

# Supporting Privacy Protection in Personalized Web Search

**Bhavana Bhalerao<sup>1</sup>, Prof. Pankaj Kulkarni<sup>2</sup>**

<sup>1</sup>Research Scholar, Computer Engineering, , Nasik, India.

<sup>2</sup>Lecturer, Dept. of Computer Engineering, Shatabdi Institute of Tech. (Polytechnic), Nasik, India.

bhavana\_bhalerao@yahoo.com, pankajkulkarni1985@gmail.com

**Abstract** –*Personalized Web Search (PWS) has defined a great effectiveness for users to retrieve the useful information for them quickly specially according to their interest (which are stored in database). Personalized search results are sorted and are arranged according to the priority given by the users so that after firing the same query for the next time user can obtain the required result more conveniently and quickly. Also unnecessary data is avoided and relevant data is displayed. Users are provided more convenient services by providing them creating their own account feature. This helps Personalized Web Search to search user's query efficiently. This paper describes the design and implementation of Personalized Web Search based on a specific field (here Books).*

## I. INTRODUCTION

As the importance of information age arrived Internet enabled people to access this information very easily but due to the sudden increase in today's information age knowledge, the search engine have become more important in day today's life.

Web search engines have become popular among the users for obtaining the required information from on the web. Along with the results, sometimes user experienced failure too. As these search engines fetches the information from a great number of sources, whenever the user enters a query on a general web search engine, along with the necessary data or information a lot of unnecessary stuff, which does not meet the real intention of user interest, are also returned.

So, to avoid this unnecessary exposure of necessary data or information Personalized Web Search is introduced. Personalized Web Search is specially designed to provide user better search results which focus on the user's interest. Surfing between the unnecessary stuff for obtaining the appropriate data or information consumes a lot of time of user which results in user irritation. Personalized Web Search overcomes this main drawback of general web search engines.

The aim of Personalized Web Search is to filter the useful data or information for user by using user's interest, which is mentioned in user account. Personalized Web Search will sort the result according to the user's interest priority, so it is very convenient for user to pick useful data or information.

In this paper we will introduce the design and implementation of Personalized Web Search.

## II. LITERATURE SURVEY

During literature surveys, many researchers investigated many search engines and thus discovered that these search engines are still ineffective. Also these search engines made users to compromise with their personalized details privacy. Researchers also discovered that many personalized web search needed installation of some proxy server or the bot.

Some of the surveys done by other researchers are as follows:

M. Spertta and S. Gach, surveyed that user profile should be of user interest that can be used by the search engine to result in personalized form.

B. Tan, X. Shen and C. Zhai studied that the existing search engines contain history about user search likings which can be used to provide enhanced results.

L. Shou, He Bai, Ke Chen and Gang Chen concluded that 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. For example, all the sensitive topics are detected using an absolute metric called surprisal based on the information theory, assuming that the interests with less user document support are more sensitive.

In general, users have to compromise between their account privacy and search quality.

## III. EXISTING SYSTEMS

Now-a-days many web search engines exist. But it also includes some drawbacks. Examples of existing

systems can be given as: Google, Yahoo, Bing, Ask etc.

These search engines have a common fact among their results and it is that they all return a bunch of results which includes the required data as well as unwanted stuff due to which user has to deal a lot of time. For example, if a user wants to search details about “cookies” (which is related to IT) he/she gets a lot of pages which contains the required data as well as unwanted data like details about “cookies” (which relates to recipe). Though the user gets the detail that he/she wants but the user will have to surf through many pages which is very time consuming.

Also if the same query is fired again the result is same as the previous result. The order of the link or web page is also the same as the pages are ranked according to their popularity and not based on the user priority. Hence, the provided results are based on clicks or hits made by the other users and which sometimes acts as a barrier between the users and the relevant result.

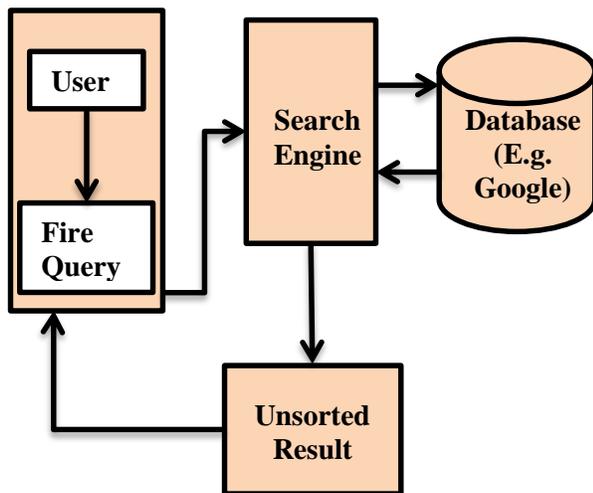


Fig III.1. Basic Structure of Existing System

For example, if the page or link that user needs is on the 4<sup>th</sup> or 5<sup>th</sup> page then the user will have to surf till 4<sup>th</sup> or 5<sup>th</sup> page. As well again for the next time user will only be able to access his last page if he/she have bookmarked it else again user will have to surf through the whole stuff again.

#### IV. PROBLEM DEFINITION

Users have to deal with information that is unnecessary. Hence it creates confusion to select the appropriate result which is needed by the user. Along with it as the users have to search a number of pages to search the accurate result a lot of time is consumed.

One of the major insufficiencies of the previous search engine is that they follow the model of “One Size Fit All”. This is not adaptive to individual user.

The Solution for the above problem is our proposed system i.e. “Personalized Web Search.”

