Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 5, Issue. 2, February 2016, pg.146 – 149

A REVIEW ON PERSONALIZED WEB SEARCH ALONG WITH PRIVACY PROTECTION

Rupali S. Atote[1], Prof. R. R. Shelke[2]

[1, 2] Department of Computer Science and Engineering, H.V.P.M.'s College of Engineering & Technology, Amravati, India

[1] rupali.atote@gmail.com; [2] rajeshrishelke@reddiffmail.com

Abstract— As in the Todays world use of web search is growing on increasing. The Personalized web search (PWS) has demonstrated its effectiveness in improving the quality of various search services on the Internet. 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. We study privacy protection in PWS applications that model user preferences as hierarchical user profiles. We propose a PWS framework 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 some most useful techniques 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.

Keywords— Personalized Web Search (PWS), Web Search Engine, user query logs, UPS (User customizable Privacy-preserving Search), Profiles.

I. INTRODUCTION

The web search engine is widely used by the users for searching useful information on the web. But the amount of information on the web grows continuously so it becomes very difficult for web search engines to find information that satisfies user's individual needs. Due to the enormous variety of user's contexts and backgrounds, as well as the ambiguity of texts, search engines return irrelevant results that do not meet the user's real intentions.

Personalized Web Search (PWS) is a general category of search techniques aiming at providing better search results, which are provided for individual user's needs. The solutions to PWS can be categorized into two types click-log-methods and profile-based-ones. 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, it can only work on repeated queries from the same user, which is a strong limitation confining its applicability. In contrast, profile-based methods improve the search experience with complicated user-interest models generated from user profiling techniques. [1]

Although there are reasons and considerations for both types of PWS techniques, the profile-based PWS has proved its more effectiveness in improving the quality of web search recently, with increasing usage of one's

personal and behavioural information to profile its users, which is usually gathered implicitly with the help of query history, browsing history, click-through data , bookmarks, user documents , and so on Unfortunately, such type of collected personal data can easily reveal an entire scope of user's private life. Protecting privacy issues rising from the lack of protection for such data, for example the AOL query logs scandal, not only raise panic among individual users, but also downs the data publisher's enthusiasm in offering personalized service. In fact, privacy concerns have become the major barrier for wide use of PWS services [4].

Existing system have a privacy-preserving personalized web search framework UPS. User specifies the privacy requirements and according to the requirements user profiles are generalized. The problem of privacy preserving personalized search is formulated as Risk Profile Generalization, by using two conflicting metrics, personalization utility and privacy risk, for hierarchical user profile. To enhance the stability of the search results and to avoid the unnecessary exposure of the profile an inexpensive mechanism is used for deciding whether to personalize a query in UPS. UPS allows customization of privacy needs; and it does not require iterative user interaction. The ultimate goal of data mining is prediction and predictive data mining is the most common type of data mining and one that has the most direct business applications [7].

The exhaustive and widespread use of computers and the improvement in Data base technology have provided large Data. The flourishing growth of Data in Databases has generated an urgent need for efficient data mining techniques to discover useful information and Knowledge [8]. cautions committee as indicated on the conference website. Information about final paper submission is available from the conference website.

II. BASICS OF PERSONALIZED SEARCH

An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.

A. Creation of User Profile

To provide personalized search results to users, personalized web search maintains a user profile for each individual. A user profile stores information about user interests and preferences. It is generated and updated by exploiting user-related information. Such information may include:

- Information about the user like age, gender, education, language, country, address, interest areas, and other information.
- Search history, including previous queries and clicked documents.
- Other user documents, such as bookmarks, favourite web sites, visited pages, and emails.

B. Server-Side and Client-Side Implement

Personalized web search can be implemented on either server side (in the search engine) or client side (in the user's computer or a personalization agent). For server-side personalization, user profiles are built, updated, and stored on the search engine side. User information is directly incorporated into the ranking process, or is used to help process initial search results. For client-side personalization, user information is collected and stored on the client side (in the user's computer or a personalization agent), usually by installing a client software or plug-in on a user's computer. In client side, not only the user's search behaviour but also his contextual activities (e.g., web pages viewed before) and personal information (e.g., emails, documents, and bookmarks) could be incorporated into the user profile. This allows the construction of a much richer user model for personalization. Privacy concerns are also reduced since the user profile is strictly stored and used on the client side. Another benefit is that the overhead in computation and storage for personalization can be distributed among the clients [5].

III. MOTIVATIONS

To protect user privacy in profile-based PWS, researchers have to consider two contradicting effects during the search process. In an ideal case, significant gain can be obtained by personalization at the expense of only a small (and less-sensitive) portion of the user profile, namely a generalized profile. A better approach is to make an online decision on A: whether to personalize the query (by exposing the profile) and

B: what to expose in the user profile at runtime.

To the best of our knowledge, no previous work has supported such feature. We need predictive metrics to measure the search quality and breach risk after personalization, without incurring iterative user interaction.

