Personalized search engine

[1] Pawar Uddhav G, [2] Bhandari Renuka G, [3]Gaikwad Snehal K , [4]Mali Vaibhav B., [5]Prof.A. V. Patil

Department of Information Technology, Zeal College of Engineering & Research, University of Pune.

Abstract- In this research of search engine, area of interest of user plays important role. As different user have different search goals during firing a query to the search engine. A major problem with existing search engine is that user search queries are usually short and single word different meaning, and thus are insufficient for specifying users need. Effective personalization cannot be achieved without accurate user profile by using traditional search engine. Traditional systems have not implemented the effective user specific domain approach for personalized search engine. In this project work we have proposed the new user specific domain approach by studying user's behavior while submitting a query. IN our proposed work we have taken user click through data to obtain the user interest in specific domain. User clicks through data is re-rank the search result obtained. An example of click-through data for the query "virus" which have different meaning as it may be computer virus, bacteria, etc, so query is ambiguous. We get user interest when user click on any link recommendation provided by our search engine so according to clicked recommendations we have provided web page of that recommendation, interested recommendations related to clicked recommendation, non-interested recommendations and also previous 10 day user history on same window .

KEYWORDS- **Information Retrieval, Ranking, Click through, Query clustering.**

1. INTRODUCTION

The web provides an extremely large and dynamics source of information, and the continuous creation and updating of web pages magnifies information over-load on the web server. Both casual users often use search engines to find a needle in this constantly growing haystack. ,who define a knowledge worker as someone whose paid work involves significant time spend in capturing, finding, analyzing, creating, producing or achieving information, report that of the task performed on the web by sample of knowledge workers fall into the categories of information capturing and finding, which require an active use of web search engines. The identical query from different users or in different contexts will generate the same set of result displayed in the same way for all users, a so called one-size-fit-all approach. Furthermore, the number of search results returned by a search engine is often so large that the results must be partitioned into multiple result pages. In the addition individual differences in information need, multiple meaning of the same word and multiple words with same meaning pose problem in that a user may have to go through many irrelevant result or try several queries

before finding when the desired information. The organization can then use such information as an input to a system based on our proposed approach. For a given query, a personalized search can provide different result for different user organizes the same results differently for each user. Processing includes result filtering, such as removal of some results and keyword matching, such as re-rank, clustering and categorizing the results.

2. EXSTING SYSTEM

In a web based application, queries are submitted to the search engines to represent information needs of the user. Yet some queries may not exactly user specific information needs since it is ambiguous and may cover a broad topic and different user may need to get information on different aspects when they submit the query. For e.g. when the query the sun is submitted to the search engine, some user are need of Indian Newspaper, some are need of natural language about sun. Hence it is necessary to capture the goal of different user in the case of information retrieval.

Disadvantage
-Recommendations are not filtered based on user interest.
-Time consuming i.e. every time we need to go on back page to see other recommendations.

3. AIM AND MOTIVATION

Our proposed work aims to develop online techniques concepts from the web-snippets of the search returned from a query and use the concepts to identify related queries for for that query.

We propose a framework for personalized web search which considers individual interest into mind and enhances the traditional web search by suggesting the relevant pages of his/her interest and also we are proving non-interested recommendation because user interest may change simultaneously. We have proposed a simple and efficient model which ensures good suggestions as well as promises for effective and relevant information retrieval. In addition to this, we have implemented the proposed framework for suggesting web pages to the user and also maintain user history.

4. PROPOSED SYSTEM

In proposed system create a user profile and monitor search session any time and also from anywhere. The advantages of our proposed system wad we can change the interest of user dynamically, user based search result will be get

focused first, URL ranking in takes place search history will be worldwide process. We are providing window that divide in four blocks.

