MJRET
Open Access
Multidisciplinary Journal of Research in Engineering and Technology
Journal homepage: www.mjret.in
ISSN:2348 - 6953

INFORMATION PRIVACY PROTECTION IN PERSONALIZED WEB SEARCH FOR IMPROVING RETRIEVAL EFFECTIVENESS

Ravindra S Ambekar, Deepak D. Devde, Yoginath A Kasture, Bhushan M. More
Department of Computer Engineering, JSPM's Imperial college of Engineering and Research, Wagholi, Savitribai Phule Pune University,
Pune, India
ravindra12amb@gmail.com, devdedd@gmail.com, yoginath97kasture@gmail.com
bhushanm336@gmail.com

Abstract: *There are lots of Web search engines such as Yahoo, Google, Redifmail, etc. These web search engines are widely used to search huge amount of data on Internet. Search Engines are the basic tool of the internet from where user can collects needed information by searching on Internet. Now a day's user spends lots of time on internet for searching accurate result. So we study privacy protection in PWS (Personalized Web Search) application that model create users profile and we proposed a PWS framework called UPS that is general category which is use for searching and aim is to provide better search result. For the generalization purpose we represent two algorithms Greedy IL & Greedy DP. To check whether a personalization query is helpful or not, so we use online prediction mechanism for run time generalization.*

Keywords:*Personalized web search, utility, privacy protection, risk, profile Introduction*

1. INTRODUCTION

Now a day's searching is one of the most common tasks performed on the internet. Web search engine become the most common important portal for ordinary people looking for useful information on the web. But some time users have experience about failure when search engine return irrelevant results that do not satisfy their real intention. Personalized search is a promising way to improve the search quality by customizing search result for people with different information goals. PWS is a searching technique which aims to provide good search result.

The PWS can generally divide into two method name as Click-log based and Profile based method. Click-log based is straightforward method in that clicked pages are directly goes into

users query history. Although this strategy has represented to perform consistently and considerable well it can work on repeated queries from same user. Which is a strong limitation confining its applicability. In Profile based method is to improve search experience with complicated users interest models generated from users profile technique. This method improves quality of web search as well as protects browsing history such as click through data, bookmark and user document. In PWS as per the user requirement we create user profile according to user interest we search query from the internet.

There are advantage and disadvantage for both types of PWS techniques. The profile based PWS has represented more effectiveness in improving quality of web search recently, with increasing usage of personal and behaviour information to profile its user which is usually collected from browsing history [5],bookmarks [9], user document [2] [10], click through data [7],[8],[1], query history [2],[3],[4], privacy issues are raising from the lack of protection for search the data, for instance the AOL query log scandal [11], not only raise panic among individual users. In fact privacy concerns have become the major drawback for wide proliferation of PWS service.

2. RELATED WORK

There are two important part in PWS system first is profile-based personalization and Second privacy protection.

2.1 Profile Based Personalization

The basic goal of Profile Based Personalization is to change the quality of search result as the provide better less search result as they desired. In Profile Based Personalization before searching result user have to create the Profile. In that Profile user have to add basic information about his/her and add interest and user have their unique id and password. If user want to search information then first user have to login with their unique id and password. If user login with incorrect id and password then user unable to search the result. The Profile based Personalization is very main module in the Personalized Web Search.

In the remaining section , we see the previous solutions to PWS on mainly two parameter name as measure the personalization effectiveness and Profile representation. There are many Profile Representation available to provide different personalization method. The term bag of words or vectors to represent their profile. There is lot of work is done on better scalability, strong descriptive ability and higher access efficiency field.

2.2 Privacy Protection in PWS System

In Personalized Web Search there are two big problem. First is consider sensitivity of the data particularly user profile extended to server of PWS and second is those treat privacy as individual identification as mentioned in. The Protecting user identification problem try to solve on pseudo identity, identity group, no personal information on different level. In the PWS System Privacy protection is very important issue. There are many ways to protect the privacy in Personalized web Search by provide unique id and password before searching any information. In privacy protection in PWS system privacy is maintain by using protecting the user search result this history is very important to the user because there is very secure

or important information he searched by protecting the user history security of PWS become very strong.

