International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 4 Issue: 1

ISSN: 2454-4248
385 – 387

Personalized Search Engine: A Review

Ms. Kajal Thakur, Prof. PragatiPatil

Department of Computer Science and Engineering

AbhaGaikwadPatil College of Engineering Mohgaon Nagpur Maharashtra 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 users' interests and personalize the search results according to the user's 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, In this paper PMSE classifies these concepts into content concepts and location concepts To characterize the diversity of the concepts associated with a query and their relevance's to the user's need, four entropies are introduced to balance the weights between the content and location facets.

Keywords:--*Clickthrough data, concept, location search, mobile search engine, ontology, profiling.*

*****

I. INTRODUCTION

Social Network is a social structure made of individuals called nodes, which are connected by one or more specific types of interdependency, such as friendship, kinship, financial exchange, dislike, sexual relationships, or relationships of beliefs, knowledge or prestige [1]. Social Network's link represents not only the flow between personal information, but the relation status through quantitative expression. The overall graph model of Social Network is composed of many nodes and the links that connect them, and each node's direct/indirect connection forms the entire network. However, the current Personalized Systems based on Social Network were designed and constructed under the PC and it didn't provide the step by step transferring methods from PC to Smartphone. To solve these problems, this research actively analyzes an individual's characteristic based on the Social Network environment and develops a Personalized Information Retrieval System which can search for what a user wants accurately. Personalized Information Retrieval System for efficient personalized information provision proposed in this study differs from existing ones in methodology as follow: Firstly, as the system is built on the basis of NFC (Near field communication), it attempts to provide its own custom service fast and easily using its information stored in NFC. Once SNS and NFC Smartphone are associated with each other, payment is made by touching a NFC tag when visiting well known restaurants, and the information recorded in SNS is supposed to provide search results customized to individual's tastes and preferences when carrying out a search in individualized search system. That is, typing the same search keyword may bring different search results on NFC Smartphone as individuals have different preferences. Secondly, the existing Personalized Information Retrieval System fails to analyze the search system using Smartphone in Social Network environment. With an increasing number of web users using Smartphone and its individualized service under research, Smartphone environment does not provide user's search rankings suited to personal preferences. For example, when a user who wants to come by a pasta restaurant offering pasta for about 10$ and listens to rock music asks for information search via Smartphone, search results should also be prioritized and provided in favor of user's personalization taste. But, the existing systems do not show search rankings in consideration of individual's tastes and tastes. Therefore, in this study differentiated search results are provided on the basis of personalization information in User Profile Registry when a user carries out a search using Smartphone in Social Network environment. Finally this research attempts to correct uncertain or vague relation between users on the existing Social Network environment and promote a more accurate and personalized information feeding, by suggesting a Personalized Information Retrieval System using Social Network's quantitative model..

II. LITERATURE SURVEY

Many existing personalized web search systems [2], [3], [4], [5] are based clickthrough data to determine users' preferences. Joachims [3] proposed to mine document preferences from clickthrough data. Later, Ng et al. [4] proposed to combine a spying technique together with a novel voting procedure to determine user preferences. Search queries can be classified as content (i.e., non-geo) or location (i.e., geo) queries. Examples of location queries are "hongkong hotels," "museums in london," and "virginia historical sites." In [6], Gan et al. developed a classifier to

 ISSN: 2454-4248
385 – 387

classify geo and non-geo queries. It was found that a significant number of queries were location queries focusing on location information. In order to handle the queries that focus on location information, a number of location-based search systems designed for location queries have been proposed. Yokoji [8] proposed a location-based search system for web documents. Location information was extracted from the web documents, which was converted into latitudelongitude pairs. When a user submits a query together with a latitude-longitude pair, the system creates a search circle centered at the specified latitude-longitude pair and retrieves documents containing location information within the search circle.

More recently, Li et al. [9] proposed a probabilistic topicbased framework for location- sensitive domain information retrieval. Instead of modeling locations in latitude-longitude pairs, the model assumes that users can be interested in a set of locationsensitive topics. It recognizes the geographical influence distributions of topics, and models it using probabilistic Gaussian Process classifiers.

III. BASICS OF PERSONALIZED SEARCH

A. Creation of User Profile

To provide personalized search results to users, personalized web search maintains a user profile for each individual. A user profile stores information about user interests and preferences. It is generated and updated by exploiting userrelated information. Such information may include:

¬ Information about the user like age, gender, education, language, country, address, interest areas, and other information.

¬ Search history, including previous queries and clicked documents.

¬ Other user documents, such as bookmarks, favorite web sites, visited pages, and emails.

B. Server-Side and Client-Side Implement Personalized web search can be implemented on either server side (in the search engine) or client side (in the user's computer or a personalization agent)

[1]. For server-side personalization, user profiles are built, updated, and stored on the search engine side. User information is directly incorporated into the ranking process, or is used to help process initial search results. The advantage of this architecture is that the search engine can use all of its resources, for example link structure of the whole web, in its personalization algorithm. Also, the personalization algorithm can be easily adapted without any client efforts. This architecture is adopted by some general search engines such as Google Personalized Search. The disadvantage of this architecture is that it brings high storage and computation costs when millions of users are using the search engine, and it also raises privacy concerns when information about users is stored on the server. For client-side personalization, user information is collected and stored on the client side (in the user's computer or a personalization agent), usually by installing a client software or plug-in on a user's computer. In client side, not only the user's search behavior but also his contextual activities (e.g., web pages viewed before) and personal information (e.g., emails, documents, and bookmarks) could be incorporated into the user profile [2].