#### V. PROPOSED SYSYTEM

We proposed a search engine which has been personalized that focuses on user profile. This search engine provides a generalized result according to user’s interest in prioritized manner. Along with the user satisfaction it also maintains the privacy of user profile. The result of this personalized search engine meets the complete requirement of the end user & maintains a high quality of this search engine.

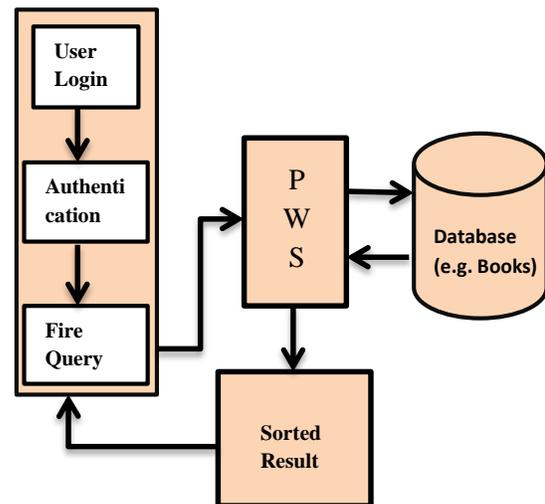


Fig.V.1. Proposed System Architecture.

From the above architecture of proposed system i.e. Personalized Web Search we can understand the working of our proposed system easily.

Here PWS (Personalized Web Search) is the main component in the working of our proposed system. As the other components are the supporting modules of the architecture.

In its initial stage user will be asked to login and the process of authentication will be carried out by the

system. If the user is authenticated then he/she can fire the query. This query is forwarded to our PWS Database.

Once the results are obtained from the database, it is provided in sorted manner which helps the user to determine appropriate result in a short period of time.

Therefore the Personalized Web Search provides the results based on user interest and priority. Hence improving the search quality with the personalization utility of the user profile.

We used two simple but effective generalization algorithms, GreedyDP and GreedyIL, to support runtime profiling. 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.

## VI. SYSTEM DESIGN

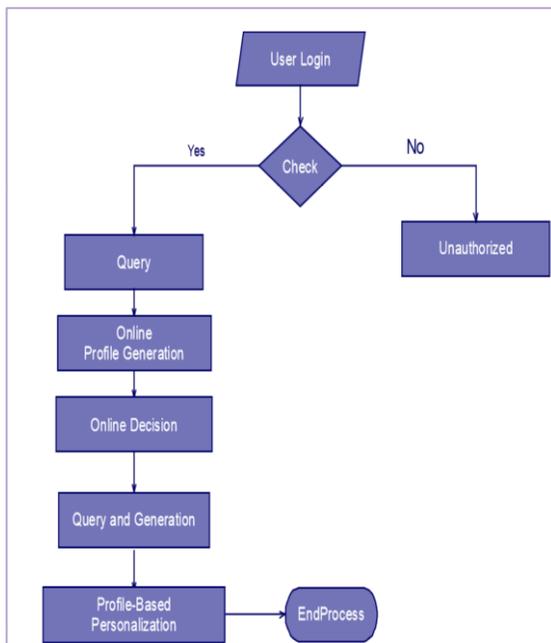


Fig.VI.1. DFD for User

The Data Flow Diagram (DFD) is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of the input data to the system, various processing carried out on these data, and the output data is generated by the system.

## VII. GREEDY ALGORITHM

A greedy algorithm is a mathematical process that recursively constructs a set of objects from the smallest possible constituent parts. Recursion is an approach to problem solving in which the solution to a particular problem depends on solutions to smaller instances of the same problem.

Greedy algorithms look for simple, easy-to-implement solutions to complex, multi-step problems by deciding which next step will provide the most obvious benefit. Such algorithms are called greedy because while the optimal solution to each smaller instance will provide an immediate output, the algorithm doesn't consider the larger problem as a whole. Once a decision has been made, it is never reconsidered.

The advantage to using a greedy algorithm is that solutions to smaller instances of the problem can be straightforward and easy to understand. The disadvantage is that it is entirely possible that the most optimal short-term solutions may lead to the worst long-term outcome.

Greedy algorithms are often used in ad hoc mobile networking to efficiently route packets with the fewest number of hops and the shortest delay possible. They are also used in machine learning, business intelligence (BI), artificial intelligence (AI) and programming.

## VIII. CONCLUSION

Thus this paper presents prioritized and personalized results to the use provided by the Personalized Web Search.

This system allows user to obtain the relevant results in shortest period of time by using the Greedy Algorithm.

This paper reveals that Personalized Web Search could achieve quality search results while preserving users customized privacy requirements. Also these results are effective and achieved with great efficiency.

In simple words, the accuracy and quality of the search is improved, which leads to better search results are less time.

## 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] M. Spertta and S. Gach, "Personalizing Search Based on User Search Histories," Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence (WI), 2005.
  - [3] ] 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.
  - [4] X. Shen, B. Tan, and C. Zhai. Implicit user modeling for personalized search. In Proceedings of CIKM 2005, pages 824–831, 2005.
  - [5] Personalised Search engine design and Implementation. Jiandong Cao, Yang Tang, Binbin Lou, 2010, IEEE.
  - [6] Automatic identification of user interest for personalized search, Feng Qiu, Junghoo Cho, MAY 2006, uk.
  - [7] Supporting privacy protection in personalized web search, Lidan Shou, He Bai, Ke Chan, And Gang Chen, 2014