3. PROPOSED SYSTEM

As given in Fig.1, UPS consists of a non trusty search engine server and a number of clients. Each client (user) accessing the search service trusts no one but himself/ herself. The key component for privacy protection is an online profiler implemented as a search proxy running on the client machine itself. The proxy maintains both the complete user profile, in a hierarchy of nodes with semantics, and the user-specified (customized) privacy requirements represent a set of sensitive- nodes. The framework works in two phases, namely the offline and online phase, for each user. During the offline phase a hierarchical user profile is constructed and customized with the user-specified privacy requirements. The online phase handles queries as follows:

1) When a user issues a query qi on the client, the proxy generates a user profile in runtime in the light of query terms. The output of this step is a generalized user profile Gi satisfying the privacy requirements. The generalization process is guided by considering two conflicting metrics, namely the personalization utility and the privacy risk, both defined for user profiles.
2) Subsequently, the query and the generalized user profile are sent together to the PWS server for personalized search.
3) The search results are personalized with the profile and delivered back to the query proxy.
4) Finally, the proxy either presents the raw results to the user, or re ranks them with the complete user profile.

UPS is distinguished from conventional PWS in that it

A. Provides runtime profiling, which in effect optimizes personalization utility while representing user's privacy requirement.
B. Allow for customization of privacy needs
C. Doesn't require user interaction.

We proposed PWS that gives desired user result. After successful login user has to select three fields. First user has to select Interest field in Interest field there are so many interest according to user. After selecting Interest field user has to select second field that is Sub Interest field. After selecting Sub Interest field user has to select third Field that is Su Sub Interest filed. These entire three field are very important for searching the result. According to these field PWS search the result and give better search result. This is our proposed system.

3.1 SYSTEM ARCHITECTURE

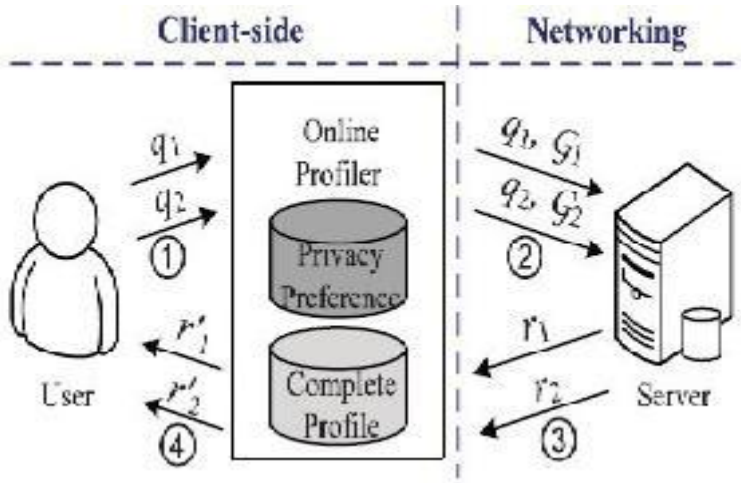

Fig.1: System Architecture

We proposed two simple but effective generalization algorithms, GreedyDP and GreedyIL, to support run time profiling. While the former tries to maximize the discriminating power (DP), the latter attempts to minimize the information loss (IS). The result of our experiments represents efficiency and effectiveness of our UPS framework.

4. RESULT DISCUSSION

Following diagram shows the User Satisfaction. There are three field first is Total site in database second is search web site or search result and third is user satisfaction. In first field there are many sites are stored in database. As per the web site stored in database user see web site. It is very easy to add and remove web site in the database. The second field that is search result that shows the all search result according to the user interest. And Third field that is User satisfaction that shows the al information based on Interest field.

