

Providing Privacy in Personalized Web Search(PWS) Queries

Sagar Singh¹, Khushal Mistri², Lokesh Joshi³, Y.N.S.Prathyusha⁴, Mrs. A.S. Deoker⁵

^{1, 2, 3, 4}A.I.S.S.M.S, Pune, ⁵ Asst. Prof., A.I.S.S.M.S, PUNE

Department of Computer Engineering
AISSMS College of Engineering Affiliated to
Savitribai Phule Pune University Pune, Maharashtra, India

Abstract - As increasing rate of users of internet massive amount of data transaction is taken place, due to this there comes a need of data security with high quality and effective searching techniques. PWS application shows the effectiveness and help to improve the qualities of searching services over the web sites and search engines. Study shows there is a lots of drawbacks in security terms while profile of users generated in personalize web search. We are going to study security in PWS applications that model user queries as hierarchical user profile. We demonstrate the PWS framework called UPS which can used for generalize profile by user's queries. We demonstrate two greedy algorithms that are known as greedyDP and greedyIL, this two algorithm is beneficial for run time generalization. Online prediction are another motive to provide by us to predict whether personalizing a query is beneficial or not. This search also relevant the result about GreedyIL significantly outperforms and GreedyDP shows the efficiency.

Keywords- Privacy Protection, Personalize web search, User profile, Personalization techniques.

I. INTRODUCTION

At this era there is tremendous demand of internet, and this is necessary or future data is also a very popular topic among all. Search results are according to the users requirements. For such results there is a necessary to implement personalized web environment. Web search results should adapt to user requirements. Main issue in search engines that there may not provide accurate and efficient user requirements even they produce same and confusing results.

Main motive should be given to areas like question answer methodology, better browsing, virtual results, localization and modified web search. Search engines should display results depending upon the user. Many technologies have been implemented for distinct user as per their needs.

Currently search engines are dealing with all users same search results regardless of user as per their special needs. Though indexing algorithms are used in traditional search engines has difficulty in generating efficiency expected by

user. To obtain data related to user's interest personalization this is being generated. Generic web search differs from personalized web search because it will provide the different result for same users queries. Information intended from different users may be different for the same queries fired. Mixed type of search results will be retrieved. Because of this required time will be less to choose information that particular user wants. By knowing some things about user, a search engine might refine user results to make them more relevant is the concept behind personalized search. The web search engine has long become the most important portal for ordinary people looking for useful information on the web. Personalizing web search is the process of obtaining web search results depending upon history or past behavior of user. Depending upon conditions, Personalization can be done at server or client side . User may be aiming to achieve many goals in single query. Hence efficiency of personalization of web search is depended on user behavior and query. In web personalization, reading minds of users is

challenging task to perform. Also there is limitation of words used to search.

In personalized web search there is essential aspect of security. Sometimes User's queries. This cause primary issues while profiling in the user's personalized web search. Generalizing profile should provide for user's information and this is best way to provide security. More the search engine knows about user information, more accurate search results will be obtained by search engine or web browsers. User may not fully trust on search it can get misuse by third person. Search engines can provide specific data and more accurate data if users trust and provide all required information to search engine. Therefore, search engines can provide mechanism such that user's keep trust and there's information will remain safe and under full security.

Research can be focused on following areas:-

- By using personalization of the user profile, that will improve search result quality.
- To maintain privacy risk under control

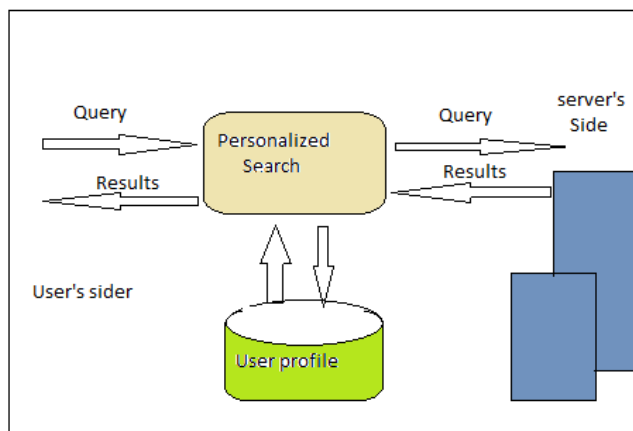


Fig 1: Personalized Search Engine Architecture

II. LITERATURE REVIEW

In [5] this paper, author study this problem and provide some primary conclusions over some searching strategies. It demonstrates a large-scale framework for personalized search on the basis of user query logs and also established as click and profile based strategies. As per their result , author exposes that

