Proceedings of
National Conference on Innovative Trends in Computer Science Engineering (ITCSE-2015)
held at BRCMCET , Bahal on 4th April 2015

A Review on Personalization Techniques

[1]Ankur Mittal , [2]Shilpa Sethi
[1]Research scholar , [2]Assistant professor
[1,2]YMCA University of Science & Technology , Faridabad, Haryana
[1]ankur89mittal@gmail.com , [2]munjal.shilpa@gmail.com

ABSTRACT- The World Wide Web is commonly known as "web" is a huge collection of interlinked billions of HTML documents. The automated searching tools such as search Engines are used to retrieve information from such a huge collection on web. Although the present search engines are using sophisticated ranking and indexing algorithms, but they still provide a long list of documents, most of which are not relevant to the user's need. One of the reason behind this is each user has different requirement for information search. So, it becomes vital in many areas to consider the user interests and preferences for the search engine while retrieving and maintaining its database. Personalization is taken as one of the method to provide relevant information according to the user's need. This paper conducts a survey of how personalization can provide useful knowledge to the user. Several user personalization approaches and techniques are illustrated in this paper.

KEYWORDS Information retrieval, search engine, quality information, User Preferences, profile, Personalization

I.INTRODUCTION

The World Wide Web (WWW) is the most interactive world where people share information via Internet. There has been an incredible growth of the Internet in last decade. Ten years ago we were in the era of the dot-com boom (and bust), the Web was strictly 1.0, and Google was just a baby. Since then people have actually welled onto the Internet making it a Global Village. But, in many respects the World Wide Web has become a victim of its own success. The rapid growth in the volume of available information is making it difficult for the search engine to fulfill the user's need. One of the reason behind this is traditional search engines retrieve the information based upon the keywords present in user query and no attention is given to the context in which user is searching. For example if a user submit a query "Information about Pluto" then search engine must be able to provide information about Pluto planet if the previous searches of user are focused on solar system and on Pluto cartoon, if the previous search are more focused on entertainment.In such cases personalization can help to find valuable information and resolve the problem, Web personalization is the system where the information is retrieved according to user interest, past browsing histories, preferences etc. This paper conducts a survey of several web personalization techniques. The paper is organized as follows. Section 2, provides basics of Web Personalization. Section 3, describes web personalization techniques and conclusions are given in Section 4.

II. WEB SEARCH PERSONALIZATION

Web Personalization can be defined as a process of helping users by providing customized or relevant information on the basis of web experience to a particular user or set of users [2]. So, Web Personalization means customizing information retrieval according to the needs of users. Web search is effective only when it is both context and interest-oriented. The factors that affect web search are : queries submitted by the user, query usage time page visited ,semantics between the search query and visited pages, page view time, sequence of the page access, context of search , relation between context of search and the information currently examined, , user interest, user's shift in interest.

Out of these factors relation between search query and page visited by the user, user actions on clicked page and time spent on web page will prove to be the most important factor of personalization [6].

The benefits of web personalization are:

1 **Save Time**: - By sorting the retrieved results, according to user's interest fulfils information need in lesser time.

2. **Better Information quality**: - By using the user profile and history search engine provide more relevant information to user.

3. **Save money**: - By preventing redundant work, lowers the cost to meet the solution.

Number of web search personalization techniques are developed which are aiming to provide the relevant result to the users. But not all the personalization techniques are applicable to all the users at the same time. Because users preferences may change with time so, user profile must be created based upon short term interest and long term interest [11]. For example, it may be possible that a user is interested in sports when cricket world cup is going on; otherwise its most searches are focused on share market. So, a sport is its short term interest and share market is its long term interest. The study of various personalization techniques are given in next section.

III. SEARCH ENGINE PERSONALIZED TECHNIUES

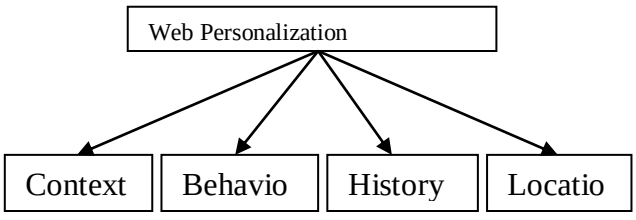

Fig 1: Classification of Personalization

In past few years, many web personalized search system has been proposed and developed based on different factors. But they all require gathering information about the user for providing the most relevant results while searching. Depending upon these different factors the personalization techniques can be broadly classified into four categories as shown in fig 1:

Each of these techniques is discussed one by one in following sections:

3.1 Context oriented personalization

As the users of search engine are from different backgrounds, and not expert in constructing the queries, so most of the time the context of user's search is not clear from the short query submitted by the user at search engine interface. So, the idea behind the context oriented personalization is to apply the context of user search while providing the results to the user. Several researches have been conducted in this area. It is observed that the idea of context can be applied at two different points:

(i) At the level of query submission
(ii) At the level of ranking the results

The work in [8] shows that if context of query is filtered out at the level of query submission, then the right query will select the right documents. It uses WordNet 3.1 to find out different synonyms of query keywords to form alternate queries. The user is required to select one of the alternate options provided by the search system.

The works in [12] apply the context of query while matching the query keyword with web document present in search engine database. So, all the documents having synonyms of user query also get selected which otherwise in traditional search systems were missing.

3.2 Behaviour based personalization:

Some personalization system focuses on capturing the behaviour of user such as short term and long term interest [7]. So, here personalization is based on the user behaviour this approach is different from previous one. [7] Determine the user short term interest by using the concept of sliding window and long term interest by applying the forgetting factors to the interest which have not changed for a long time.

According to [7] a interest factor is computed as given in equation (1) as follows:

Interest factor= a (short interest) + b (long term interest) --- (1)

Where:

a+b=1; and the value of a is determined by a unit step function as given by equation (2).

μ_a= { >0.5 if short <0.5 if long ----

3.3 History Based Personalization

Personalization when focus on the previous history of user is referred as history oriented web personalization. History based approach uses the factors such as user query logs history, pages visited by user , action performed on that page, time spent on that page etc.In recent years Google also developed the personalized and customized technique for users for better search result. Which record the user browsing history of user while user sings in and surfing internet. For the singed out user Google enables the customization search result based upon their search activity linked to cookie saved in browser [10]. It is completely separated from the signed-in user. Google save the user history no of page clicked and viewed. [6] Proposed the personalized search system which is based on user previous history. Here system build the user profile when user surfing the internet here it record the number of query submitted by user, user id, and page clicked etc. Here some information is explicitly collected. Previous history is used to re–rank the search results. System keep records of user previous browsing history Such as action (print, save, bookmark etc.) and time spent on each web page. By using theses no. of factors, more relevant results are obtained.

3.4 Location Based Personalization

Location based approach is distinguished from other types of personalization in the sense it refine the user original query according to user's current location. For example, if the user search for the "Haldiram" and user current location is "Faridabad", then modified search query is "Haldiram Faridabad" according to current location of user keeping in mind that may be user searching for near current location. Result is given to user as a option to original query.. Here some time it not useful so much and it have some disadvantages.

Mohamed [9] suggest a model for web personalization based on current location of users. Author here used the Location – based services (LBS) 2.0, in which personalization, socialization, and recommendations have the important role in location based services. Location based services is deals with Web 2.0 where content sharing, social networks all becomes common for all web services. Location based provides the nearest neighbour queries to more personalized best neighbour queries where user preference and context are taken under consideration to answer the query. Here location information are extra with query to provide better result.

Some authors had also proposed mix and match of above techniques. For example Qian Gao [7], proposed a method using the four agent in the system, which all communicate each other, each have own responsibility. Agents are client agent, feature word extraction agent, personalized ontology profile agent, user preference profile construction agent. In method first apply the word stemming to extract stemmed word, from these stemmed word extract the feature word by applying similarity matching on stemmed word and ontology profile. By using the feature word user preference profile construction agent construct the user preference profile. After that we create the user long term preference by using the user browsing behaviour, feature words and time. Here to create long term preference used the dynamic update strategy called Time-based-forgetting functions update strategy.

