International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 06 | June-2016 www.irjet.net p-ISSN: 2395-0072

CUSTOMIZED PERSONALIZED WEB SEARCH WITH PRIVACY PROTECTION

Rakesh G H [1], Mrs. Pratibha Mishra [2]

[1] *M.Tech Student, Department of CSE, BITM College, Karnataka, INDIA*
[2] *Associate Professor, Department of CSE, BITM College, Karnataka, INDIA*

Abstract - *Internet now a day's growing widely and continuously, as it is increasing in its size the users who use the services of it demanding search results for their queries accurately and quickly. This raises concerns of privacy issues however as users are typically uncomfortable to provide their personal information to the service provider on the Internet. The aim is to deal with the privacy issues surrounding personalized search and discusses ways that privacy can be improved so that users can become more comfortable with the release of their personal data in order to receive more accurate search results.*

Key Words: Privacy protection, Personalized Web Search (PWS), risk, profile.

1. INTRODUCTION

Now-a-days people depend much on the internet for any information they required about their education, business and lifestyle etc. It is one of the service providers, which provide the information that is required to the user with the help of Web search engine. Search Engine is the tool which allows the user's to find the information from the World Wide Web. Web contains large information which will be made available to the user through the search engines. When user request information about any particular topic, the search engine gives data in large quantity which are useful and also provides the data that are not much related to the topic that is requested by user.

This useless data is mainly because of various users who may search the same topic but with different meaning at the same time.

In order to avoid this kind of search process, technique called Personalized Web Search (PWS) is used. In this kind of web search, as the name indicates personalized web search, services will be provided personally for each user with better search results for their quires.

There are mainly two categories of PWS, namely click-log-based and profile-based. The click-log based method provides service to user based on the user's query history that is based on the pages that are searched by the user repeatedly; it will concentrate on those clicked pages and provide service. While profile-based methods improve the search experience to the user by providing the service based on the information that is given by the user in their profile.

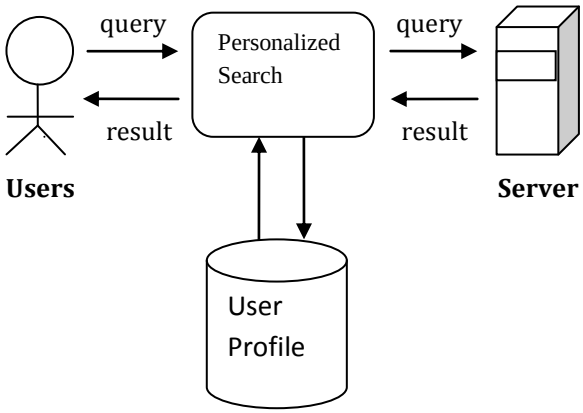

Fig -1 Architecture of PWS

2. PROPOSED APPROACH

The customized PWS with privacy protection will allow the user to browse the internet in a personalized web search with comfort and convenient manner. The approaches or methodology proposed are:

A. User registration:

The user will be provided with option and will be asked to provide information related to their personal information

in the user profile during registration, these information will be used to provide the accurate result during web search.

B. Displaying output:

Here in this proposed approach the output for the user requested queries will be displayed differently from the existing systems and here some customized output will be shown to the user which will make user to feel comfort.

3. SCOPE OF THE PROJECT

It is designed to provide protection for the user's personal information that is provided to the web search and also provides privacy protection for the user's profile with minimizing the risk that may occur during search with the help of a Personalized Web Search frame work called UPS.

4. OVERVIEW OF MODULES

MODULES DESCRIPTION:

A. Profile-Based Personalization
B. Generalizing User Profile
C. Online Decision
D. Privacy Protection in PWS System

A. Profile-Based Personalization:

This module basically produces the personalized page for the user based on the user information that will be provided to the personalized web search while registration. In order to perform this process of personalization, two mechanisms are used: a profile generator that automatically creates user profiles representing the user preferences, and content based recommendation algorithm that estimates the user's interest in unknown content by matching the profile to the detail descriptions of the content. Both features are put together into a personalization system.

