A PERSONALIZED WEB SEARCH MODEL USING USER IDENTIFICATION TECHNIQUE

Nidhi Saxena*, Vinodini Katiyar & Mohd. Husain***

* Assistant Professor, Department of Information Technology, Shri Ramswaroop Group of Professional Colleges, Lucknow, India.
** Head, Department of Information Technology, Dr. Shakuntala Misra National Rehabilitation University, Lucknow, India.
*** Director/Professor, JIT, Barabanki, Lucknow, India.

Abstract:
   The use of search engine is now a days common among each and every society. On the other side the web sites are available in bulk and a single search can give various different results. There exist problem of fetching results based on user interest in order to save time and to reduce complexity while searching. The personalized search based on user unique identification can solve the current problem to large extent. In this paper we have taken a novel personalization approach. We identify user and makes search according to user interest based on previous searches made by him. We present a model used for proposed technique.

Index Terms: Personalization, User Search Log & Ranking

1. Introduction:

   There are various queries made to a search engine which are short and not properly specified. Different users have different intentions for same query. For example a user having interest in biology may issue a query "kingfisher" in order to retrieve information about bird named kingfisher, while user searching about airlines may issue the same query for kingfisher airlines. To consider two different interests and to customize result based on user interest there is a necessity of personalization in search engine. With the increasing demand of personalization, number of search engines developed which provides user interest before making search. Some of them were explicitly collecting user interest before making search and some implicitly store user searches in search log in order to find user interest.

   Unfortunately, it was found that explicit collection of user feedback is not supported by all users. On the other hand the implicit methods for finding interest proved better. Most of them are based on storing search record in user search log. But there exist a major problem that the user is identified as a system and not as person and all searches made from that system are recorded in same search log. Our system is an enhancement where user logins in to search engine by a user recognition technique which creates search log and all searches made by the person are recorded, there by ranking results based on user interest.

2. Background:

   Personalization is the process of providing information to the user on the basis of user interest. User's interest can be collected explicitly by feedback or it can be implicitly collected based on user behaviour. Such information are stored in user profile, analysed and used as a sample for future search results. Several approaches were developed using explicit construction of user profile. [1]. some of them uses document present at user machine [2]. It is based on the assumption that user have interest in those documents.Another research is based on search history [3].It investigate three conditions- Session, Historic and Aggregate.

   Currently some algorithms are designed which keep record of each page visited by the user and time spend on that web page [4] [5] [6]. These algorithms predict the

interest and give results based on data collected in user search log. An optimized version of earlier algorithm was whenbeside above record; click event on the each page by the user is also stored and used for ranking [7]. In 2013 Ratio Rank [8] , a new algorithm considersin link weights and out link weights along with number of visit count. An enhanced page rank algorithm [9] came with better approach to personalization. Most of the above approaches uses implicit data collection for the purpose of personalization and there by ranking. The basic requirement in such algorithm was to have a watch over Visiting frequency and time spend on a web page but they all lack the meaning of personalization, were they only stores system search data in a user search log and do not uniquely identify each user on single machine for personalization.

3. Personalized web search model:
Our model is an extension to the existing user profile development technique used in implicit personalization. Apart from creating history about page visited and time spent, the unique identification record is also maintained in search history. This record distinguish two user's search on the same machine. The results so received will be more users specific and ranking will be improved.
The model comprises of four phases-

3.2 User identification:
At the time of login in web search engine user need to be identified. This identification will be carried with the help of face recognition by taking snap at login or by voice recognition. This biometric method of identification will uniquely identify each user and maintain their log.

3.3 User search log creation:
The system works on user identification and maintains log based on user identification. Separate log will be developed for each user on a machine. Each log having record of page visited frequency and time spend.

