Location and Content Based Mobile Search Engine

Y. Sam Josuva[1], V. Belmer Gladson[2]

[1]Assistant Professor, [2] Research Scholar

Department of Computer Science & Engg, Manonmaniam Sundaranar University, Tamilnadu,

Abstract - A location and content based mobile search engine (LCMSE) is a personalized form of search engine where users' preferences were captured by analyzing their behavioral data and by mining through their click through information. It can further be broken down into content based and location based where the system gathers more information about the user using their location (GPS) data. The system follows a client-server model and the client saves the users data locally to protect privacy and the content extraction and ranking is done on the server side to ease the load on the client and the need for updating and refreshing the server is eliminated and also to make the process faster and seamless. The prototype was done on the Android platform and system has got significant improvement over the traditional model. 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.

Keywords - Search Engine, Image, Location, Content.

I. INTRODUCTION

The proposed Location and Content Based Mobile Search Engine is an innovative approach for personalizing web search results. By mining content and location concepts for user profiling, it utilizes both the content and location preferences to personalize search results for a user. A Location and Content Based Mobile Search Engine (LCMSE) that captures the user's preferences in the form of concepts by mining their click through data. Due to the importance of location information in mobile search, LCMSE classifies these concepts into content concepts and location concepts. In addition, user's locations (positioned by GPS) are used to supplement the location concepts in LCMSE.[5] 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 query and their relevance's to the user's need, four entropies are introduced to balance the weights between the content and location facets. Based on the client-server model, the project presents a detailed architecture and design for implementation of LCMSE. In our design, the client collects and stores locally the click through data to protect privacy, whereas heavy tasks such as concept extraction, training, and re ranking are performed at the LCMSE server. Moreover, the system address the privacy issue by restricting the information in the user profile exposed to the LCMSE server with two privacy parameters. The prototype LCMSE on the Google Android platform. Experimental results show that LCMSE significantly improves the precision comparing to the baseline.

Query processing is a major bottleneck in standard web search engines, and the main reason for the thousands of machines used by the major engines.[1] Geographic search engine query processing is different in that it requires a combination of text and spatial data processing techniques. They propose several algorithms for efficient query processing in geographic search engines, integrate them into an existing web search query processor, and evaluate them on large sets of real data and query traces. a new approach to mining user's preferences on the search results from click through data and using the discovered preferences to adapt the search engine's ranking function for improving search quality.[2] We develop a new preference mining technique called SpyNB, which is based on the practical assumption that the search results clicked on by the user reject the user's preferences, but it does not draw any conclusions about the results that the user did not click on.

Essentially, the RSCF algorithm takes the click through data containing the items in the search result that have been clicked on by a user as an input, and generates adaptive rankers as an output[3]. By analyzing the click through data, RSCF ¯rest categorizes the data as the labelled data set, which contains the items that have been scanned already, and the unlabelled data set, which contains the items that have not yet been scanned. The labelled data is then augmented with unlabelled data to obtain larger data set for training the rankers.

A scalable way for users to automatically build rich user profiles. These profiles summarize user's interests into a hierarchical organization according to specific interests.[4] Two parameters for specifying privacy requirements are proposed to help the user to choose the contented degree of detail of the profile information that is exposed to the search engine. Experiments showed that the user profile improved search quality when compared to standard MSN rankings.

II. EXISTING SYSTEM

Most existing location-based search systems, such as, require users to manually define their location preferences (with latitude-longitude pairs or text form), or to manually prepare a set of location sensitive topics. Existing works on personalization do not address the issues of privacy

preservation. The number of users and queries in the experiments are small. This means that the results from the experiments cannot be constructed as representative in diverse situations. Since users are given with predefined queries and topical interests, they have to synthesize their information needs from the given queries and topical interests and conduct their searches correspondingly. Thus, their search behaviors in the experiments may be quite different from what they might have exhibited when they attempt to resolve real-life information needs.

III. PROPOSED SYSTEM

LCMSE profiles both of the user's content and location preferences in the ontology based user profiles, which are automatically learned from the click through and GPS data without requiring extra efforts from the user. The system proposes and implements a new and realistic design for LCMSE. To train the user profiles quickly and efficiently, our design forwards user requests to the LCMSE server to handle the training and reranking processes. LCMSE addresses this issue by controlling the amount of information in the client's user profile being exposed to the LCMSE server using two privacy parameters, which can control privacy smoothly, while maintaining good ranking quality. The proposed location and content based mobile search engine is an innovative approach for personalizing web search results. By mining content and location concepts for user profiling, it utilizes both the content and location preferences to personalize search results for a user. It studies the unique characteristics of content and location concepts, and provides a coherent strategy using client-server architecture to integrate them into a uniform solution for the mobile environment.LCMSE incorporates a user's physical locations in the personalization process. The system conduct experiments to study the influence of a user's GPS locations in personalization. The results show that GPS locations help improve retrieval effectiveness for location queries.

