

A Survey of Web Page Personalization in Web Search Engine

Sofia Sayed¹, Reeba R²

¹ PG Scholar, Department of Computer Science & Engineering,
Sree Buddha College of Engineering, Alappuzha, India
sofivm.s@gmail.com

² Assistant Professor, Department of Computer Science & Engineering,
Sree Buddha College of Engineering, Alappuzha, India
reeba.amjith@gmail.com

Abstract

Web search engines play an important role in web life. Generic web search engines are not suitable for identifying different needs of different customers. If users enter an improper or ambiguous keyword and the inability of users to express what they need are some of the main challenges faced by generic engines. We have to personalize the search results to address this issue. Personalized web search (PWS) is designed to provide different search results for different users. Major barrier for the wide proliferation of PWS is user's unavailability with exposing their private information during search. For keeping users' information safe and ensuring privacy, search engines should provide security mechanisms. Search engines can provide more accurate and specific data if users trust search engines and provide more information. In this paper we will discuss on different techniques used in personalized web search and securing personalized information.

Keywords: *Personalized web search, Taxonomy, Personalization Techniques, Privacy, Information Retrieval*

1. Introduction

Internet is used widely by different types of users to satisfy various information needs. However, an equivocal query/topic submitted to search engines may not satisfy user information needs, because each user has different intentions for querying. So, we have to first identify their search goals. That evaluation can be very useful in improving search engine relevance and user knowledge. Personalization technologies are used to present right information to the right user at a right moment. Personalized Web search is a method used for providing a better search results.

PWS is used to improve the search quality of users by customizing the search results for people with different information goals. However, users might experience error when search engines return unsatisfactory results that do not meet their real goal. Such unsatisfactory of users is largely due to the enormous variety of users' contexts and backgrounds.

Need of security in the personalized web search is another important factor. Users are not ready to disclose their information during web search. This has become a major issue in profiling the user in personalized web search. There should be a mechanism which considers profiles according to information provided by user. If the search engine knows more about the user, then more accurate results will be obtained by search provider. But trust of users is an important factor here. Search engines can provide better results if users trust search engines and provide more information. Search engines should have to provide security mechanisms then only users will be ensuring privacy and their information should be kept safe.

Personalized web search is categorized into two types, namely click-log-based methods and profile-based ones. The click-log based methods simply impose bias to clicked pages in the user's query history. The main limitation for this is it can only work on repeated queries from the same user. In profile-based methods improve the search experience with complicated user-interest models generated from user profiling techniques. These methods can be effective for almost all sorts of queries, but in some conditions it will be unstable but are reported to be unstable under some circumstances, these privacy issues make the user uncomfortable.

In personalized web search, user information is collected and analyzed in order to find intention behind issued query fired by user. Typically search is performed by providing queries to retrieval system in form of set of words. If different users enter same query, the system will produce same results without considering the user. But search results should be produced by taking the user in the equation, so that different users can get different search results for same query. By keeping track of user's personal information and interests.

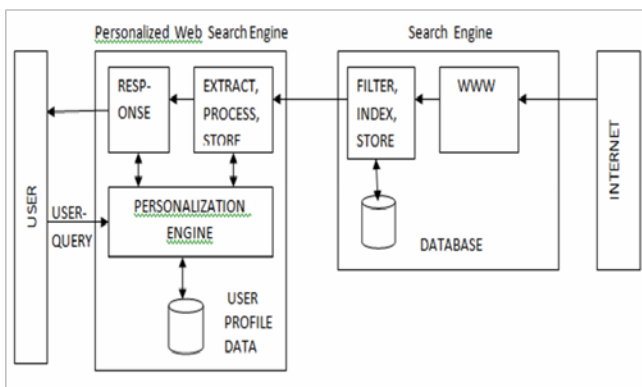


Figure: Personalized Web Search

2. Related Works

Web page personalization uses different techniques for personalizing the web page and keeping privacy over there. Users have to describe their general interests is one in that. For example, Google Personal asks users to create a profile of themselves by selecting categories of interests. This profile can then be used for personalizing search results by mapping web pages to the same categories. Another method is the use of implicit representations of a user's long and short term interests. By the use of this, no need for the user to specify a profile of interests. Kelly, D. and Teevan, J et al.,[1], they studied about how to improve search accuracy and intelligently personalize information retrieval by exploit implicit user modeling. Unlike previous work, it give preference to the use of immediate search context and implicit feedback information. It eagerly updating of search results to provide maximum benefit to user. Author presented a

decision-theoretic framework for optimizing interactive information retrieval based on eager user model updating, in which the system responds to every action of the user by choosing a system action to optimize a utility function.

Koenmann, J. and Belkin, N. et al., [2],proposed specific techniques to exploit and capture two types of implicit feedback information:(1)identifying related immediately preceding query and using the query and the corresponding search results to select appropriate terms to expand the current query, and (2) exploiting the viewed document summaries to immediately re-rank any documents that have not yet been seen by the user.

J.Lai et al., [3] compared personalized search results with user profile results. The large amount of information is available on web. When user searches for different type of queries of same topic some time it provide same result. So it is difficult for user to get proper and relevant results because it does not consider user preferences and area of interest.