3.4 Searching/Ranking:
After login user enters the keyword to search .The keyword will be checked in that user search log. If found the results will be based on frequency of search and time spend. But if the search is for the first time crawler will not do ranking. The ranking will start by the second time user make the earlier keyword search.
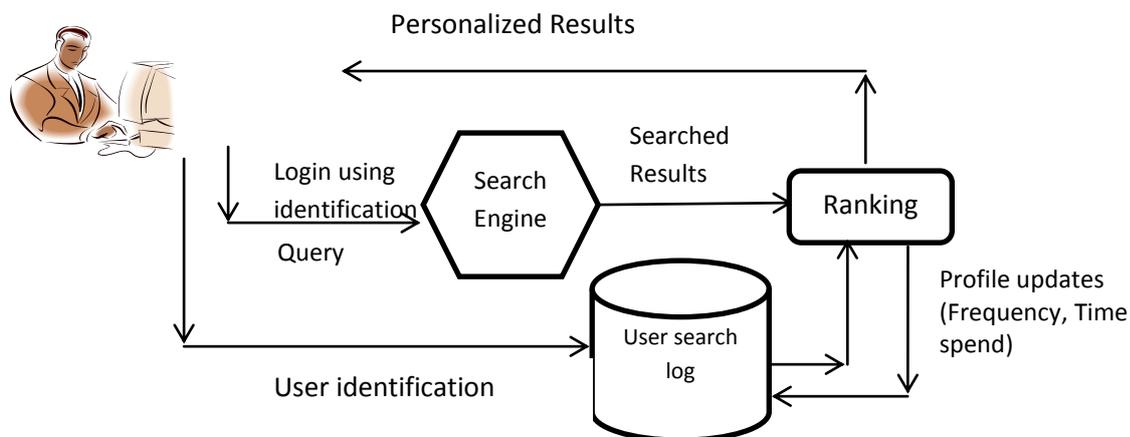

Figure 1: Personalized Web search Model

In the given Model the user login using Identification technique. For the purpose of easy identification face recognition can be used. A single system can be used by not

more than 3 or 4 persons therefore it is easy to maintain User search log for each user based on identification.

Identified user logins the search engine and enters the query. The search engine performs searching and ranking the result based on the user's earlier search record present inindividual search log. Ranking will be based on personalization technique keeping frequency of page visited and time spent on them.

4. Conclusion and Future Work:

In this paper we examined the major problem faced in personalization techniques. The problem of creating user search log, having user search history and not the system search history. This current problem will be reduced to certain extent by the proposed search model.

The system will give more accurate and relevant information to the user by considering not only the visiting frequencyand time spend but also combining the featuresof user identification and user login in search engine .This results in creating separate log and more accurate personalization. It will prove as a good tool for search engine by limiting the problems faced by previous ranking system. The next generation of WWW will be information oriented and to satisfy the customers web mining is a promising answer. In future, this study may develop strong approaches for implementation to enhance the process of personalization and ranking process in web searching.

5. References:

1. C. Srinvas, "Explicit User Profiles for Semantic web search using XML", IJERA, Volume 4, Page Number 234-241, 2012.
2. Robin Singh Bhadoria, Deepak Sain, Rahul Moriwal, " Data Mining Algorithm for Personalizing user's profile on Web", IJCTEE, Volume 1, Issue 2, 2011.
3. Paul N. Bennett, Ryen W. White, Wei Chu, Susan T. Dumais, Peter Bailey, Fedor Borisyuk, and Xiaoyuan Cui, "Modeling the Impact of Short- and Long-Term Behavior on Search Personalization", 35th Annual ACM SIGIR Conferences 2012.
4. Haibin Liu, VladoKeselj, "Combined Mining of Web Server logs and web contents for classifying user navigation patterns and predicting user's future requests", Elsevier Data and Knowledge Engineering, 2007.
5. R. Khanchan, M. Punithavalli, "An Efficient Web Page Prediction based on Access Time Length and Frequency", IEEE 3rd International Conference on Electronics Computer Technology (ICECT) , 2011.
6. Niclaas Matthijs, Filip Radlinski, "Personalizing Web Search using Long Term Browsing History", ACM WSDM'11, Page Number 9–12, 2011.
7. Rohit Agarwal, K. V. Arya, Shashi Shekhar, "An Efficient Weighted Algorithm for Web Information Retrieval System", IEEE International Conference on Computational Intelligence and Communication Systems, 2011.
8. Ranveer Singh , Dilip Kumar Sharma, "Ratio Rank: Enhancing theImpact of In links and Out links", IEEE International Advance Computing Conference, 2013
9. Ranveer Singh, Dilip Kumar Sharma,, "Enhanced Ratio Rank Enhancing Impact of In links and Out links", IEEE Conference on Information and Communication Technologies, 2013
10. Jaime Teevan, Susan T. Dumais, Eric Horvitz," Personalizing Search via Automated Analysis of Interests and Activities", 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM 2005.