Personalized Web Search Techniques - A Review

By T. Sathiyabama, Dr. K. Vivekanandan

Bharathiar University, Coimbatore

Abstracts - Searching is one of the commonly used tasks on the Internet. Search engines are the basic tool of the internet, from which related information can be collected according to the specified query or keyword given by the user, and are extremely popular for recurrently used sites. With the remarkable development of the World Wide Web (WWW), the information search has grown to be a major business segment of a global, competitive and money-making market. A perfect search engine is the one which should travel through all the web pages in the WWW and should list the related information based on the given user keyword. In spite of the recent developments on web search technologies, there are still many conditions in which search engine users obtain the non-relevant search results from the search engines. A personalized Web search has various levels of efficiency for different users, queries, and search contexts. Even though personalized search has been a major research area for many years and many personalization approaches have been examined, it is still uncertain whether personalization is always significant on different queries for diverse users and under different search contexts. This paper focuses on the survey of many efficient personalized Web search approaches which were proposed by many authors.

Keywords: Personalized Web Search, Ontology-based Multi-Facet (OMF), Personal Agent for Web Search (PAWS).

GJCSST Classification: H.3.3, H.3.5
Personalized Web Search Techniques - A Review

T. Sathiyabama¹, Dr. K. Vivekanandan²

Abstract - Searching is one of the commonly used task on the Internet. Search engines are the basic tool of the internet, from which related information can be collected according to the specified query or keyword given by the user, and are extremely popular for recurrently used sites. With the remarkable development of the World Wide Web (WWW), the information search has grown to be a major business segment of a global, competitive and money-making market. A perfect search engine is the one which should travel through all the web pages in the WWW and should list the related information based on the given user keyword. In spite of the recent developments on web search technologies, there are still many conditions in which search engine users obtains the non-relevant search results from the search engines. A personalized Web search has various levels of efficiency for different users, queries, and search contexts. Even though personalized search has been a major research area for many years and many personalization approaches have been examined, it is still uncertain whether personalization is always significant on different queries for diverse users and under different search contexts. This paper focusses on the survey of many efficient personalized Web search approaches which were proposed by many authors.

Keywords : Personalized Web Search, Ontology-based Multi-Facet (OMF), Personal Agent for Web Search (PAWS).

1. Introduction

Search engines present a well-organized way to search the relevant information from the Web. However, the search results acquired might not always be helpful to the user, as search engine fail to recognize the user intention behind the query. A particular word could mean many things in dissimilar contexts and the anticipated context can be determined by the user alone. For illustration, specified a search keyword "apple", a user might be searching for fruit apple or for apple computers. A typical search engine provides similar set of results without considering of who submitted the query. Therefore, the requirement arises to have personalized web search [1] system which gives outputs appropriate to the user as highly ranked pages. With the growth of Internet, people are becoming more and more dependant on the Web search engines for their various information needs. In spite of the extensive use, there are still several challenges for search engine. Particularly, when queries are entered to a search engine, similar results are returned to different users. To avoid this difficulty, personalized Web search [1, 16] has been developed. In personalized search, how to efficiently acquire user’s real-time information requirement is a key issue. User’s query inputted to the search engine is the most important source of evaluating information need. However, the query has the characteristics of shortness, ambiguousness and incompleteness which control the clear expression of user’s information requirements and thus influences the qualification for personalized search. Consequently, it is far from sufficient to achieve user’s requirement only from the query.

The most common problem of most Web search systems is that they do not think about the differences among individual user needs. The result of the search process relies exclusively upon the query. The underlying search engine presents the same search results for different users when they submit the same query, not considering about their different information needs and preferences.

