SMART AND SECURE PERSONALIZED WEB SEARCH

Avinash Gholap, Tushar Burde, Vikas Chaudhari,
Nikhil Dalavi, Prof.Laxmi Madhuri

Computer Engineering, Dr. D.Y.Patil, School of Engineering
Pune, India.

Abstract: *As the size of the Internet continues to expand, the users of different search providers continuously demand search results which meets their need. Personalized Search is one of the options available to users in order to sculpt that the result return to them is accurate. This raises concerns of privacy issues however as users are usually uncomfortable revealing personal information to the third party, like service providers. This paper aims to deal with the privacy issues revolving around personalized search and discusses ways that privacy can be enriched to make the users become more comfortable with the release of their personal data in order to gain more accurate search results.*

Keywords*: PWS, O.S Authentication, UPS.*

1. INTRODUCTION

Today a web search engine has accomplished a lot of popularity and importance for users seeking information on the web. Since the contents available on web is immense and cryptic, users most of the time experience failure when an irrelevant result of user's query is returned from the search engine [1][3].The search engines usually cannot distinguish different users' needs well. For example, a user of computer science may use the search query "leopard" to locate information on Apple OS X Leopard and a user from biology filed may use the same query for the animal leopard; however a search engine does not treat both the query differently. Alternatively, personalized search provides customized results.[2]

Personalized web search (PWS) is a broad type of search techniques aiming at providing better search results, which are tailored for individual user needs. The user information has to be collected and analyzed to figure out the exact intention of user behind issued the query. [3]The solutions to PWS can generally be categorized into two types naming as click-log-based methods and profile-based. In click-log based method — they simply enforce preference to clicked pages in the user's query history. Whereas, the profile-based methods improve the search experience with complicated user-interest models generated from user profiling techniques [1] [3].

In spite of having pros and cons as the types of PWS techniques, the profile-based PWS has more potency in remodeling the quality of web search recently, with growing usage of personal and behavioral data to profile its users, which is consistently gathered from query history. Unfortunately, such implicitly collection of personal data can easily reveal a gamut of user's private life. Privacy issues is emerging from the lack of security for such data, for instance the AOL query logs scandal, not only scare individual user, but also dampen the data-publishers have high intensity in offering personalized service. In fact, privacy concerns have become the major obstacle for wide proliferation of PWS services.[3]

DRAWBACKS:

- All the sensitive topics are disclosed using an unconditional metric known as surprisal based on the information theory.[1]
- Run-time profiling is not supported by existing profile-based PWS. [3][4]
- The existing methods do not take into account the customization of privacy requirements.[1][4]
- Many personalization techniques require constant user interactions when creating personalized search results. [7]

2. PROPOSED WORK

2.1. Profile-Based Personalization

This paper introduces an approach to personalize digital multimedia information based on user profile information. For which two main mechanisms were developed: a profile generator that automatically creates user profiles defining the user preferences, and a content-based recommendation algorithm that derives the user's interest area in unknown content by matching their profile to metadata storage of the content. Both features are integrated into a personalization system. [1][3]

2.2. Privacy Protection in PWS System

This paper proposes a PWS framework called UPS that can derive profiles in for each query according to user-specified privacy requirements [8]. Two predictive metrics are offered to calculate the privacy breach risk and the query utility for hierarchical user profile. It develops two simple but effective abstraction algorithms for user profiles allowing for query-level customization using our proposed metrics. This paper also provides an online prediction mechanism based on query utility for deciding whether to personalize a query in UPS [8] [10]. Overall experiments demonstrate the efficiency and effectiveness of our framework [4].

2.3. O.S Authentication

O.S also has a unique characteristic in its communication implementation that serves to identify the uniqueness of the O.S over a network. By analyzing certain protocol flags, options, and data in the packets a device sends on network. It can increase reliability and accuracy in guesses about the O.S that sent those packets. O.S Authentication is a process of that analysis. By analyzing initial values of packets, we can determine O.S installed on host.