UPS framework works in two phases:

1. Offline phase
2. Online phase

In offline phase hierarchical user profile is constructed and customized with the user-specified privacy requirements.

Online phase handles queries as follows:

The generalization process is guided by considering two conflicting metrics, namely the personalization utility and the privacy risk, both defined for user profiles. The query and the generalized user profile are sent together to the PWS server. UPS is distinguished from conventional PWS in that it

1) provides runtime profiling, which in effect optimizes the personalization utility while respecting user's privacy requirements;

2) Allows for customization of privacy needs; and

3) Does not require iterative user interaction.

Or details must not show any professional title (e.g. Managing Director), any academic title (e.g. Dr.) or any membership of any professional organization (e.g. Senior Member IEEE).

To avoid confusion, the family name must be written as the last part of each author name (e.g. John A.K. Smith).

Each affiliation must include, at the very least, the name of the company and the name of the country where the author is based (e.g. Causal Productions Pty Ltd, Australia).

Email address is compulsory for the corresponding author.

IV. CONTRIBUTIONS

The proxy maintains both the complete user profile, in a hierarchy of nodes with semantics, and the user-specified (customized) privacy requirements represented as a set of sensitive-nodes. While the former tries to maximize the Discriminating Power (DP), the latter attempts to minimize the Information Loss (IL). By exploiting a number of heuristics, GreedyIL outperforms GreedyDP significantly [2].

V. ADVANTAGES

- It achieves better search results.
- It achieves the privacy results when applying the background knowledge to the user profiling results.
- It has less computational time and communicational time.
- It achieves better accuracy when compared with the Existing Works [6].
- It protects personal privacy without compromising search quality.
- It is consistently effective on different queries for different users, and under different search contexts.
- Personalized Search has significant improvement over common web search.
- It is used to find certain data among a huge amount of information in a minimal amount of time.
- Proposed protocol improves the existing solutions in terms of query delay [3].

VI. CONCLUSION

In this paper we are presenting a client-side privacy protection framework called UPS for personalized web search. UPS could potentially be adopted by any PWS that captures user profiles. The framework allowed users to specify customized privacy requirements via the hierarchical profiles. To protect user privacy in profile-based PWS, researchers have to consider some contradicting effects during the search process. In addition, UPS also performed generalization on user profiles to protect the personal privacy without compromising the search quality. We proposed greedy algorithms, for the profile generalization. We think that UPS could achieve quality search results while preserving user's customized privacy requirements.

It supports runtime profiling. It is more efficient. The literature results also confirmed the effectiveness and Inefficiency of our solution. For future work, we will try to resist adversaries with broader background knowledge, such as richer relationship among topics or capability to capture a series of queries from the victim. We will also seek more sophisticated method to build the user profile, and better metrics to predict the performance of UPS.

REFERENCES

[1] Reshma Y. Totare, Meenakshi Rathod, Prajakta Raskar, "Personalized Web Search protecting privacy using Greedy Algorithm", International Journal of Emerging Technology and Innovative Engineering ,Volume I, Issue 2, February 2015 ,ISSN: 2394 – 6598.

[2] Lidan Shou, He Bai, Ke Chen, and Gang Chen, "Supporting Privacy Protection in PersonalizedWeb Search", IEEE Transactions On Knowledge And Data Engineering, Vol.26, No.2,February 2014.Guilherme B. Castro, José Sidnei C. Martini, "Biologically-Inspired Neural Network for Traffic

Signal Control", 17th International Conference on Intelligent Transportation Systems (ITSC), October 8-11, 2014.

[3] V.Aruna1 , K.Palanivel , "A Privacy-Preserving Personalized Web Search Framework UPS", Globala Journal of Advanced Engineering Technologies, Vol.3, Issue4- 2014 ISSN (Online): 2277-6370 & ISSN: 2394-0921.

[4] Ms. A. S. Patil,Prof. M.M.Ghonge,Dr. M. V. Sarode, "User customizable Privacy-preserving Search Framework-UPS for Personalized Web Search", International Journal of Research in Advent Technology, Vol.2, No.4, April 2014 ,E-ISSN: 2420-14141.

[5] Chanchala Joshi, Teena Jaiswal, Himanshu Gaur, "An Overview Study of Personalized Web Search", International Journal of Scientific and Research Publications, Volume 3, Issue 1, January 2013 ISSN 2250-3153.

[6] Archana Ukande, Nitin Shivale, "Supporting Privacy Protection in Personalized Web Search with Secured User Profile", International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064

[7] R. R. Shelke, Dr. R. V. Dharaskar, Dr. V. M. Thakare, "Data Mining For Mobile Devices Using Web Services", International Conference on Industrial automation And Computing (ICIAC - 12th & 13th April 2014) , Jhulelal Institute of Technology, Nagpur.

[8] R. R. Shelke ,Dr. V. M. Thakare, Dr. R . V. Dharaskar, "Study of Data Mining Approach for Mobile Computing Environment",International Journal on Computer Science and Engineering (IJCSE) , ISSN : 0975-3397, Vol. 4, 12 Dec 2012.