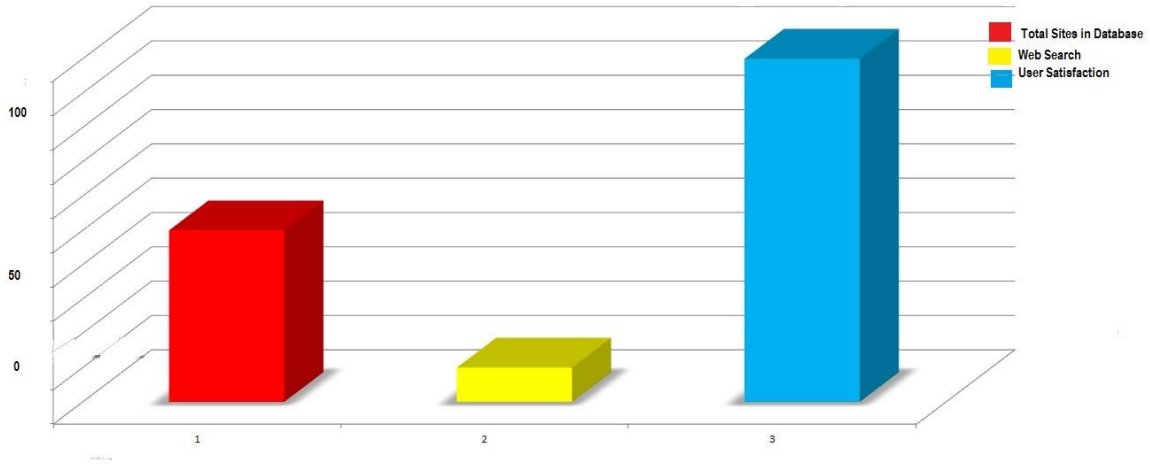

Fig.2: Result Graph of User searching result

5. CONCLUSION

This paper provides information on User customizable privacy preserving search framework called UPS for Personalized Web Search. In also UPS performed online generalization on

user profiles to secure the personal privacy without compromising the search quality. Basically, UPS capture the user profile of the PWS in a hierarchical classification. We suggest two greedy algorithms, namely GreedyDP and GreedyIL, for the online generalization.

ACKNOWLEGEMENT

We are always thankful to our guide Prof. V.S.Wadne for his valuable discussion and constructive suggestions, which greatly contributed to our Paper. We sincerely thank our Head of Department (Comp.) Prof.S.R.Todmal for his reassuring encouragement throughout the preparation of our Paper. We are also grateful to our respected Principal Dr. Sachin Admane sir for his co-operation despite his busy schedule. Last but not the least we thank all the staff members of Computer Department for providing an excellent environment to undergo this Paper.

REFERENCES

[1]. Z. Dou, R. Song, and J.-R. Wen, "A Large-Scale Evaluation and Analysis of Personalized Search Strategies," Proc. Int'l Conf. World Wide Web (WWW), pp. 581-590, 2007.

[2]. J. Teevan, S.T. Dumais, and E. Horvitz, "Personalizing Search via Automated Analysis of Interests and Activities," Proc. 28th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), pp. 449-456, 2005.

[3]. M. Spertta and S. Gach, "Personalizing Search Based on User Search Histories," Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence (WI), 2005.

[4]. B. Tan, X. Shen, and C. Zhai, "Mining Long-Term Search History to Improve Search Accuracy," Proc. ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD), 2006.

[5]. K. Sugiyama, K. Hatano, and M. Yoshikawa, "Adaptive Web Search Based on User Profile Constructed without any Effort from Users," Proc. 13th Int'l Conf. World Wide Web (WWW),2004.

[6]. X. Shen, B. Tan, and C. Zhai, "Implicit User Modeling for Personalized Search," Proc. 14th ACM Int'l Conf. Information and Knowledge Management (CIKM), 2005.

[7]. X. Shen, B. Tan, and C. Zhai, "Context-Sensitive Information Retrieval Using Implicit Feedback," Proc. 28th Ann. Int'l ACM SIGIR Conf. Research and Development Information Retrieval (SIGIR),2005.

[8]. F. Qiu and J. Cho, "Automatic Identification of User Interest for Personalized Search," Proc. 15th Int'l Conf. World Wide Web (WWW), pp. 727-736, 2006.

[9]. J. Pitkow, H. Schu¨tze, T. Cass, R. Cooley, D. Turnbull, A. Edmonds, E. Adar, and T. Breuel, "Personalized Search," Comm. ACM, vol. 45, no. 9, pp. 50-55, 2002.

[10]. Y. Xu, K. Wang, B. Zhang, and Z. Chen, "Privacy-EnhancingPersonalized Web Search," Proc. 16th Int'l Conf. World Wide Web (WWW), pp. 591-600, 2007.

[11]. K. Hafner, Researchers Yearn to Use AOL Logs, but They Hesitate,New York Times, Aug. 2006.

[12]. A. Krause and E. Horvitz, "A Utility-Theoretic Approach to Privacy in Online Services," J. Artificial Intelligence Research,vol.

[13]. J.S. Breese, D. Heckerman, and C.M. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," Proc. 14th Conf. Uncertainty in Artificial Intelligence (UAI), pp. 43-52, 1998.