

# Privacy Protection in Personalized Web Search

<sup>1</sup>Mrs. Manisha Deshmukh, <sup>2</sup>Prof. Umesh Kulkarni

<sup>1,2</sup> Department of Computer Engineering, ARMIET, Department of computer Engineering, VIT, India

**Abstract:** The quality of various search services on the Internet is effectively improved by using personalized web search (PWS). Personalized web search is a promising way to improve search quality by customizing search results for people with individual information goals. However, evidences show that user’s reluctance to disclose their private information during search has become a major barrier for the wide proliferation of PWS. Privacy protection in PWS applications model user preferences as hierarchical user profiles. PWS framework called UPS can adaptively generalize profiles by queries while respecting user specified privacy requirements. Runtime generalization aims at striking a balance between two predictive metrics that evaluate the utility of personalization and the privacy risk of exposing the generalized profile. Two greedy algorithms, namely Greedydp and GreedyIL, are used for runtime generalization. An online prediction mechanism for deciding whether personalizing a query is beneficial is provided. Extensive experiments demonstrate the effectiveness of the framework. The experimental results also reveal that GreedyIL significantly outperforms GreedyDP in terms of efficiency.

**Keywords:** Greedy dp and Greedy IL, PWS, UPS, hierarchical profiles.

## 1. INTRODUCTION

To protect user privacy in profile-based PWS, to improve the search quality with the personalization utility of the user profile [1]. To protect user privacy without compromising the personalized search quality. User information are collected and analyzed to fetch the user intention behind the issued query. User customizable Privacy-preserving Search (UPS) is used to generalize profiles by queries with user privacy requirements. The click-log based methods are straightforward they simply impose bias to clicked pages in the user’s query history[5]. It can only work on repeated queries from the same user, which is a strong limitation confining its applicability. Profile-based methods can be potentially effective for almost all sorts of queries, but are reported to be unstable under some circumstances. Improve the search experience with complicated user-interest models generated from user profiling techniques.[7]. The UPS framework allowed users to specify customized privacy requirements via the hierarchical profiles. In addition, UPS also performed online generalization on user profiles [ 10] .

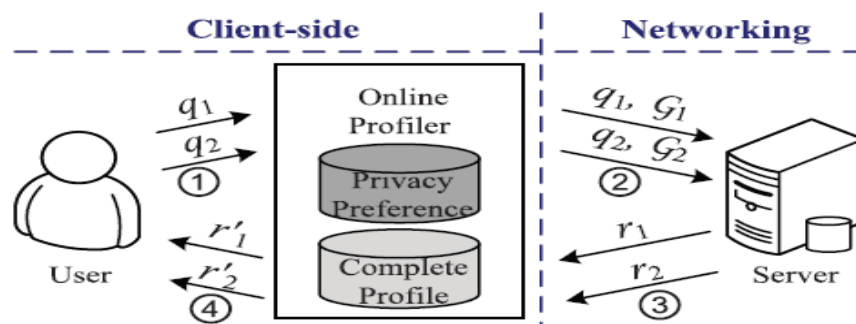


Fig.1.Block Diagram

## 2. LITERATURE SURVEY

Dou, Song and Wen 2007 [1] Personalized web search is a promising way to improve search quality by customizing search results for people with individual information goals. Private information during search has become a major barrier for the wide proliferation of PWS. We studied Privacy protection in PWS model user preferences as hierarchical user profiles.

Teevan, Dumais and Horvitz 2005 [2] personalized search has been under way for many years and many personalization algorithms have been investigated, In this paper, we study this problem and provide some findings. We present a large-scale evaluation framework for personalized search based on query logs and then evaluate five personalized.

Speritta and Gach 2005 [3] Personalized web search (PWS) has demonstrated its effectiveness in improving the quality of various search services on the Internet. We study privacy protection in PWS applications that model user preferences as hierarchical user profiles.

We propose a PWS framework called UPS that can adaptively. Sugiyama, Hatano, and Yoshikawa 2004 [4] In this paper, we first propose several approaches to adapting search results according to each user's need for relevant information without any user effort, and then verify the effectiveness of our proposed approaches.

