International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

PMSE: A Personalized Mobile Search Engine Using Content and Location Concept

Deepa S[1], Arun P S[2]

[1,2] Department of Computer Science & Engineering, Sree Buddha College of Engineering, Alappuzha, Kerala, India

Abstract: ***Now a days there is 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 user's interests and personalize the search results according to the user's profiles. In this paper, A Personalized Mobile Search Engine (PMSE) using content and location concept, that captures user's 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 user profile, which are used to adapt a personalized ranking function for rank adaptation of future search results.***

Keywords: Click through data, concept, location search, mobile search engine, personalization

1. Introduction

Mobile devices have evolved to provide bigger full-color screens, enhanced processing power and faster and permanent broadband Internet connections. These technologies have brought the World Wide Web to mobile devices introducing new requirements and expectations. But the most majority of web sites and search engines are usually designed with desktop computers in mind. For that reason, current mobile search experience is far from satisfactory .Search engine analysts, being aware of this problem, have designed mobile-oriented views to provide the same service from a smaller interface.

Mobile Web Search introduces new challenges not present in traditional web search. Users normally own modern cell phones which allows them to be permanently online anywhere, anytime. A typical mobile web search scenario consists of a user outdoors with an information need. At this point he takes his phone and uses a web search engine to find an answer to a query. Furthermore, he is probably doing something else at the same time, like walking or talking to a friend. In such situation the user needs a short, fast but also accurate answer to his query. Most of the current web information is developed based on HTML. Semantic assigns a meaning to a document and the Semantic Web is a high end automated intellectual technology that allows not only humans but machines to understand information. In order for a machine to take information from web and work, there has to be simple semantics for the machine to process. Semantic Web exists to express such semantics in a standardized method.

Observing the need for different types of concepts, in this paper present a personalized mobile search engine [PMSE] which represents different types of concepts in different ontologies. By mining content and location concepts for user profiling, it utilizes both the content and location preferences to personalize search results for a user.

2. Related Works

A major problem in mobile web search is that the interaction between the users and 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 practical approach to capturing a user's interests for personalization is to analyze the user's clickthrough data. Clickthrough data have been used in determining the users' preferences on their search results. Many existing personalized web search systems [2],[3], [4], are based clickthrough data to determine users' preferences. In paper [1], proposed a method that provides personalized query suggestions based on a personalized clustering technique. An algorithm named agglomerative clustering algorithm to exploit query document relationship from clickthrough data. In [5] presents a new approach for situation aware personalized search. Case-Based Reasoning [CBR] approach is used.

In [6] proposed a system based in a semantic context-aware framework, which helps the user to build personalized search queries by means of an auto completion mechanism. However, most of the previous work assumed that all concepts are of the same type. In this paper separate concepts into location concepts and content concepts to recognize information importance So far there have been many papers written and researched on search engines. Most commercial search engines return roughly the same results to all users. However, different information needs even for the same query. PMSE profiles both of the user's content and location preferences in the ontology based user profiles, which are automatically learned from clickthrough and GPS data without requiring extra efforts from the user.

In PMSE propose a realistic design for PMSE by adopting the metaearch approach which relies on one of the commercial search engines, such as Google or Yahoo to perform an actual search. The client is responsible for

receiving the user's requests, submitting the requests to the PMSE server, displaying the returned results, and collecting his/her clickthrough in order to derive his/her personal preferences. The PMSE server, on the other hand, is responsible for handling heavy tasks such as forwarding the requests to a commercial search engine, as well as training and reranking of search results before they are returned to the client. The user profiles for specific users are stored on the PMSE clients, thus preserving privacy to the users. PMSE has been prototyped with PMSE clients on the Google Android platform and the PMSE server on a PC server to validate the proposed ideas 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. By mining content and location concepts for user profiling, it utilizes both the content and location preferences to personalize search results for a user. The differences between existing work and PMSE using content and location concept are:

1. Most existing location-based search systems require users to manually define their location preferences or to manually prepare a set of location sensitive topics. PMSE profiles both of the user's content and location preferences in the ontology based user profiles, which are automatically learned from the clickthrough and GPS data without requiring extra efforts from the user.
2. Propose and implement a new and realistic design for PMSE. This helps to train the user profiles quickly and efficiently.

3. System Design

The PMSE's architecture meets two 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.

In the PMSE's client-server architecture, PMSE clients are responsible for storing the user clickthrough and the ontologies derived from the PMSE server. Simple tasks, such as updating clickthoughs 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.

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. PMSE's design addressed two issues:

1. Limited computational power on mobile devices.
2. Data transmission minimization.

PMSE consists of two major activities:

Reranking the search results at PMSE server: When a user submits a query on the PMSE client, the query together with the feature vectors containing the user's content and location preferences (i.e., filtered ontologies according to the user's privacy setting) are forwarded to the PMSE server, which in turn obtains the search results from the back-end search engine (i.e., Google). The content and location concepts are extracted from the search results and organized into ontologies to capture the relationships between the concepts.

The server is used to perform ontology extraction for its speed. The feature vectors from the client are then used in RSVM training to obtain a content weight vector and a location weight vector, representing the user interests based on the user's content and location preferences for the reranking. Again, the training process is performed on the server for its speed. The search results are then reranked according to the weight vectors obtained from the RSVM training. Finally, the reranked results and the extracted ontologies for the personalization of future queries are returned to the client.