Some information systems planned to understand such adaptive systems have been developed that personalize information or present more appropriate information for users. Three categories of Web search systems present such information: (a) systems that make use of relevance feedback, (b) systems in which users register their interest or demographic information, and (c) systems that suggest information according to users’ ratings. In these systems, users are required to register personal information for instance, their interests, age, and so on, or users are required to give feedback on relevant or irrelevant results, ratings on a scale basis from 1 (very bad) to 5 (very good). This is time consuming and hence users desire to use some other easier methods. To avoid these difficulties, some new approaches have to be proposed to reduce the difficulties of users.

In most of these systems and techniques, however, only users’ static information, like interest, information [17], such as feedback, browsing, etc., are utilized to carry out personalized web search,

Author ¹ : Research Scholar, Department of Computer Science and Engineering, Bharathiar University, Coimbatore. sathiyamaha@yahoo.co.in
Author ² : Reader, BSMED, Bharathiar University, Coimbatore. vivek_bsmmed@gmail.com
and the situation/circumstance of users are not taken into account. On the other hand, the user’s information needs completely depend on his situations/circumstances, so such information should be considered to realize personalized web search.

Through the application of high-speed internet, high capacity networks, and highly developed interactive websites, like Facebook, Youtube and blogging, searching has grown to be much easier for the internet users to publish data over the Web. With this information flooding era, it has become very difficult for a user to recognize the right information over the Web.

Personalized Web search is to perform retrieval for each user by taking his/her own situations/circumstances into account. As the competition in search market grows rapidly, some search engines have presented the personalized search service. For instance, Google’s Personalized Search permits users to specify the Web page categories of interest. Some Web search systems use relevance feedback to process user requirements or inquire users to register their demographic information earlier so as to present better service. Since these approaches require users to engage in extra activities to specify their preferences manually further than search, approaches that are capable of implicitly recognizing users’ information needs should be developed. Since the need for Personalized Web search is increasing, many researches have to be done to provide the relevant information by considering the users situations. The next section describes the various Personalized Web search approaches proposed by many authors.

II. Literature Survey

Nauman et al., [1] used personalized web search for enhancing common sense and folksonomy based intelligent search systems. A huge division of the contemporary web is characterized by user generated content classified using collaborative tagging or folksonomy. It makes very tricky to search for appropriate content because of ambiguity in lexical illustration of concepts and variances of preferences of users. With additional services relying on tags for content classification, it is significant that search approaches progress to better suit the scenario. A promising technique in avoiding these difficulties is to use machine common sense in combination with folksonomy. A past effort to use this technique has shown encouraging results in obtaining relevant content but it does not deal with the issue of noise in search results. In this paper, the authors make use of the personalized web search approach of conventional web demographic information, etc., and their web activities search systems to concentrate on the issue of irrelevant search outcomes in common sense and folksonomy dependent search systems. In personalized web search, outcomes are reflective of user’s favorites, which are based on the search history and kind of interest shown by the user. This paper proposes alterations to personalized web search approach. Using this personalized approach, the authors extend the fundamental common sense and folksonomy dependent search systems to deal with the problem of noise in search results.

Zhengyu Zhu et al., [2] proposed query expansion based on a personalized web search model. An innovative query expansion approach is proposed in this paper. 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 favorite implicitly and then produce the user profile automatically. When the user enters query keywords, more personalized expansion words are produced by the proposed approach, and then these words in common with the query keywords are forwarded to a famous search engine such as Baidu or 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 of the proposed work for personalized information service of a search engine.

P. Palleti et al., [3] developed personalized web search using probabilistic query expansion. The Web comprises of enormous amount of data and search engine recommends professional ways to assist navigate the Web to obtain the relevant information. Most common search engines, on the other hand, provide query results without taking user’s purpose behind the query. Personalized Web search system intends to offer relevant outputs to users by considering user interests into account. In this approach, the authors developed a personalized Web search system applied at proxy which changes to user interests perfectly by generating user profile with the use of collaborative filtering. A user profile basically consists of probabilistic correlations among query terms and document terms which are utilized for providing personalized search results. Experimental outcomes prove that this proposed personalized Web search system is very effective and efficient.