Sugiyama, Hatano, and Yoshikawa 2004 [5] In this paper, we first propose several approaches to adapting search results according to each user's need for relevant information without any user effort, and then verify the effectiveness of our proposed approaches.

## 3. EXISTING SYSTEM

The solutions to PWS can generally be categorized into two types, namely click-log-based methods and profile-based ones. [1][2].The click-log based methods are straightforward— they simply impose bias to clicked pages in the user's query history. Although this strategy has been demonstrated to perform consistently and considerably well [1] it can only work on repeated queries from the same user, which is a strong limitation confining its applicability [4]In contrast, profile-based methods improve the search experience with complicated user-interest models generated from user profiling techniques.[7] Profile-based methods can be potentially effective for almost all sorts of queries, but are reported to be unstable under some circumstances .We propose a privacy-preserving personalized web search framework UPS, which can generalize profiles for each query according to user-specified privacy requirements.[8]

## 4. PROPOSED SYSTEM

The UPS framework allowed users to specify customized privacy requirements via the hierarchical profiles. In addition, UPS also performed online generalization on user profiles to protect the personal privacy without compromising the search quality. [8]The two greedy algorithms were proposed, namely Greedy DP and Greedy IL, for the online generalization. [9]. an online prediction mechanism is provided for deciding whether personalizing a query is beneficial. Utility of personalization and the privacy risk of exposing the generalized profile metrics are used to analyze the system.

Existing system was having some drawbacks like:-

Controlling Session attacks

Capability to capture the series of queries

Log file optimization and maintenance was not considered

## 5. PROBLEM DEFINATION

To protect user privacy without compromising the personalized search quality. User information are collected and analyzed to fetch the user intention behind the issued query. User customizable Privacy-preserving Search (UPS) is used to generalize profiles by queries with user privacy requirements. Two Greedy algorithms are used to generalize the user query profiles. Greedy discriminating power algorithm (Greedy DP) is used to maximize Greedy Information Loss (Greedy IL) is used to minimize the information loss in user profiles. Greedy IL algorithm achieves high efficiency than

the Greedy DP algorithm. An online prediction mechanism is provided for deciding whether personalizing a query is beneficial. Utility of personalization and the privacy risk of exposing the generalized profile metrics are used to analyze the system. The following drawbacks are identified from the existing system.

#### **SOFTWARE REQUIREMENTS:-**

Operating System: Independent of Operating System

Application Libraries: Java and J2EE

Language: J2EE and Java

Front End: Net Beans

Database 1.,Data base 2.

#### **HARDWARE REQUIREMENTS:-**

Processor: Pentium IV. (& onwards).

Memory (RAM):1GB RAM (32 bit) or 2 GB (64 bit)

Hard disk: 40GB

Internet access

## **6. CALCULATION**

In this project Personalized web search (PWS) is used to improve the quality of various search services on the Internet. Privacy preserved PWS methods are used to protect the disclosure of personal information in search process. User customizable Privacy-preserving Search (UPS) framework is used to support privacy in search process. The UPS scheme is enhanced with attack resistant methods. Personalization utility is high in the personalized web search scheme. The system reduces the generalization risk levels. The system increases the attack control rate. Priority based user profile construction process is supported by the system

## **ACKNOWLEDGEMENT**

I would like to express my sincere gratitude towards my guide Prof. Umesh Kulkarni for the help, guidance and encouragement in the development of this methodology. They supported me with scientific guidance, advice and encouragement, and were always helpful and enthusiastic and this inspired me in my work. I have benefitted from numerous discussions with guide and other colleagues.

## **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. Schütze, 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-Enhancing Personalized 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. 39, pp. 633-662, 2010.
- [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.
- [14] P.A. Chirita, W. Nejdl, R. Paiu, and C. Kohlschütter, "Using ODP Metadata to Personalize Search," Proc. 28th Ann. Int'l ACM SIGIR Conf. Research and Development Information Retrieval (SIGIR), 2005.
- [15] .A. Pretschner and S. Gauch, "Ontology-Based Personalized Search and Browsing," Proc. IEEE 11th Int'l Conf. Tools with Artificial Intelligence (ICTAI '99), 1999.