personalized search has significant efficient and perform well as compare to the simple web search on some queries but it show little effect on some other queries while searching on the simple web search . Author also exposes that both long term and short-term contexts are essential in improving search performance for both profile-based personalized search strategies as well as click log based strategies. In this paper, author tries to experiments to investigate whether personalization is consistently effective under distinct situations. In this paper they proposed a Profile-based searching strategy but there may chances to lack in stability as that of click log based strategies for some of user queries. The author could improve the search accuracy on some queries, but it also hazards some of queries. Since these Strategies are not as much to optimal, they will keep trying to search to improve them in future [10]. Author searched for profile-based methods, both long-term and short-term contexts are essential in improving search performance and bring efficiency in Web search. .

In paper [1], Author developed new personalized approach which uses online decision on the query personalization. Previous problem has been overcome by this approach and it gives

- Offline profiling. It gives personal user profile because of which gives better search result.
- It support customization of privacy requirement, and personalized inconsistency.
- User interaction while creating search results . May lead challenge of efficiency.

On user-level and query-level it gives user requirement generalization which allows user to add requirement

Gang in [2], proposed a UPS (User customizable privacy preserving search) framework where according to user's privacy requirement and query content while submitting query proxy generalizes user profile. Next query along with user profile is sent to PWS server. Results will personalized with user profile and send back to proxy server again so it can show result to the user. By providing Security of privacy is online generalizer which carries offline and online phases.

Paper[1],[2] focuses on the literature of profile-based personalization which improves the search utility and privacy-protection in PWS. The aim of user profiling is to collect the information about the user and his/her interests. In related paper [3], user profiling process consist of Data Collection, Profile Constructor and Application or Algorithm which utilize the user profile information in order to provide personalized services. User identification is crucial for system that constructs user profile. Many research has been done on constructing this user profile in different manner such as Weighted Keyword profile, Semantic Network profiles, Concept profiles [3], Search History, and many more. Yabo Xu in [4] has implemented Split and BuildUP algorithms to construct user profile hierarchically. Through this profile, user can control the information which will be publicly available to server.

This paper [6] was motivated by two emerging trends: web users want personalized services and web users want privacy. One challenge is that personal information must be made anonymous under the assumption that the participating parties, including the web service, are not completely trusted, due to systematic collection of personal information in addition to queries. Another goal is the generation of online and dynamic nature of searching result of user's queries. Author proposed the notion of online anonymity to protect web users and proposed an approach to maintain online anonymity through time. This approach makes use of a third party called the user pool and it does not require the user pool to be trusted. The simulation study on real US demographics showed promising results: it is feasible to achieve personalization for reasonable privacy settings.

III. PROPOSED WORK

We suggest a privacy-preserving personalized web search framework UPS. According to user-specified privacy requirements it can simplify profiles for each query. We also formulate the problem of privacy-preserving personalized search as -Risk Profile Generalization, with its NP hardness proved relying on the meaning of two contradictory metrics, that are named as personalization utility and privacy risk on that

personalization, and also that hierarchical based user profile. To support runtime profiling a GreedyDP and GreedyIL generalization algorithms develop which is simple but effective. While the former tries to maximize the discriminating power (DP), the latter attempts to minimize the information loss (IL). GreedyIL outperforms GreedyDP considerably, by developing a number of heuristics. We give a reasonably priced mechanism for the client to make a decision whether to personalize a query in UPS. This choice can be made previous to each runtime profiling to improve the constancy of the search results while keep away from the needless contact of the profile. Our extensive experiments show the competence and efficiency of our UPS framework.

The UPS framework works on two phases that is offline and online.

