ISSN:0975-9646

Implementation of Personalized User Model Based on Ontology: A Review Study

Shamkala Waghmare[1], Prof. R. K. Krishna[2]

Final Year M.Tech, Department of Computer Science,
RCERT, Chandrapur, Gondwana University (M.S.) India[1]
Associate Professor, Department of Electronics Engineering,
RCERT, Chandrapur, Gondwana University (M.S.) India[2]

***Abstract*- As the number of Internet users and the number of accessible Web pages grows, it is becoming increasingly difficult for web users to find specific documents that are relevant to their particular needs. Therefore, a new web search personalization approach has been proposed that captures the user's interests and preferences in the form of concepts by mining search results and their clickthroughs. Since the user's location information play a very important role in web search, it is needed to separate the concepts into content concepts and location concepts, and then organizes them into ontologies to create an *ontology-based* profile to precisely capture the user's content and location interests and hence improve the search accuracy. In personalized search systems, the search results are ranked according to user's interest or the searchable documents are arranged according to user-defined concepts for obtaining the desired information faster. This paper will make an extensive review of the personalization consideration; it proposes a personalized user model to provide more user-oriented information considering context information such as a personal profile with preferences and location. The system architecture is designed to support an effective execution usage on Web services and client applications.**

Keywords— **Web search, Clickthrough data, concept, location search, ontology, personalization, user profile.**

I INTRODUCTION

In the today's era of global communication and information sharing, internet has emerged as a fundamental information and communication medium that has generated extensive enthusiasm. As the number of Internet users and the number of accessible Web pages grows, it is becoming increasingly difficult for users to find documents those are relevant to their particular needs. Users must either browse through a large hierarchy of concepts to find the information for which they are looking for or submit a query to a publicly available search engine and search through hundreds of results, most of them are irrelevant. Search engines are a very popular way to locate information, most existing Web search engines return a list of search results based on a user's query but ignore the user's specific interests and/or search context. Therefore, the identical query from different users or in different contexts will generate the same set of results displayed in the same way for all users, a so called one-size-fits-all [Lawrence 2000] approach.

In the modern Web, as the amount of information available causes information overloading, the demand for personalized approaches for information access increases. Personalized systems address the overload problem by building, managing, and representing information customized for individual users. This customization may take the form of filtering out irrelevant information and/or identifying additional information of likely interest for the user. Research into personalization is ongoing in the fields of information retrieval, artificial intelligence, and data mining, among others.

Personalized search is one way to resolve the problem. By capturing the users' interests in user profiles, a personalized search system is able to adapt the search results obtained from general search engines such as Google to the users' preferences through personalized reranking of the search results. In the personalization process, user profiles play a key role in reranking search results. Several personalization techniques have been proposed to model users' content preferences via analysis of users' clicking and browsing behaviors [17], [18].

Ontological concepts and methods in the computer field have been used for knowledge representation, knowledge sharing and reuse. Ontology concepts are utilized to represent user profiles. Ontologies provide a common understanding of topics for communication between system and users, and enable Web-based knowledge processing, sharing, and reuse between applications. Ontologies enable intelligent agents to gather Web information for users in knowledge-based Web gathering.

User profile is a representation of a user in an information system. User profiles were used in web information gathering to interpret the semantic meanings of queries and capture user information needs. User profiles reflect as the interesting topics of a user's information needs. When users read through a document, they can easily determine whether or not it is of their interest or relevance to them. Hence, user profiles are used in web information gathering to capture user information needs from user submitted queries, in order to gather personalized web information for users.

In personalized search systems, the search results are ranked according to user's interest or the searchable documents are arranged according to user-defined concepts for obtaining the desired information faster. Personalized information system provides more user-oriented information. A practical approach to capturing a user's interests for personalization is to analyze the user's click

through data, developed a search personalization method based on users' concept preferences and showed that it is more effective than methods that are based on page preferences. Observing the need for different types of concepts, present a personalized user model which represents different types of concepts in different ontologies. In particular, recognizing the importance of location information in user search, separate concepts into location concepts and content concepts. For example, a user who is planning to visit India may issue the query "hotel," and click on the search results about hotels in India. From the click through of the query "hotel," user can learn the user's content preference (e.g., "room rate" and "facilities") and location preferences ("India"). The proposed framework is capable of combining a user's GPS locations and location preferences into the personalization process.

In this paper, the importance of location information in personalized search has been recognized and hence proposed to incorporate the user's *location preferences* in addition to content preferences in user profiles. User profiling strategy thus incorporates to capture both of the users' content and location preferences for building a personalized user model. The propose a realistic design for Personalized User Model by adopting the search approach which relies on one of the commercial search engines, such as Google, Yahoo, or Bing, to perform an actual search. Personalized user model consists of two major activities: Reranking and Profile Updating.

