clickthrough Personalization is necessary to maintain a more suitable result. Among many approaches, the fuzzy concept network according to a user profile can characterize a user's subjective interest appropriately. The paper proposes another search engine that utilizes the fuzzy concept network to personalize the outputs from a link-based search technique. Depending on a user profile, the fuzzy concept network rearranges five outputs of the link-based search engine, and the system presents a personalized

superior quality result. Experimental observations with the three subjects show that the system proposed searches not only appropriate but also personalized Web pages on a user's preference.

Thorsten Joachims. This paper presented an approach to mining locales of WWW search engines with the goal of improving their retrieval performance automatically. The key insight is that such clickthrough data can provide training data in the form of relative preferences. Based on a new formulation of the learning problem in information retrieval, this paper derives an algorithm for learning a ranking function. Taking a Support Vector approach, the resulting training problem is tractable even for large numbers of queries and large numbers of features. Experimental results show that the algorithm performs well in practice, successfully adapting the retrieval function of a meta-search engine to the preferences of a group of users. Q. Tan, X. Chai, W. Ng, and D. Leel, propose a new algorithm called Ranking SVM in a Co-training Framework (RSCF), which is able to improve the retrieval quality of search result by learning from clickthrough data. The RSCF algorithm relies only on implicit user feedback and does not add any burden to the users during the process of web searching. A key insight of utilizing clickthrough data is that the data only provides a small set of training data in the form of relative preferences. In RSCF, we resolve this problem by augmenting the training data set using a co-training framework.

We show by an extensive set of online experiments that RSCF improved the retrieval quality. There are still a lot of issues that de- serve further study in order to improve the performance of RSCF. Later on, Chen et al. [8] studied the problem of efficient query processing in location-based search systems. A query is assigned with a query footprint that specifies the geographical area of interest to the user. Several algorithms are employed to rank the search results as a combination of a textual and a geographic score.

Personalized Web search with location preferences is recommended by K.W.-T. Leung et al., [1]. Since the amount of Web information develops at very fast pace, search engines must be capable of retrieving information based on the user's preference. In this paper, the authors recommended a novel web search personalization technique that recognizes the user's interests and preferences with the help of concepts by mining search out- puts and their clickthrough. Because of the significant role played by the location information in mobile search, the author planned to divide concepts into content concepts and location concepts, and categorize them into ontologies to generate an ontology based, multi-facet (OMF) profile to exactly recognize the user's content and location interests and hence enhance the search accuracy.

Furthermore, identifying the information that different users and queries may have different importance on content and location information, this paper introduces the idea of content and location entropies to determine the amount of content and location information linked with a query, and click content and location entropies to calculate how much the user is concerned in the content and location information in the results. As a result, the authors proposed to define personalization efficiency based on the entropies and utilize it to equal the weights among the content and location facets. Finally, based on the resultant ontologies and personalization efficiency, the author trained an SVM to adapt a personalized ranking function for re-ranking of

future search. Many experiments were carried out to evaluate the precision produced by this OMF profiles and that of baseline method. Experimental out- come shows that OMF enhances the precision considerably compared to the baseline

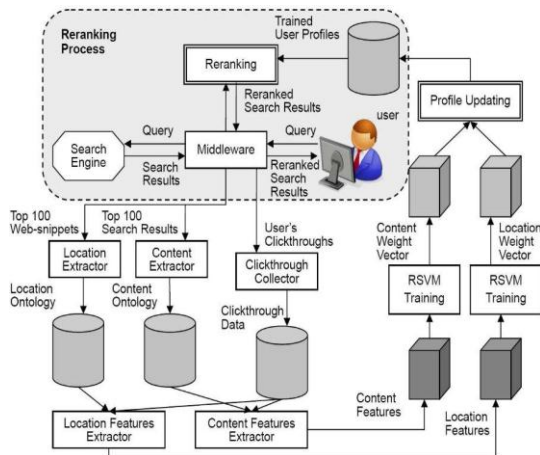


Fig.1. The general process flow of PMSE

III. Proposed system

Fig. 1 shows PMSE’s client-server architecture, which meets three important requirements. First, computation-intensive tasks, such as RSVM training, should be handled by the PMSE server due to the limited computational power on mobile devices. Second, data transmission between client and server should be minimized to ensure fast and efficient processing of the search. Third, clickthrough data, representing precise user preferences on the search results, should be stored on the PMSE clients in order to preserve user privacy.

