International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 04 | April -2020 ISSN: 2582-3930

Supporting Privacy Protection in Personalized Web Search

B. Sudhanthira[1], Master of Computer Application,
Mr. S. Barath[2], M.C.A., M.Phil., Assistant Professor,
Krishnasamy college of Engineering and Technology, Cuddalore.

Abstract

Personalized web search (PWS) has demonstrated its effectiveness in improving the quality of various search services on the Internet. However, evidences show that users' reluctance to disclose their private information during search has become a major barrier for the wide proliferation of PWS. 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 generalize profiles by queries while respecting user-specified privacy requirements. Our 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. We present two greedy algorithms, namely GreedyDP and GreedyIL, for runtime generalization. We also provide an online prediction mechanism for deciding whether personalizing a query is beneficial. Extensive experiments demonstrate the effectiveness of our framework

Introduction

The web search engine has gained a lot of popularity and importance for users seeking information on the web. Since the contents available in web is very vast and ambiguous, users at times experience failure when an irrelevant result of user query is returned from the search engine. Therefore, in order to provide better search, result a general category of search technique Personalized Web search is used. In personalized web search, user information is collected and analyzed in order to find intention behind issued query fired by user. There are two categories of PWS, namely click-log-based and profile-based. The click-log based methods are straightforward— they simply impose bias to clicked pages in the user's query history. This strategy has been performing well but it work on repeated queries from same user which is a strong limitation to its applicability. While profile-based methods improve the search experience generated from user profiling techniques. Profile-based methods can be potentially effective for almost all sorts of queries, but are reported to be unstable under some circumstances. There are both advantages and disadvantages for both type of PWS technique, profile based PWS is more effective for improving search result. The user profile is made from information gathered Supporting Privacy Protection in Personalized Web Search- A Review from query history, browsing history, click-through data bookmarks, user documents and so forth. Unfortunately, such implicitly collected personal data can easily reveal a gamut of user's private life. Profile-based methods prove to be effective for almost all sorts of queries, but they may become unstable under some circumstances. Though there are limitations in this approach, it has demonstrated more effectiveness in improving the quality of web

search. In this paper we will study how to provide privacy for the personalized web search applications that model the user preferences as hierarichal user profiles. We provide a framework called UPS (User customizable Privacy-preserving Search) which provides the search results by adapting to the user's information needs and also provides privacy according to the user specified privacy requirements which help the user to choose content and degree of detail of the profile information that is exposed to the search engine. An online prediction mechanism is provided for deciding whether to personalize the query (by exposing the profile) and what to expose in the user profile at run time.

Existing System

The existing works concentrate on server-side personalized search services in preserving privacy, it provide a less security to the user. To provide a security to the user from the profile-based PWS from the client side, many researchers have to deem two challenging effects during the search process of the user Profile-based PWS mainly focus on improving the search utility. The basic idea of these works is to tailor the search results by referring to, often implicitly, a user profile that reveals an individual information goal. In the remainder of this section, we review the previous solutions to PWS on two aspects, namely the representation of profiles, and the measure of the effectiveness of personalization. Many profile representations are available in the literature to facilitate different personalization strategies

Disadvantage

1. All the sensitive topics are detected using an absolute metric called surprisal based on the information theory.

2. The existing profile-based PWS do not support runtime profiling.

3. The existing methods do not take into account the customization of privacy requirements.

4. Personalization techniques require iterative user interactions when creating personalized search results

Proposed System

This paper proposes a privacy- preserving personalized web search framework called UPS i.e User customizable Privacy- preserving Search, that generalize profile for every query as per user privacy specification. Based on personalization and privacy risk metric, this paper formulate Risk Profile Generation, with its NP- hardness proved. It develops two simple but effective generalization algorithms, GreedyDP and Greedy IL, to support runtime profiling. GreedyDP maximize the discriminating power (DP) while GreedyIL minimize the information loss (IL). This paper also provide a mechanism for the client to decide whether or not to personalize a query in UPS. This decision is made before each runtime profiling to enhance the stability of the search results.

Advantages

1. It enhances the stability of the search quality.

2. It avoids the unnecessary exposure of the user profile.

3. The 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

Modules Description

Profile-Based Personalization:

This paper introduces an approach to personalize digital multimedia content based on user profile information. For this, two main mechanisms were developed: a profile generator that automatically creates user profiles representing the user preferences, and a content-based recommendation algorithm that estimates the user's interest in unknown content by matching her profile to metadata descriptions of the content. Both features are integrated into a personalization system.