Jie Yu et al., [4] suggested mining user context based on interactive computing for personalized Web search. Personalized Web search is a successful way of same query. How to achieve user’s real-time information requirement is a key subject in personalized search. Existing approaches focus more on constructing user profile which depends on Web pages/documents which influences the effectiveness [5] of search engine.
Additionally, dynamics of user profile is frequently ignored. To deal with this problem, the authors have introduced a technique that acquires the user context to perfectly present preferences of users for successful personalized search in this paper. Initially, small-term query context is created from Web-snippets to take part a role of semantic background of user’s search behavior, recognizing associated concepts of the query. Then, user context snap is constructed depending on query context based on user’s interactive search behavior. Finally, development of user context is taken into account by introducing forgetting factor to combine the independent user context snap in a user session. The experimental outputs completely reveal that this technique can effectively construct user context based on individual user information need.

Fang Liu et al., [5] recommended personalized Web search for improving retrieval effectiveness. Modern Web search engines are generated to serve all users, independent of the particular requirements of any individual user. Personalization of Web search is to perform retrieval for every user according to his/her interests. In this paper, the authors propose a novel approach to learn user profiles from users’ search histories. The user profiles are then utilized to enhance retrieval efficiency in Web search. A user profile and a common profile are studied from the search history of the user’s and a category hierarchy, respectively. These two profiles are integrated to map a user query into a group of categories which corresponds to the user’s search intention and provide a context to disambiguate the words in the user's query. Web search is performed according to both the user query and the group of categories. A number of profile learning and category mapping approaches and a fusion algorithm are presented and evaluated. Experimental outcomes show that this approach to personalize Web search is very effective.

Xuwei Pan et al., [6] proposed context-based adaptive personalized Web Search for improving information retrieval effectiveness. Recent Web search/meta search engines are constructed to serve all users, independent of the particular requirements of any individual user in dissimilar situations. Personalization of web search is to perform retrieval for each user integrating his/her interests. In this approach, the authors proposed a novel adaptive personalized technique based on context to adapting search outputs consistent with each user’s requirement in different situations for relevant information with slight user effort. Personalized Web search is a successful way of offering precise results to different users when they present the Following to the process of the context-based adaptive personalized search investigation, three important technologies to execute this method are provided, which are semantic indexing for Web resources, modeling and obtaining user context and semantic resemblance matching among Web resources and user context. Experimental observations prove that the adaptive personalized search system is executed by most of users and the approach to personalize Web search is effective.

Kyung-Joong Kim et al., [7] developed a personalized Web search engine using fuzzy concept network with link structure. Most of the famous search engines make use of link structure to discover precision result. Typically, a link-based search engine provides superior-quality outputs than a text-based search engine. On the other hand, they have complexity in providing the result that satisfies the specific user's preference. Personalization is necessary to maintain a more suitable result. Among many approaches, the fuzzy concept network according to a user profile can characterize a user's subjective interest appropriately. The paper proposes another search engine that utilizes the fuzzy concept network to personalize the outputs from a link-based search technique. Depending on a user profile, the fuzzy concept network rearranges five outputs of the link-based search engine, and the system presents a personalized superior quality result. Experimental observations with the three subjects show that the system proposed searches not only appropriate but also personalized Web pages on a user's preference.

Chen Ding et al., [8] suggested personalized Web search with self-organizing map. The commonly used Web search engines provide the similar answer set for different users, even though the users possibly will have different preferences. Personalized Web search performs the search for all users according to their preference. With the intention of minimizing the consumption of time on browsing irrelevant documents, this paper suggests an intelligent Personal Agent for Web Search (PAWS). The PAWS cleverly utilizes the self-organizing map (SOM) as the user’s profile and therefore, is capable of providing high quality answer set to the user.