Click through collection at LCMSE client.

The ontologies returned from the LCMSE server contain the concept space that models the relationships between the concepts extracted from the search results. They are stored in the ontology database on the client. when the user clicks on a search result, the click through data together with the associated content and location concepts are stored in the click through database on the client. The click through are stored on the LCMSE clients, so the LCMSE server does not know the exact set of documents that the user has clicked on. This design allows user privacy to be preserved in certain degree.

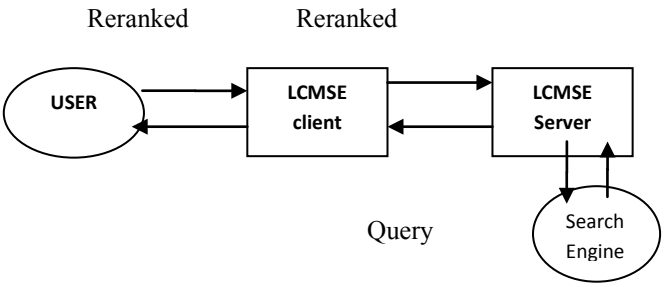
When a user submits a query on the LCMSE client the query forwarded to the LCMSE server .It obtains the search results from the back-end search engine .The content and location concepts are extracted from the search results and organized into ontologies to capture the relationships between the concepts.

After clicking on

Ontology+ Search Result

Re-ranking the search results at LCMSE server:

The search results are then re-ranked according to the weight vectors obtained from the RSVM training. Finally, the re-ranked results and the extracted ontologies for the personalization of future queries are returned to the client.



User Interest Profiling:

LCMSE uses "concepts" to model the interests and preferences of a user. The concepts are further classified into two different types, namely, content concepts and location concepts. The ontologies indicate a possible concept space arising from a user's queries, which are maintained along with the click through data for future preference adaptation.

Diversity and Concept Entropy:

LCMSE consists of a content facet and a location facet. In order to seamlessly integrate the preferences in these two facets into one coherent personalization framework. In this, weights of content preference and location preference based on their effectiveness in the personalization process.

The notion of personalization effectiveness is derived based on the diversity of the content and location information in the search results.

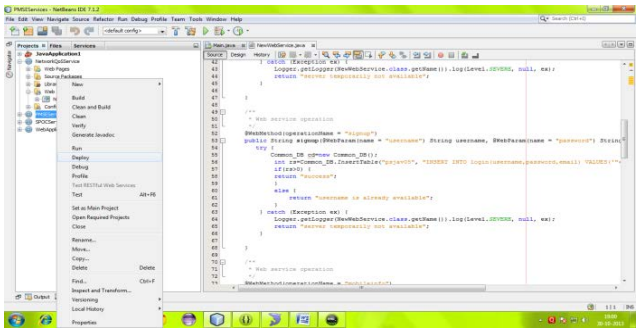
IV. CONCLUSION

We proposed LCMSE to extract and learn a user's content and location preferences based on the user's click through. To adapt to the user mobility, we incorporated the user's GPS locations in the personalization process. We observed that GPS locations help to improve retrieval effectiveness especially for location queries. System also proposed two privacy parameters, min Distance and expiration, to address privacy issues in LCMSE by allowing users to control the amount of personal information exposed to the LCMSE server. The privacy parameters facilitate smooth control of privacy exposure while maintaining good ranking quality. In our design, 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 LCMSE server. Moreover, we address the privacy issue by restricting the information in the user profile exposed to the LCMSE server with two privacy parameters. The prototype of LCMSE on the Google Android platform. Experimental results show that LCMSE significantly improves the precision comparing to the baseline.