Privacy protection in PWS system:

We propose a PWS framework called UPS that can generalize profiles in for each query according to user-specified privacy requirements. Two predictive metrics are proposed to evaluate the privacy breach risk and the query utility for hierarchical user profile. We develop two simple but effective generalization algorithms for user profiles allowing for query-level customization using our proposed metrics. We also provide an online prediction mechanism based on query utility for deciding whether to personalize a query in UPS. Extensive experiments demonstrate the efficiency and effectiveness of our framework

Generalizing User Profile:

The generalization process has to meet specific prerequisites to handle the user profile. This is achieved by preprocessing the user profile. At first, the process initializes the user profile by taking the indicated parent user profile into account. The process adds the inherited properties to the properties of the local user profile. Thereafter the process loads the data for the foreground and the background of the map according to the described selection in the user profile. Additionally, using references enables caching and is helpful when considering an implementation in a production environment. The reference to the user profile can be used as an identifier for already processed user profiles. It allows performing the customization process once, but reusing the result multiple times. However, it has to be made sure, that an update of the user profile is also propagated to the generalization process. This requires specific update strategies, which check after a specific timeout or a specific event, if the user profile has not changed yet. Additionally as the generalization process involves remote data services, which might be updated frequently, the cached generalization results might become outdated. These selecting a specific caching strategy requires careful analysis.

Online Decision:

The profile-based personalization contributes little or even reduces the search quality, while exposing the profile to a server would for sure risk the user's privacy. To address this problem, we develop an online mechanism to decide whether to personalize a query. The basic idea is straightforward. if a distinct query is identified

during generalization, the entire runtime profiling will be aborted and the query will be sent to the server without a user profile.

Client:

Every client have no trust on anyone except on himself/herself.
Profiler gives strong strength as a proxy search which runs as clients system Whole user profile is perfectly maintained by proxy
A hierarchy of nodes by own language and also privacy. Framework work in phase i.e. in two phase, viz., offline and another phase is online for every user.

Server:

To maintain the privacy and confidentiality by encrypting the user profile at the server side. Secure and Efficient Manner

Conclusion

In this paper we presented a client-side privacy protection framework called UPS (User Customizable Privacy Preserving) Search for personalize websearch.UPS could likely be adopted by any PWS that capture user profiles in a hierarchical taxonomy. Our proposed framework provided customized privacy requirements via the hierarchical profiles to the users. Through this profile, users control what portion of their private information is exposed to the server and the users can specify to which degree the content should be protected. Web users were increases because of available of information's from the web browser based on the search engine. With the increasing number of user service engine must provide the relevant search result based on their behavior or based on the user performance. Providing relevant result to the user is based on their click logs, query histories, bookmarks, by this privacy of the user might be loss. For providing relevant search by using these approaches the privacy of the user may loss. Most existing system provides a major barrier to the private information during user search. That approaches does not protect privacy issues and rising information loss for the user data.
For this issue this paper proposes client based architecture based on the greedy algorithm to prevent the user data and provide the relevant search result to the user in future it can include this work in mobile application.
.

Future Enhancement

For generalizing retrieve data by using the background knowledge. Through this we can resist the adversaries. Privacy protection in publishing transaction data is an important problem. A key feature of transaction data is the extreme scarcity, which renders any single technique ineffective in anonymizing such data. Among recent works, some incur high information loss, some result in data hard to interpret, and some suffer from performance drawbacks. This paper proposes to integrate generalization and compression to reduce information loss .The integration is non-trivial. We propose novel techniques to address the efficiency and scalability challenges. Our proposed system gives better quality results and gives more efficiency. Privacy is too good when compared with the Existing system. In the

Existing System, only generalization technique is used. Our String-matching algorithm gives more accuracy when compared with the Greedy IL algorithm. Generalization and suppression technique achieves better privacy when compared with the existing system.

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]. Xu, Yabo, et al. "Privacy-enhancing personalized web
search." *Proceedings of the 16th international conference on World
Wide Web*. ACM, 2007

[3]. 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), vol. 11, no. 9, pp. 449-456, 2005.

[4] 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.

.
[5] 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.