This allows the construction of a much richer user model for personalization. Privacy concerns are also reduced since the user profile is strictly stored and used on the client side. Another benefit is that the overhead in computation and storage for personalization can be distributed among the clients. A main drawback of personalization on the client side is that the personalization algorithm cannot use some knowledge that is only available on the server side (e.g., PageRank score of a result document). Furthermore, due to the limits of network bandwidth, the client can usually only process limited top results.

IV. CONTENT V/S HYPERLINK BASED ANALYSIS

A. Content Based Personalized Search By checking content similarities between web pages and user profile personalized search can be improved [3]. User's interests can be automatically learned by classifying implicit user data. Search results are filtered or re-ranked by checking the similarity of topics between search results and user profiles. User-issued queries and user-selected documents are categorized into concept hierarchies that are accumulated to generate a user profile. When the user issues a query, each returned result is also classified. The documents are re-ranked based upon how well the document categories match user interest profiles. Chirita et al. [4] use the ODP (Open Directory Project, http://www.dmoz.org/) hierarchy to implement personalized search. User favorite topics nodes are manually specified in the ODP hierarchy. Each document is categorized into one or several topic nodes in the same ODP hierarchy. The distances between the user topic nodes and the document topic nodes are then used to re-rank search results.

B. Hyperlink Based Personalized Search Hyperlink Analysis significantly improves the relevance of the web search results so that all major search engines claim to use some type of hyperlink analysis. Web information retrieval mainly focuses on hyperlink structures of the Web, like with Web search engine Google. In personalized Web searches, the hyperlink structures of the Web are also becoming important. The use of personalized PageRank to enable personalized Web searches was first proposed in [6], where it was suggested as a modification of the global PageRank

 ISSN: 2454-4248
385 – 387

algorithm, which computes a universal notion of importance of a Web page. The computation of (personalized) PageRank scores was not addressed beyond the original algorithm. Experiments[7] concluded that the use of personalized PageRank scores can improve a Web search. Crawling (process of gathering the web pages by the search engine) and ranking are the main uses of hyperlink analysis. In this approach, web crawler which is a software program to browse WWW in automated methodical manner find more and more web pages linked to the source page with the assumption of nearly all the linked web pages are on same topic. This process repeats for each set of web pages until no more linked pages. Then crawler of the search engine orders the web pages by the quality. To judge high quality pages, hyperlink analysis is used. In this case, search engine assumes that the source page pointed to many pages is of higher quality than a source page pointed to few numbers of pages. For this ranking process, hyperlink analysis involves with connectivitybased ranking, PageRank and HITS (Hyperlink-Induced Topic Search) algorithms etc. In addition to produce a quality and relevant web results, hyperlink analysis have several advantages like finding mirrored hosts, web page categorization and identify the geographical scope of the search etc. But in this approach, search engine has to deal with more details consist even with unnecessary stuffs also. It becomes wastage of the resources.

III.CONCLUSION

There are number of researches which are conducting to minimize some of the drawbacks of Personalized Web Search. Hyperlink based personalization algorithms work only for repeated queries, they are simple and stable. The topical interestbased personalized search algorithms implemented were not as stable as the click-based. They could improve search accuracy for some queries, but they harmed performance for more queries. Personalized Web search yields significant improvements over generic Web search for queries with high click entropy. For the queries with low click entropy, personalization methods performed similarly or even worse than generic search. As personalized search had different effectiveness for different kinds of queries, queries should not be handled in the same manner with regard to personalization. No personalization algorithms can outperform others for all queries. Different methods have different strengths and weaknesses.

REFERENCES

[1] Ji-Rong Wen, Zhicheng Dou, Ruihus Song, “Personalized Web Search”, Microsoft Research Asia, Beijing, China, 2009

[2] P.A. Chirita, C. Firan, and W. Nejdl, “Summarizing Local Context to Personalize Global Web Search”, Proc. ACM Int’l Conf. Infor. and Knowledge Management (CIKM), 2006.

[3] Zhongming Ma, Gautam Pant and Olivia R. Liu Sheng. “Interest-Based Personalized Search,” ACM Transactions on Information Systems, Vol 25, Issue 2, Article 5, February 2007.

[4] Chirita P.A., Nejdl W., Paiu R., and Kohlschu¨tter C. “Using ODP metadata to personalize search”. In Proc. 31st Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval, 2005, pp. 178–185.

[5] Aktas M. S., M. A. Nacar and F. Menczer, 2004. “Using hyperlink features to personalize web search”. Adv. Web Mining Web Usage Anal., 3932: 104-115.

[6] Page L., Brin S., Motwani R., and Winograd T. “The Pagerank Citation Ranking: bringing order to the Web”, Technical report, Stanford Dig. Lib. Tech. Project, 1998.

[7] K. Sugiyama, K. Hatano, and M. Yoshikawa. “Adaptive web search based on user profile constructed without any effort from users”. In Proceedings of the 13th international conference on World Wide Web, pages 675–684, 2004.