B. Generalizing User Profile:

The generalization process is to generalize 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, use of references which will be stored in cache 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.

C. Online Decision:

Online decision mainly helps the user to provide the information about whether to personalize the profile for the given query. Here the online decision is made by the user after specifying the query, if the user thinks the query is personal than query will be sent to the server with limited information about the user profile. This will avoid the risk that may occur on the user's privacy.

D. Privacy Protection in PWS System:

PWS framework called UPS that can generalize profiles for each query according to user-specified privacy requirements. Here two methods will be used that evaluate the privacy risk and the query usefulness for user profile. And also provide an online prediction mechanism based on query utility for deciding whether to personalize query in UPS. UPS makes use of two algorithms namely GreedyDP and GreedyIL for runtime generalization. While the first algorithm tries to maximize the discriminating power (DP), the other attempts to minimize the information loss (IL).

5. LITERATURE SURVEY

Personalized web search (PWS) differs from normal web search, which returns identical research results to all users for identical queries, regardless of varied user interests and information needs. Personalized Web Search can be categorized into two types; one is click-log-based methods and other profile-based ones.

The Click log based methods are based on just selecting the clicked pages in the user's search history. Profile based method has more effectiveness in improving the quality of web search with increasing usage of personal and behavior information to profile its users, which is usually gathered completely from query history, browsing history, click-through data, bookmarks, user documents and so forth. Several research and implementations on personalized web search has been carried.

In [1] methods for personalization has been developed that uses online decision from the user before performing any personalization process on the queries provided. The approach will try to solve the previous problems of generating user profile offline and using this same generated profile for every query.

It mainly supports:

- Online profiling. This method creates separate user profile for each query that will be provided and hence improves search results.
- Considers customization of privacy requirement, and also supports personalization features.

This made a new challenge in terms of efficiency. In order to fulfill these features, the framework customizes user generalization on user-level and query-level which allows user to add their requirement and varies the generalization based on the query contents respectively.

In [2], proposed a UPS (User customizable Privacy preserving Search) framework in which the server creates user profile according to user specified privacy requirement and query content while submitting query. Then query along with this generalized user profile is sent together to PWS server. The results are personalized with profile and delivered back to server which then shows results to user.

Creating profile for the user process consists of collecting the data [3]. Profile Constructor and Algorithm which makes use of this information from the user profile in order to provide personalized services. User identification plays crucial role from the system point of view that constructs user profile. Many researches 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.

In [4], proposed a method, here when a given a query submitted to a search engine, suggests a list of related queries. The related queries are based in previously issued queries, and can be issued by the user to the search engine to redirect the search process. The method proposed is based on a query gathering process in which groups of related and similar queries which are identified. The method not only discovers the related queries, but also ranks them according to a relevance criterion.

6. EXISTING SYSTEM

The existing profile-based Personalized Web Search does not support runtime profiling. A user profile is typically generalized for only once offline, and used to personalize all queries from a same user indiscriminatingly. Such "one profile fits all" strategy certainly has drawbacks given the variety of queries. The existing methods do not take into account the customization of privacy requirements. This probably makes some user privacy to be overprotected while others insufficiently protected.

7. LIMITATIONS OF EXISTING SYSTEM

Some of the limitations of the existing systems are:

- The existing profile-based PWS do not support runtime profiling. A user profile is typically generalized for only once offline, and used to personalize all queries from a same user indiscriminatingly. Such "one profile fits all" strategy certainly has drawbacks given the variety of queries.
- The existing methods do not take into account the customization of privacy requirements. This probably makes some user privacy to be overprotected while others insufficiently protected.

8. PROPOSED SYSTEM APPROACH

Features of the proposed PWS system will improve its performance and quality of service. Performance of the proposed system will improve as in this system the user will be allowed to register with some more features and more information about his personal interest which will improve search engine performance. The system uses Privacy Preserving Personalized web search framework UPS, which can generalize profiles for each query according to user-specified privacy requirements. UPS makes use of two algorithms namely GreedyDP and GreedyIL for its runtime generalization. While the first algorithm tries to maximize the discriminating power (DP), other attempts to minimize the information loss (IL).