P.Paleti et al., [4] by using probabilistic query expansion author developed a personalized web search. In this approach, the authors created a personalized web search system applied at proxy which changes to user interests completely by generating user profile with the help of collaborative filtering. A user profile basically consists of probabilistic correlations among query and document terms which are utilized for providing personalized search results. Experimental result prove that this proposed personalized Web search system is very useful and effective.

Kyung-Joong Kim et al., [5] implemented a personalized Web search engine by using fuzzy concept network with link structure. Many search engines uses link structure to find out precision. Output of link based search engine comparatively high than text based search engine but complexity is there. One of the best way to obtain more effective and desirable results is personalisation. Fuzzy concept with network and link structure is helpful to find user query's subjective interest. The proposed approach is used to make results personalized by using link based search techniques.

Matthijs and Radlinski et al.,[6] proposed another method which collect web usage data that is URL of page, page session date and time, duration of page visit, length of the source HTML using Firefox add on. Peng et al. [7] construct user profile by gathering search result used by user with reference of Google directory. In this architecture tree structure is maintaining and each topics are linked in tree. In that tree directory users search topic is keeping. Link Visited count is maintaining here and it shows degree of interest.

Acharyya and Ghosh [8] propose a general personalization framework based on the conceptual modeling based on the users' navigational behavior. The proposed methodology consists of three steps involves mapping each visited page to a particular topic or concept, imposing a tree structure (taxonomy) on these topics, and then estimating the parameters of a semi- Markov process defined on this tree based on the observed user paths.

3. Conclusion

This paper covers issues like need of personalized web search, how personalized web search can be implemented, and existing system of personalized web search. It also gives a survey report of different ways to maintain privacy in personalized web environment. The future scope of this paper is to design a framework to provide a complete privacy of the users so that they can work without any fear of working in personalized web environment.

References

[1] Kelly, D. and Teevan, J. (2003). Implicit feedback for inferring user preference: A bibliography. *SIGIR Forum*, 37(2): 18-28.

[2] Koenmann, J. and Belkin, N. (1996). A case for interaction: A study of interactive information retrieval behavior and effectiveness. In *Proceedings of CHI '96*, 205-212.

[3] J.Lai and B. Soh, "Personalized Web search results with Profile comparisons," Third International Conference on Information Technology and Applications (ICITA 2005), Vol. 1, Pp. 573 – 576,2005.

[4] P. Palleti, H. Karnick and P. Mitra, "Personalized Web Search Using Probabilistic Query Expansion," International Conferences on Web Intelligence and Intelligent Agent Technology Workshops (IEEE/WIC/ACM), Pp. 83 – 86, 2007.

[5] Kyung-Joong Kim and Sung-Bae Cho, "A personalized Web search engine using fuzzy concept network with link structure," Joint 9th IFSA World Congress and 20th NAFIPS International Conference, Vol. 1, Pp. 81 –86, 2001.

[6] N.MatthijsandF. Radlinski. (2011) "Personalizing Web search using long term browsing history ". In Proceedings of the ACM WSDM Conference on Web Search and Data Mining, pp. 25–34.

[7] Peng, Xueping, ZhendongNiu, Sheng Huang, and YuminZhao."Personalized Web Search Using Clickthrough Data and Web Page Rating."Journal of Computers 7, no. 10 (2012): 2578-2584.

[8] S.Acharyya, J. Ghosh, "Context-Sensitive Modeling of Web Surfing Behaviour Using Concept Trees", in Proc. of the 5th WEBKDD Workshop, Washington, August 2003.

[9] S.E.Middleton, N.R. Shadbolt, D.C. De Roure, "Ontological User Profiling in Recommender Systems", ACM Transactions on Information Systems (TOIS), Jan. 2004/ Vol.22, No. 1, 54-88.

[10] P. Kearney, S. S. Anand, "Employing a Domain Ontology to gain insights into user behaviour", in Proc. of the 3rd Workshop on Intelligent Techniques for Web Personalization (ITWP 2005), Endinburgh, Scotland, August 2005.

[11] Shen, X. and Zhai, C. X. (2003). Exploiting query history for document ranking in interactive information retrieval. In Proceedings of SIGIR '03 (Poster), 377-378.

[12] Sparck Jones, K., Walker, S. and Robertson, S. A. (1998). Probabilistic model of information retrieval: Development and status. Technical Report TR-446, Cambridge University Computer Laboratory.

[13] Speretta, M. and Gauch, S. (2004). Personalizing search based on user search history. Submitted CIKM'04. <http://www.itc.ku.edu/keyconcept/>.

[14] Shen, Xuehua, Bin Tan, and Cheng Xiang Zhai. "Context-sensitive information retrieval using implicit feedback." Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2005.

[15]. Xu, Yabo, et al. "Online anonymity for personalized web services." Proceedings of the 18th ACM conference on Information and knowledge management. ACM, 2009.

[16]. A. Viejo and J. Castell_a-Roca, "Using Social Networks to Distort Users' Profiles Generated by Web Search Engines," Computer Networks, vol. 54, no. 9, pp. 1343-1357, 2010.

[17] G. Chen, H. Bai, L. Shou, K. Chen, and Y. Gao, "Ups: Efficient Privacy Protection in Personalized Web Search," Proc. 34th Int'l ACM SIGIR Conf. Research and Development in Information, pp. 615-624, 2011.