Zhengyu Zhu et al., [5] proposed query expansion based on a personalized web search model. It depends on a representation of personalized web search organization. The novel system, as a middleware connecting a user and a web search engine, is fixed on the client machine. It can study the

user's favourite implicitly and then produce the user profile automatically. When the user enters query keywords, more personalized expansion words are produced by the proposed approach, then these words in common with the query keywords are forwarded to a famous search engine such as Google. These expansion words can facilitate search engine retrieval information for a user based on his/her implicit search objectives. The novel web search representation can build an ordinary search engine personalized, specifically all the way through personalized query expansion the search engine can provide different search results to different users who enter the equivalent keywords. The experimental observations demonstrate the consequence and use the proposed work for personalized information service of a search engine. [5]

J. Saravanakumar [8] proposed a system which based on the semantic information retrieval. When query given by user first query selection process generate the alternate query by query pre processing, word stemming. Then result of main query and alternate query is compared according to weight assigned each result and ranked accordingly. After that on user selection assign new weights and re-ranked according the new weight. By the assign new weight achieves the personalization. This system used in finding large amount of information. [8] The comparison of above discussed techniques is given in table 1.

Techniques / Comparison terms	Context Oriented	Behaviour Oriented	Location Oriented	History Oriented
Factors	Synonyms Polysemy of keywords in user query	Interest area of user	Geographic location	Click(u_i,q_i,p_i)*,Action(p_i) ** Time spent(u_i,q_i,p_i)***
User involvement	Inferred automatically by the system(implicit)	May be implicit or explicit	Inferred automatically by the system(implicit)	Inferred automatically by the system(implicit)
Nature of success	Simple, easy to implement.	More relevant results but user rarely interested in giving feedback	Easy to implement but sometimes misleading results	Complex to implement but provides relevant results
#Search Engines	Google Yahoo	Google Bing Yahoo	Google	Google Bing

Table1: Comparisons of web search personalization techniques

*Click (u_i,q_i,p_i) : proportion of no. of clicks made by user ui on web page P_i for query q_i w.r.t total no. of clicks made by user ui on all the pages for query q_i.

**Action (p_i): denotes the action performed on the page p_i.

***Time spent (u_i,q_i,p_i): time spent by the user u_i on the web page p_i.

Search engine: name of search engine using the personalization techniques.

IV. CONCLUSION

A web personalization information retrieval system provides the more relevant results to user as compared to non personalized search system. In this paper we reviewed several personalization techniques and come to the conclusion that although many personalization techniques had been proposed and personalization systems had been developed but hybrid approach of these techniques can provide the better results than considering only individual approach.

REFERENCES

[1] M. Albanese, A. Picariello, C. Sansone, and L. Sansone,"Web personalization based on static information and dynamic user behavior," 6th annual ACM international workshop on Web information and data management, WIDM '04. New York, NY, USA: ACM, 2004, pp. 80–87.

[2] K.Sridevi, Dr. R.Umarani, Web Personalization Approaches: A Survey, International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 3, March 2013

[3] Zeeshan Khawar Malik, Colin Fyfe," Review of Web Personalization" Journal of Emerging Technologies in Web Inttelligence,Vol. 4,No. 3,August 2012

[4] Chao Xu , "Task-Based User Profiling for Personalized Query Refinement" Department of Information System. New Jersey Institute of Technology

[5] Zhengyu Zhu, Jingqiu Xu, Xiang Ren, Yunyan Tian and Lipei Li, —Query Expansion Based on a Personalized Web Search Model, Third International Conference on Semantics, Knowledge and Grid, Pp. 128 – 133, (2007).

[6] Shilpa Sethi, Ashutosh Dixit, "Design of Personalised Search System Based on User Interest and Query Structuring" , International Conference on "Computing for Sustainable Global Development" 2015

[7] Qain Gao, Su Mei Xi, Young Im Cho, "A Multi Agent Personalized Ontology Profile Based User Preference Profile Construction Method", 44th International Symposium On Robotics, 2013

[8] K. Saravanakumar, K. Deepa, "Alternate Query Construction Agent for Improving Web Search Result using WordNet", International Conference on Computational Intelligence and Communication Systems, 2011

[9] Mohamed F. Mokbel, Jie Bao Ahmed, Mohamed Sarwat, "Personalization, Socialization, and Recommendations in Location based Services 2.0"

[10] Hamid Rastegari,Siti Mariyam Shamsuddin, "Web Search Personalization Based on Browsing History by Artificial Immune System", Int. J. Advance. Soft. Comput. Appl., Vol. 2, No. 3, November 2010

[11] Shilpa Sethi, Ashutosh Dixit, "Personalization: An Exploration Into The New World 0f Searching"

[12] Payal Gulati, Dr. A.K. Sharma "Ontology Driven Query Expansion for Better Image Retrieval)" International Journal of Computer Applications (0975 – 8887)Volume 5– No.10, August 2010