1 User clicked webpage (Where the user has gone).

2 Interested recommendations (related links of user clicked links).

3 Non interested recommendations (where user is not interested to go).

4 User history (user history is according to user clicks)

If user click on non interested link then we find related links of that links and provide links in interested recommendations block and move the non interested links in non interested recommendations block, user history with modified recommendations and also webpage of user clicked.
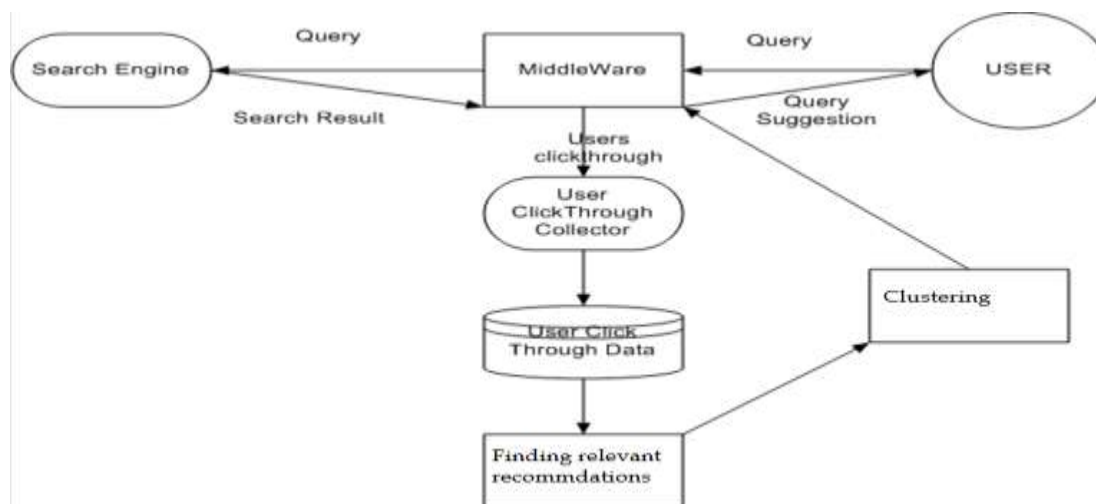
Advantages

-Time saving User gets all the recommendation on same page, so there is no need to go on back page to see other recommendations.

-Recommendations are filtered based on user interest.

5. ARCHITECTURE OF SYSTEM
The below figure shows the architecture of proposed method, in this method we have considered user click through to get user interest. As we are taking user interest at runtime, we can build the accurate user profile. User click through are then collected to get the user click through data. Using this data user's conceptual profile is build. After getting the accurate profile a similarity based clustering mechanism is applied for the query suggestions to the user. And also maintain user history.

Fig 1: Architecture of proposed system

6. ALGORITHM OF KEYWORD MATCHING

Step 1:-Store unused keywords in separate string array1

Step 2:-Copy all words of description in user clicked recommendation in string array2.

Step 3:-Compare the unused keywords in string array1 with string arry2.If there is any match occur then remove such keywords, blank spaces from array2.

Step 4:-Copy all words of description in recommendation of remaining recommendation in string arry3.

 Step 5:-Compare the keywords of array2 with array3.

Step 6:-Cluster the matched recommendations and unmatched recommendations.

Step 7:-If recommendation remaining then go to step 4 otherwise go to step 8.

Step 8:-Provide interested recommendations as matched recommendations and non-interested recommendations as unmatched recommendations to user and add clicked recommendation to user history.

Step 9:-If user clicked on any recommendation then go to step 2.

Example:-

- **Link1:** ← User clicked recommendation

    Apple, mango, banana.

- **Link2:** banana, guava.

        Score =1.

- **Link3:** apple, orange, mango.

Score=2.

- **Link4:** pineapple.

    Score=0.

Link 1 Webpage	**Interested links** 1.link 3(Score =2) 2.Link 2(Score=1)
	Non-interested links 1.link4(Score=0)
	User History 1.link1

Table 1. Graphical representation of links in proposed system

7. CONCLISION

Thus, we have developed personalized search engine based on Google API by removing some drawbacks of existing search engine and focus on each user personally. So user can get all the interested links, his/her history with web page on same window and works efficiently with speed as possible.

REFERENSES

[1] Thorsten Joachims Cornell University Department of Computer Science 2005.Optimizing Search Engine using clickthrough data.

[2] Aniko Hannak and Arash Molavi Kakhki Northeastern University 2007. Measuring Personalization of Web Search.

[3] Wai-ating Leung and Waifred NG, Dik LUN Lee 2008. Personalized Concept-Based Clustering of Search Engine Queries.

[4] Bhupesh Gupta, Sandip Goyal and Ashish Oberois CSE deptt. MM University Mullana India 2012. A Review on Query Clustering Algorithm for Search Engine Optimization.

[5] S. Preetha, PG Student, and Vimal Shanker, Assistant Professor 2012. Personalized Search engine on mining user preference using clickthrough data