In the PMSE’s client-server architecture, PMSE clients are responsible for storing the user. Click through and the ontologies derived from the PMSE server. Simple tasks, such as updating clickthrough and ontologies, creating feature vectors, and displaying reranked search results are handled by the PMSE clients with limited computational power. On the other hand, heavy tasks, such as RSVM training and reranking of search results, are handled by the PMSE server. Moreover, in order to minimize the data transmission between client and server, the PMSE client would only need to submit a query together with the feature vectors to the PMSE server, and the server would automatically return a set of reranked search results according to the preferences stated in the feature vectors. The data transmission cost is minimized, because only the essential data (i.e., query, feature vectors, ontologies and search results) are transmitted between client and server during the personalization process. PMSE’s design addressed he issues: 1) limited computational power on mobile devices, and 2) data transmission minimization.

IV. RESULTS AND DISCUSSION

One of the most important evaluation criteria is the Response time.

The response time of an algorithm is the time taken to perturb the requests. Response time is a performance measure. Number of users requesting from different queries from different locations. following table shows the time

need for getting precise response from server for serving different number of bunch of users.

Sr NO	Total no. of User Request	Average Response Time
1	10	62
2	20	41
3	30	36
4	40	34

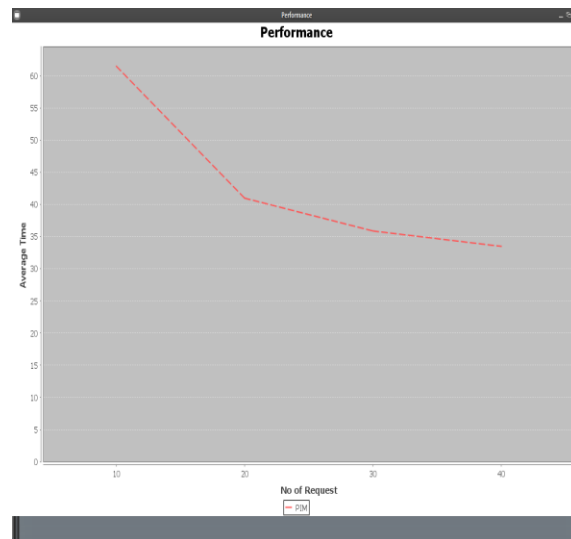


Fig 2: Graph

Figure 2 shows the graph for response time for various numbers of requests, initially response time is high after that as number of request increases, response time minimizes.

V.CONCLUSION

We proposed PMSE to extract and learn users content and location preferences based on the user’s clickthrough. To adapt to the user mobility, we incorporated the users GPS locations in the personalization process. We observed that GPS locations help to improve retrieval effectiveness, especially for location queries. the system extract the location concept and location concept from users click through data. The system has been implemented with spyNB mining algorithms. To adapt to the user mobility, the system incorporated the users GPS locations in the personalization process. It is observed that GPS locations help to improve retrieval effectiveness, especially for location queries.

REFERENCES

1. K.W.-T. Leung, D.L. Lee, and W.-C. Lee, PersonalizedWeb Search with Location Preferences, Proc. IEEE Intl Conf. DataMining (ICDE), 2010.
2. K.W.-T. Leung, W. Ng, and D.L. Lee, Personalized Concept-Based Clustering of Search Engine Queries, IEEE Trans. Knowledge an Data Eng., vol. 20, no. 11, pp. 1505-1518, Nov. 2008.

3. S. Yokoji, Kokono Search: A Location Based Search Engine, Proc. Intl Conf. World Wide Web (WWW), 2001.
4. E. Agichtein, E. Brill, S. Dumais, and R. Ragno, Learning User Interaction Models for Predicting Web Search Result Preferences, Proc. Ann. Intl ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2006.
5. T. Joachims, Optimizing Search Engines Using Clickthrough Data, Proc. ACM SIGKDD Intl Conf. Knowledge Discovery and Data Mining, 2002.
6. W. Ng, L. Deng, and D.L. Lee, Mining User Preference Using Spy Voting for Search Engine Personalization, ACM Trans. Internet Technology, vol. 7, no. 4, article 19, 2007.
7. Q. Tan, X. Chai, W. Ng, and D. Lee, Applying Co-Training to Clickthrough Data for Search Engine Adaptation, Proc. Intl Conf. Database Systems for Advanced Applications (DASFAA), 2004.
8. Y.-Y. Chen, T. Suel, and A. Markowetz, Efficient Query Processing in Geographic Web Search Engines, Proc. Intl ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2006.