C. Biancalana et al., [9] proposed a new way for personalized Web search using social tagging in query expansion. Social networks and collaborative tagging systems are quickly attaining more recognition as most important means for categorization and sharing data: users tag their bookmarks so as to make things easier for information distribution and later visit. Social Bookmarking approaches are helpful in two essential respects: first, they can permit an individual to remember the browsed URLs, and second, tags can be made by the group of people to direct users towards important content. This paper concentrates on the latter use. The authors present a novel method for personalized web search with the use of query expansion [2, 3]. This paper additionally extends the group of recognized co-occurrence matrix approach models by using a new method of exploring social tagging services. The proposed approach illustrates its
power especially in the case of disambiguation of word contexts. This paper shows steps to plan and execute such a system in practice and performed numerous experiments on a real web-dataset. This is the first study focused on the use of social bookmarking and tagging approaches for personalization of web search and its performance in a real-world application.

Personalized Web search with location preferences is recommended by K.W.-T. Leung et al., [10]. Since the amount of Web information develops at very fast pace, search engines must be capable of retrieving information based on the user's preference. In this paper, the authors recommended a novel web search personalization technique that recognizes the user's interests and preferences with the help of concepts by mining search outputs and their clickthroughs. Because of the significant role played by the location information in mobile search, the author planned to divide concepts into content concepts and location concepts, and categorize them into ontologies to generate an ontology-based, multi-facet (OMF) profile to exactly recognize the user's content and location interests and hence enhance the search accuracy. Furthermore, identifying the information that different users and queries may have different importance on content and location information, this paper introduces the idea of content and location entropies to determine the amount of content and location information linked with a query, and click content and location entropies to calculate how much the user is concerned in the content and location information in the results. As a result, the authors proposed to define personalization efficiency based on the entropies and utilize it to equal the weights among the content and location facets. Finally, based on the resultant ontologies and personalization efficiency, the author trained an SVM to adapt a personalized ranking function for re-ranking of future search. Many experiments were carried out to evaluate the precision produced by this OMF profiles and that of baseline method. Experimental outcome shows that OMF enhances the precision considerably compared to the baseline.

J. Lai et al., [11] compared personalized Web search results with user profile. The huge information available on the Internet makes it difficult for users to acquire necessary information from the Web searched results in a more personalized approach. For the same key word, most of the search engines provide the similar result to each user without considering user preference. For many users, it is no longer adequate to obtain non-customized outputs. It is vital to evaluate users' search and browsing activities [17] based on searching keywords inputted by users, the clicking rate of each link in the output and the time they used on each site. To this end, the authors have proposed a technique to obtain user searching profiles. This paper also proposed a method to obtain document profiles, according to similarity score of documents. In this paper, the authors discussed how to utilize this model to integrate the user searching profiles and the document profile, with the intention of presenting personalized search results to the users.

B. Smyth [12] proposed a community-based approach to personalizing Web search. Researchers can influence the underlying knowledge produced within search communities by gathering users' search behaviors - the queries they enter and results they choose - at the community level. They can make use of this data to construct a relevance model that provides the promotion of community-relevant results throughout standard Web search. This paper focuses on the collaborative Web search technique that encourages the suggestion that community search behaviors can offer valuable form of search knowledge and sharing of this knowledge makes adapting conventional search-engine outputs possible.

O. Shafiq et al., [13] put forth community aware personalized Web search. Searching for the accurate information over the Web is not easy. In the era of high speed internet, elevated capacity networks, and interactive Web applications, it has turned out to be even easier for the users to publish data online. A huge amount of data is published over the Internet; every data is in the form of web pages, news, blogs and other material, etc. In the same way, for search engines like Google and Yahoo, it becomes slightly hard to discover the accurate information, i.e., as per user's preferences; search results for same query would be different in priority for different users. In this paper, the authors proposed a technique to prioritize search results of search engines like Google, according to the individual interests and context of users. With the purpose of finding out personal interest and context, this paper follows a distinctive method of (a) finding out behaviors of a user of his/her social-network, (b) finding out what information does the social networks (i.e., friends and community) present to the user. Based on this information, the authors have formulated a methodology that considers the information about social networks and prioritize search outputs from Web search engine.