Reranking the search results at PMSE server: When a user submits a query on the PMSE client, the query together with the feature vectors containing the user's content and location preferences (i.e., filtered ontologies according to the user's privacy setting) are forwarded to the PMSE server, which in turn obtains the search results from the back-end search engine (i.e., Google). The content and location concepts are extracted from the search results and organized into ontologies to capture the relationships between the concepts.

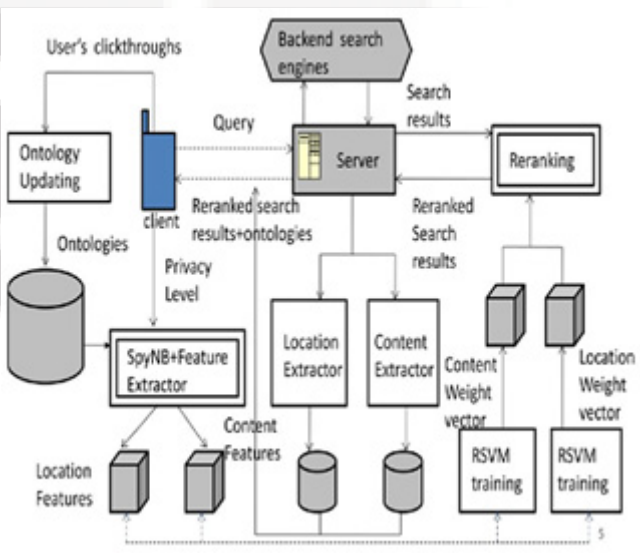

Figure 1: System design of PMSE

The server is used to perform ontology extraction for its speed. The feature vectors from the client are then used in RSVM training to obtain a content weight vector and a location weight vector, representing the user interests based on the user's content and location preferences for the reranking. Again, the training process is performed on the server for its speed. The search results are then reranked according to the weight vectors obtained from the RSVM training. Finally, the reranked results and the extracted ontologies for the personalization of future queries are returned to the client.

4. Experimental Evaluation

In the test phase, a user submits a test query and receives top 100 search results from the back-end search engine (i.e., Google) without any personalization. The user then clicks on any number of results that he/she judges to be relevant to his/her personal interest in much the same way that a standard search engine would have been used. After the users finished all of the five test queries in the test phase, the training phase begins. The clicked results from the test phase are treated as positive training samples in RSVM training. The clickthrough data, the extracted content concepts, and the extracted location concepts are employed in RSVM training to obtain the personalized ranking function. After the training phase, the evaluation phase is performed to decide if the personalized ranking function obtained in the training phase can indeed return more relevant results for the user. There are 4 factors are analyzed in this paper. They are:

1. Domain - spend time
2. Open a link
3. Explore a link
4. Ignore

For URL ranking, in old paper, only consider the click through information. But in this paper we consider the 4 factors such as domain spend time, open a link, explore a link and ignore. Fig.2 shows the performance evaluation graph. From the figure it is clear that the accuracy of PMSE using content and location concept[PMSE-CL]is greater than that of PMSE. In fig.2 the y-axis indicates the clickthrough database size and the x-axis indicates the accuracy.

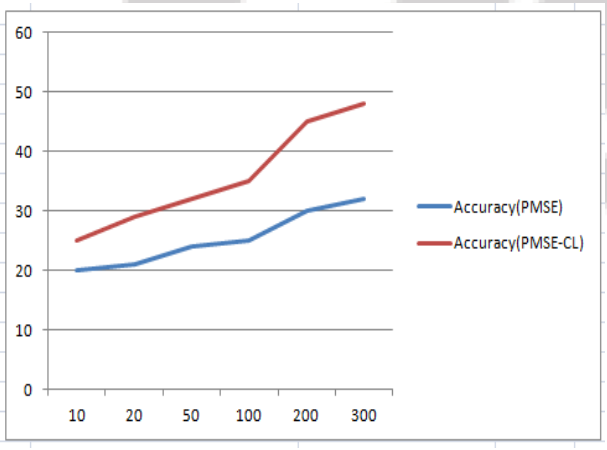

Figure 2: Performance evaluation graph

5. Conclusion

The proposed personalized mobile search engine using content and location concept 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. In this paper, personalized the search results based on user's clickthrough preferences. Search queries are classified according to preferences. Most existing location-based search systems, require users to manually define their location preferences (with latitude-longitude pairs or text form).But PMSE automatically learn from clickthrough and GPS data without requiring extra effort from the user. Here we store personal information in mobile. At the time of search we pass this information to server. Here there is chance of leaking this personal information. In future we will implement a secured personal search engine.

References

[1] S. Yokoji, "Kokono Search: A Location Based Search Engine," Proc.Int'l Conf. World Wide Web (WWW), 2001.
[2] Q. Tan, X. Chai, W. Ng, and D. Lee, "Applying Co-Training to Clickthrough Data for Search Engine Adaptation," Proc. Int'l Conf. Database Systems for Advanced Applications (DASFAA), 2004..
[3] T. Joachims, "Optimizing Search Engines Using Clickthrough Data," Proc. ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining, 2002.
[4] K.W.-T. Leung,W.Ng, and D.L. Lee, "Personalized Concept-Based Clustering of Search Engine Queries," IEEE Trans. Knowledge and Data Eng., vol. 20, no. 11, pp. 1505-1518, Nov. 2008.
[5] "Ontology Supported Personalized Search for Mobile Devices"Daniel Aréchiga, Jesús Vegas and Pablo de la Fuente Redondo,proc.int'l conf.Web search and Data Mining,2011.
[6] "semantic context aware framework for paersonalization"Y.Xu,B.Zhang,and Z.Chen,2oo7.