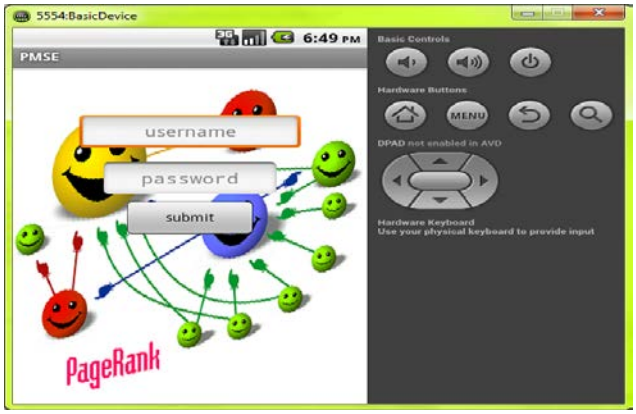
V. FUTURE ENHANCEMENT

Future work, we will investigate methods to exploit regular travel patterns and query patterns from the GPS and click through data to further enhance the personalization effectiveness of LCMSE. To maintain the good efficiency to the user preferred location search. We can use in the Tablets and the other future technologies.

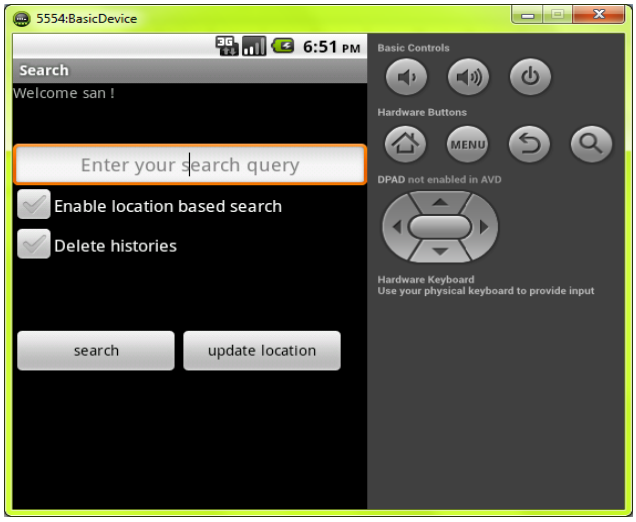
Snap Shots:

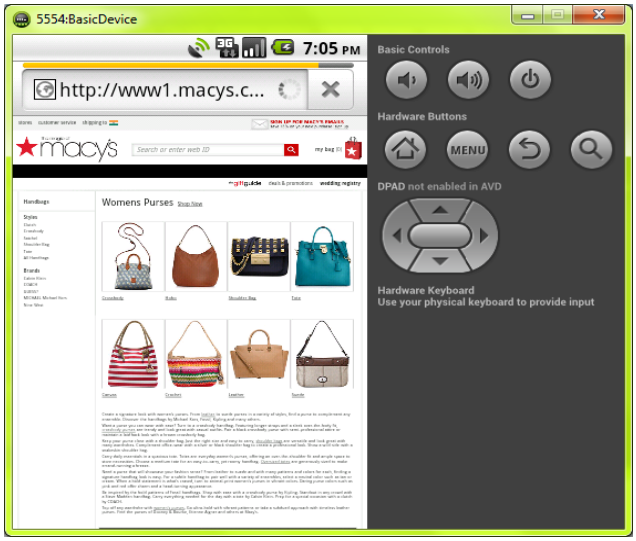


Deploy the LCMSE server in webserver using netbeans IDE



Run client app





REFERENCES

[1] Efficient Query Processing in Geographic Web Search Engines Yen-Yu Chen , Torsten Suel SIGMOD 2006, June 27–29, 2006, Chicago, Illinois, USA

[2] Minig user preference using spy voting for search engine personalization Wilfred Ng, Lin Deng and Dik Lun Lee

ACM Transactions on Internet Technologies, Vol. 7, No. 3, August 2007

[3] Applying co-training to click through data for search engine adaptation Qingzhao Tan Xiaoyong Chai Wilfred Ng , Research Grant Council of Hong Kong,Grant No HKUST6079/01E, DAG01/02.EG05, and HKUST6185/02E.

[4] Privacy enhancing personalized web search Kenneth wai-ting Leung Dik Lun Lee , Wang-Chien Lee ICDE Conference 2010 , 978-1-4244-5446-4/10/$26.00 □□ 2010 IEEE

[5] A Personalized Mobile Search Engine Kenneth Wai –ting Leung ; Dik Lun Lee; Wang –Chien Lee IEEE Transactions on Knowledge and Data Engineering (Volume: 25, Issue: 4, April 2013)

[6] E. Agichtein, E. Brill, and S. Dumais, “Improving Web Search Ranking by Incorporating User Behaviour Information,” Proc. 29th Ann. Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2006.

[7] E. Agichtein, E. Brill, S. Dumais, and R. Ragno, “Learning User Interaction Models for Predicting Web Search Result Preferences,” Proc. Ann. Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2006.