Han-joon Kim et al., [14] suggested building a concept network-based user profile for personalized Web search. This paper presents an innovative approach of building the user profile of concept network for personalized search. The user profile is characterized as a concept network, where each concept is approximately denoted with the formal concept analysis (FCA) theory. Consider a concept, called 'session interest concept', include a user's query intention throughout a query session and it can reproduce the user's preference. Every time a user enters his/her query, a session interest concept is
created. Then, these innovative concepts are combined into the existing concept network (i.e., a user profile) wherein recent user preferences are gathered. Based on FCA, a session interest concept is characterized as a pair of extent and intent where the extent integrates a set of documents preferred by the user among the search outputs and the intent comprises of a set of keyword features obtained from the chosen documents. And, with the intention of making a concept network develop, it is necessary to estimate the resemblance among a new concept and presented concepts, and to this end, the author makes use of a reference concept hierarchy called Open Directory Project. The user profile of concept network is ultimately utilized to develop a user's initial query. The experimental result proves that this approach increases the accuracy of search results based on the personal preference.

Yan Chen et al., [15] recommended a personalized context-dependent Web search agent using semantic trees. In Web searching applications, contexts and users' preferences are two significant features for Internet searches in some way that outputs would be much more appropriate to users' requests than with existing search engines. Researchers had planned a concept-based search agent which utilizes conceptual fuzzy set (CFS) for matching contexts-dependent keywords and concepts. In the CFS model, a word accurate meaning may be determined by other words in contexts. Owing to the fact that various combinations of words may become visible in queries and documents, it may be complicated to identify the relations between concepts in all possible combinations. To avoid this problem, the authors proposed a semantic tree (ST) model to identify the relations between concepts. Concepts are symbolized as nodes in the ST, and relations connecting these concepts are represented by the distances between nodes. Furthermore, this paper makes use of the users' preferences for personalizing search results. Finally, the fuzzy logic will be utilized for finding which factor, semantic relations or users' preferences will control results.

Wen-Chih Peng et al., [16] proposed ranking Web search results from personalized perspective. In this paper, the authors develop the approach of data mining to mine common access patterns from user browsing activities. By considering common access patterns, developed an approach to automatically obtain user interests. Additionally, according to the user interests mined and feedbacks of users, a new approach is proposed with the plan of dynamically altering the ranking scores of Web pages. In particular, algorithm PPR standing for Personalized Page Rank, is segmented into four stages. The first stage allocs the initial weights according to user interests. In the second stage, the virtual links and hubs are generated based on user interests. By examining user click streams; this proposed algorithm will incrementally reproduce user favors for the personalized ranking in the third stage. To enhance the accuracy of ranking, collaborative filtering is considered when the new query is entered. By carrying out simulation experiments, it is shown that algorithm PPR is not only very efficient but also very adaptive in offering personalized ranking to users.

M. Kutub et al., [17] examined the user Web search behaviour. Since the beginning of the web searching technology, people have been searching for almost everything and anything on the internet. The ever-increasing habits of users on these search engines and the dynamic nature of the World Wide Web has diminished the accurateness of the search results and raised the search time of an individual. In recent days, more than ever before, there is a necessity for search engines to be appropriate and accurate to the user's requirements and to be capable of making decisions about what the user desires to search, and should be competent to suggest him similar or related topics of his interest. This mounting necessity of the search engine to develop into a decision engine brought to fore different creative technological approaches like Tag clouds and AutoComplete. For enhanced and more appropriate search experience, it is essential that the study of existing search behavior of users and its equivalent response by the search engine. This approach considers the nature of searches made and how they change from time to time. In this paper, the authors examined and take data from different angles and then present suggestions and conclusions for improved, more personalized and relevant search.