9. GREEDY ALGORITHM

Algorithm follows solving the problem by itself by making the locally optimal choice at each stage with the hope of finding the best possible solutions to the problem. Greedy algorithm considers easy to implement and simple approach and decides next step that provide beneficial result. In many problems, a greedy strategy does not produce the solutions that are near to the given problem, but its method of discovering itself give a way to locally optimal solutions that approximate a solutions near to the given problem in a reasonable time.

A. GreedyDP algorithm

It mainly works in a bottom up manner. It starts with leaf node in a tree and for every iteration, it chooses leaf topic for removing thus trying to maximize utility of output. During iteration a best profile-so-far is maintained satisfying the Risk constraint. The iteration stops when the root topic is

reached. The best profile-so-far is the final result. These algorithms require recalculation of profiles which adds up to computational cost and memory requirement.

B. GreedyIL algorithm

GreedyIL algorithm improves generalization efficiency. GreedyIL maintains priority queue for candidate prune leaf operator in descending order. This decreases the computational cost. GreedyIL states to terminate the iteration when Risk is satisfied or when there is a single leaf left. Since, there is less computational cost compared to GreedyDP, GreedyIL outperforms GreedyDP.

10. IMPLEMENTATION DETAILS

Implementation details contain the details of working when implemented on the system with the help of hardware and software requirements that are specified.

Process flow diagram of the project implementation:

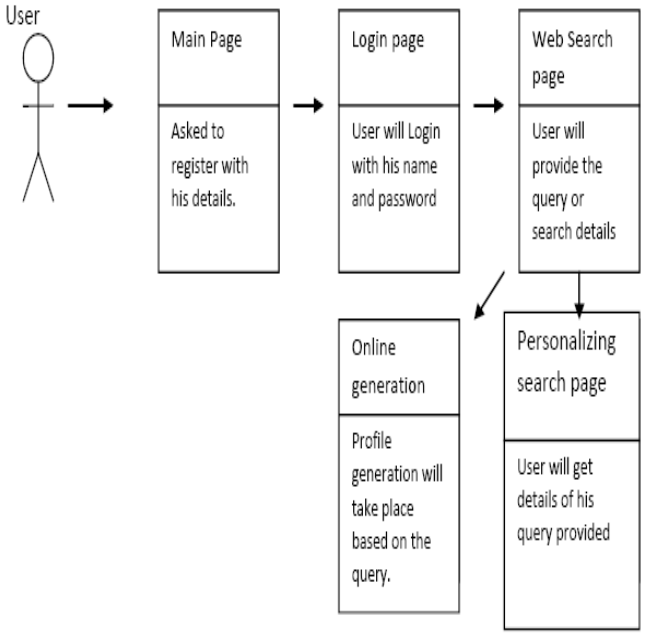

Fig -2 Process flow for PWS

It is capable of running on standard Net beans IDE tool. The interface for this system will provide a personalized web search experience for the Users with its privacy protecting policy.

11. CONCLUSION

As the resources on the web browser are increasing day by day this intern increases number of users and it also make the service engine to provide the relevant search result based on their behavior. Providing relevant result to the user is based on their click logs, query histories, bookmarks, by this privacy of the user might be loss. Most existing system provides protection for the private information during user search. The approaches that were developed provide the protection about these were lacked in features related to the user profile and his information. The proposed system will provide features to the user profile and make it easy and more interesting to the user profile during his search process based on his interest and also saves time during search.

12. 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. 615- 624, 2011.

[3] Susan Gauch, Mirco Speretta, Aravind Chandramouli, and Alessandro Micarelli, "User profiles for personalized information Access".

[4] Ricardo Baeza-Yates1, Carlos Hurtado1, and Marcelo Mendoza on "Query Recommendation using Query Logs in Search Engines," Proc. Millennium Nucleus, Center for Web Research (P01-029-F), Mideplan, Chile.