1. Offline Phase In offline phase a hierarchical user profile is build and modified with the user-specified privacy supplies.
2. Online Phase In online phase the proxy generates a user profile in runtime in the light of query terms when a user issues a query q_i on the client. Generalized user profile G_i pleasing the privacy supplies is the output of this step. By bearing in mind two conflicting metrics, namely the personalization utility and the privacy risk, both defined for user profiles are directed by generalized process. Subsequently, for personalized search, the query and the generalized user profile are sent together to the PWS server. After that the search results are personalized with the profile which is generated on the basis of user's query and bring back that query to the query proxy. Lastly, the proxy either re ranks them with the complete user profile or presents the raw results to the user and provides the security.

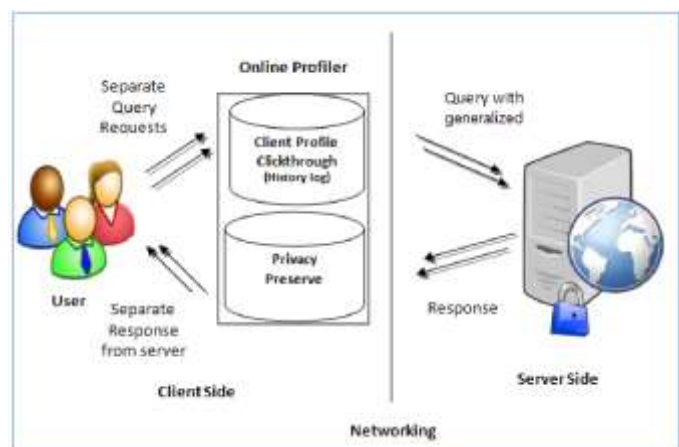


Figure 2 : System Architecture

IV. Conclusion

Thus this paper demonstrates a client side privacy protection framework that is known as UPS. PWS is initialized on UPS framework i.e. UPS framework can work for any PWS that captures user profile in hierarchical way. User can customized privacy requirements and this framework that is based on the hierarchical profiles. This framework performs online generalization on user queries to protect personal privacy without decreasing search quality. For such things we proposed 2 greedy algorithms that are greedy greedyIL and greedyDp, this two algorithm are for online generalization. We will try to reset background knowledge such as high level of relationship between queries or topics i.e. exclusive, sequentially and many more and also we are trying to build a method which will be capable to capture a series of queries. We also try to implement complex method to build the user profile, and better way to predict the performance in UPS framework.

V. References

- [1] Lidan Shou, He Bai, Ke Chen, and Gang Chen, "Supporting Privacy Protection in Personalized Web Search, Knowledge and Data Engineering, IEEE Transactions on , vol.26, no.2, pp.453,467, Feb. 2014 doi: 10.1109/TKDE.2012.201.
- [2] 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. 615624,2011.
- [3]. Xu, Yabo, et al. "Online anonymity for personalized web services." *Proceedings of the 18th ACM conference on Information and knowledge management*. ACM, 2009.
- [4] Y. Xu, K. Wang, B. Zhang, and Z. Chen, "Privacy Enhancing Personalized Web Search," Proc. 16th Int'l Conf. World Wide Web (WWW), pp. 591-600, 2007.
- [5]. Dou, Zhicheng, Ruihua Song, and Ji-Rong Wen. "A large-scale evaluation and analysis of personalized search strategies", *Proceedings of the 16th international conference on World Wide Web*. ACM, 2007.
- [6] X. Shen, B. Tan, and C. Zhai, "Privacy Protection in Personalized Search," SIGIR Forum, vol. 41, no. 1, pp. 4-17, 2007.
- [8] H.R. Kim and P.K. Chan, "Learning implicit user interest hierarchy for context in personalization, " Proc. of International conference on Intelligent User Interfaces (IUI), Miami, Florida, 2003.
- [9] Mingxuan Yuan Lei Chen Philip S. Yu, "Personalized Privacy Protection in Social Network," Proc. VLDB,2010.
- [10] 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.
- [11] Fang Liu , Clement Yu , Weiyi Meng, "Personalized web search by mapping user queries to categories". [12] J. Teevan, S. T. Dumais, and E. Horvitz. "Beyond the commons: Investigating the value of personalizing web search". In *Proceedings of PIA '05*, 2005.
- [13] 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.
- [14] M. Spertta and S. Gach, "Personalizing Search Based on User Search Histories," Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence (WI), 2005.
- [15] Priyanka Deulkar, "A Study on Securing Privacy In Personalized Web Search" IJECS Volume 4 Issue 3 March, 2015 Page No.11100-11103