B. Arzanian et al., [18] proposed a multi-agent based personalized meta-search engine using automatic fuzzy concept networks. Since the dynamic content of the web develops rapidly, the common purpose web search engines are becoming poor. Even though the meta-search engines can assist by raising the search coverage of the web, the vast number of unrelated results returned by a meta-search engine is still causing problems for the users. The personalization of meta-search engines avoids this problem by filtering results according to individual user's interests. In this paper, a multi-agent structural design is developed for personalizing meta-search engine by means of the fuzzy concept networks. The most important objective of this paper is to use automatic fuzzy concept networks to personalize outputs of a meta-search engine presented with a multi-agent architecture for searching and fast retrieving. Experimental outputs indicate that the personalized meta-search results of the system are more appropriate than the combined results of the search engines.

Dik Lun Lee et al., [19] put forth personalized concept-based clustering of search engine queries. The remarkable development of information on the Web has forced new challenges for the construction of effective
search engines. The most important problem of Web search is that search queries are typically short and ambiguous, and thus are inadequate for identifying the precise user needs. To alleviate this difficulty, a few search engines recommend terms that are semantically connected to the specified queries so that users can select from the suggestions the ones that return their information needs. In this paper, the author introduced an efficient technique that recognizes the user's conceptual preferences with the intention of providing personalized query suggestions. This objective can be realized with two new strategies. At first, develop online approaches that extract concepts from the Web-snippets of the search outputs returned from a query and utilize the concepts to recognize related queries for that query. Then, propose a novel two-phase personalized agglomerative clustering approach that is capable of creating personalized query clusters. No earlier work has focused personalization for query suggestions, according to author’s knowledge. To estimate the efficiency of this technique, a Google middleware was formulated for collecting clickthrough data to perform experimental evaluation. Experimental results show that this technique has enhanced precision than the existing query clustering approaches.

F. Akhlaghian et al., [20] proposed a personalized search engine using ontology-based fuzzy concept networks. At the moment, personalization of search engines as the only web search tool plays significant role in raising the speed of access to web information. Since the users may have various backgrounds and anticipations for a specified query, personalization of search engines outputs based on user's profile can assist to better match the overall interests of an individual user. In this paper the authors personalize the search engine outputs with the help of automatic fuzzy concept networks. The main objective is to make use of the concepts of ontology to improve the common fuzzy concept networks built according to user's profile. Experimental output shows enhancement in personalized search engine outputs using enriched fuzzy concept networks contrast to common fuzzy concept networks.

III. PROBLEMS AND DIRECTIONS

Many Personalized Web search approaches have been discussed in the literature survey. The following directions will help the researchers and guide them to develop the best Personalized Web search approach.

a) Information about the Users Community in Social Networks

A framework to rank Web search outcomes according to the information provided by the user in his/her community from the social networks. This approach can assist in providing more relevant information for a particular user by reorganizing the search results from Web search engines. Hence it enables users to acquire the right information in accordance with their interest uncomplicatedly.

b) Input Query with the Category of Search

The input query is associated with a small set of categories. For example, the input query “windows” can be entered with the category “software” or “computer”, and then the user’s objective becomes understandable. Also the construction of specific search engines for specific purpose will satisfy the users.

IV. CONCLUSION

The remarkable development of information on the Web has forced new challenges for the construction of effective search engines. The only input to the search engine is the keyword and it searches the whole WWW to provide the relevant information to the users. However, most of users are likely to use only a few keywords to convey their information requirements, and thus the search queries usually do not correspond to what the users want specifically. In addition, with the huge development of the information presented on the Web, it is very complicated for Web search engines to satisfy the user information requirement only with a short ambiguous query. To overcome such a basic difficulty of information retrieval, personalized search, which is to provide the customized search results to each user, is a very promising solution. Fundamentally, in studying how a search can be personalized, the most significant thing is to accurately identify users’ information. This survey will guide the researchers to develop a promising technique for Personalized Web search technique.

REFERENCES RÉFÉRENCES REFERENCIAS


This page is intentionally left blank