M30-2-4-10-2015

3. CONTRIBUTION

3.1. O.S Authentication

Many modern Malware implementations carry out their activities on guest operating system to escape detection from antivirus software, which is installed on host operating system. To avoid this type of activities this paper proposed the system, which identifies the packets send from client machine which is encrypted by a host operating system not by the guest operating system. The host operating system identifies the genuine packet by using information available from the TCP packet. The system distinguishes the TCP packets by analyzing certain values in the packets that client send over a network. If the system determine the packet they got differ from the operating system installed on host machine, an unauthorized operating system is likely present .The system uses TCP SYN packet's TCP/IP headers to identify the host OS generating the packet.

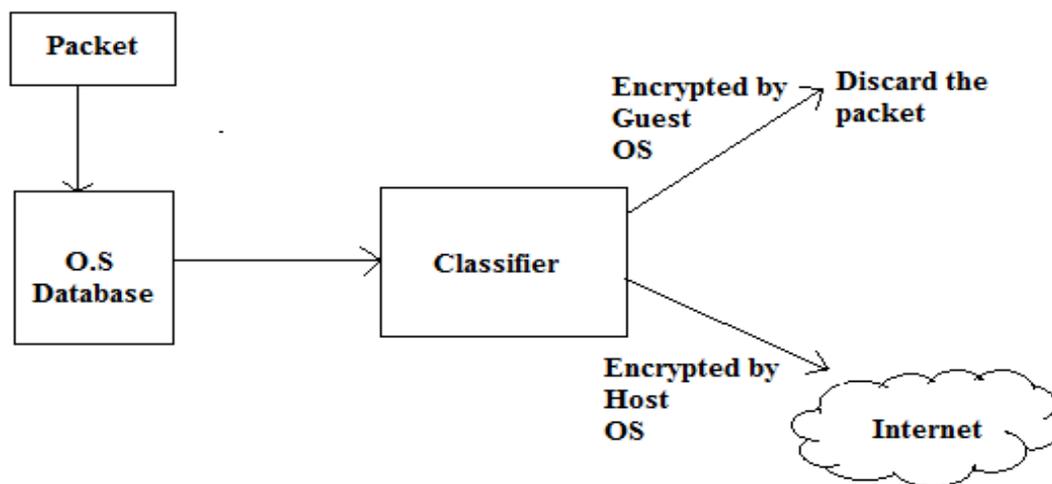### 3.2. System Architecture of O.S Authentication Technique

Fig.1: System Architecture

4. CONCLUSION

In this paper Personalized web search (PWS) proposes the improved quality of various search services on the Internet. Privacy preservation technique in PWS protects the disclosure of personal information during the User's session. User customizable Privacy-preserving Search (UPS) is a framework used to support privacy in search process. The risk levels are reduced by generalization of the system. The system increases the attack control rate. The client side system is more secure from malicious activity and different attacks over the network.

5. FUTURE SCOPE

- Enhances the stability of the search quality.
- The user profile exposure is avoided.
- Searching deep web pages and ranking to them for effective search query. It contains relevant or irrelevant search result.
- The Validation user's search result according to the links resulted by user query. For that different parameters of URL are taken into account.

ACKNOWLEGEMENT

REFERENCES

[1]. *Lidan Shou, He Bai, Ke Chen, and Gang Chen, " Supporting Privacy Protection in Personalized Web Search" , IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL.26, NO.2, FEBRUARY 2014.*

[2]. *Rakesh Kumar and Aditi Sharan, " Personalized web search using browsing history and domain knowledge" , 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*

[3]. *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.*

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

[5]. *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*

[6]. *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.*

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

[8]. *X. Shen, B. Tan, and C. Zhai, "Privacy Protection in Personalized Search," SIGIR Forum, vol. 41, no. 1, pp. 4-17, 2007.*

[9]. *Y. Xu, K. Wang, G. Yang, and A.W.-C. Fu, "Online Anonymity for Personalized Web Services," Proc. 18th ACM Conf. Information and Knowledge Management (CIKM), pp. 1497-1500, 2009.*

[10].*Y. Zhu, L. Xiong, and C. Verdery, "Anonymizing User Profiles for Personalized Web Search," Proc. 19th Int'l Conf. World Wide Web (WWW), pp. 1225-1226, 2010.*

M30-2-4-10-2015