- **Reranking:** The user submits the user's query to the server, the server displaying the returned results, and collecting his/her clickthroughs in order to derive his/her personal preferences. Result sets are displayed as URL Links, summary and titles that are relevant to the query. The results are re-ranked and combined with previous user's activities, identify relevant documents and put them on top of the result list. The clicked URLs are saved into the database for future references.
- **Profile Update:** After the search results are obtained from the search engines, the content and location concepts and their relationships are mined online from the search results and stored, respectively, as content ontology and location ontology. When the user clicks URLs in a search result, the clicked result together with its associated content and location concepts are stored in the user's clickthrough data automatically and saved into database server.

The rest of the paper is organized as follows:- In section II literature survey from previous papers is presented. In Section III, the strategy to personalize web search is introduced. Section IV gives the conclusion drawn from the review paper. Section V gives the problems and future directions that can help to explore the related issues.

II RELATED WORK

Most commercial search engines return roughly the same results to all users. However, different users may have different information needs even for the same query. For example, a user who is looking for a laptop may issue a query "apple" to find products from Apple Computer, while a housewife may use the same query "apple" to find apple fruits. The objective of personalized search is to disambiguate the queries according to the users' interests and to return relevant results to the users. Clickthrough data is important for tracking user actions on a search engine.

D. E. Rose et al. [13] and U. Lee, Z. Liu et al. [14], studied users' click-through behavior, to understand the user's intentions. Clickthrough data plays an important role for tracking user actions on a search engine. They find that over 60% of queries were informational, and nearly 40% seemed to give unrelated information as per user's request. The classification by U. Lee, Z. Liu et al. [14], uses clickthrough data to identify the information need reflected by a query.

X. Liu, N. Liu and Du[3] based on ontology, user model building and web service discovery technology, and proposes multi-level personalized service of basic data resources integration platform based on ontology, and then designs a user model based on ontology, personalized – service discovery framework and two types of personalization, including personalized service customized and personalized services recommended.

H. Kumar, P. Park and H. Kim [7] use folksonomies for building user preference list (UPL) based on user's search history, which can be exploited by intelligent systems for query recommendation, personalized search, and web search result ranking by using agglomerative clustering by employing Google Similarity Distance.

Hwang, Shin, Kim, and Lee [8] design a personalized retrieval system considering context information such as location, traffic condition, time, weather, user profile, and others and implement a simple prototype with user's location and profile based on Web services and client applications; also support an effective execution usage on Web services and client applications, and implement a map viewer using a shape type of map format files with Points of Interest information.

Leung, Lee, and W. Lee [9] design PMSE to extract and learn a user's content and location preferences based on the user's clickthrough. To adapt to the user mobility, we incorporated the user's GPS locations in the personalization process. GPS locations help to improve retrieval effectiveness, especially for location queries. Due to the importance of location information in search, PMSE classifies these concepts into content concepts and location concepts. In addition, users' locations (positioned by GPS) are used to supplement the location concepts in PMSE. The user preferences are organized in an ontology-based, multifacet user profile, which are used to adapt a personalized ranking function for rank adaptation of future search results. To characterize the diversity of the concepts associated with a query and their relevance to the user's need, four entropies are introduced to balance the weights between the content and location facets. Based on the client-server model, the client collects and stores locally the clickthrough data to protect privacy, whereas heavy tasks such as concept extraction, training, and reranking are performed at the PMSE server.

Many personalized web search systems [17], [18], [19], [20] are based on analyzing users' clickthroughs. T.Joachims [19] presented an approach to learn retrieval functions by analyzing which links the users click on in the

presented ranking. This led to a problem of learning with preference should be ranked higher than document. It studied the problem of learning a ranking function over a finite domain in terms of empirical risk minimization. A Support Vector Machine (SVM) algorithm was given that led to a convex program. It can be extended to non-linear ranking functions. Experiments showed that the method can successfully learn a highly effective retrieval function for a meta-search engine.

Later, Agichitein et al. [17] proposed a method to learn users' clicking and browsing behaviors from the clickthrough data using a scalable implementation of neural networks called RankNet. compared two alternatives of incorporating implicit feedback into the search process, namely reranking with implicit feedback and incorporating implicit feedback features directly into the trained ranking function. Implicit user feedback can further improve web search performance, when incorporated directly with popular
content- and link-based features.

More recently, W. Ng, L. Deng, and D. L. Lee et al. [18] extended Joachim's method by combining a spying technique together with a novel voting procedure to determine user preferences. They present a new approach to mining a user's preferences on the search results from clickthrough data and using the discovered preferences to adapt the search engine's ranking function for improving search quality. They develop a new preference mining technique called SpyNB, which is based on the practical assumption that the search results clicked on by the user reflect the user's preferences, but it does not draw any conclusions about the results that the user did not click on. As such, SpyNB is still valid even if the user does not follow any order in reading the search results or does not click on all relevant results. SpyNB discovers many more accurate preferences than existing algorithms do. The interactive online experiments further confirm that SpyNB and personalization approach are effective in practice and show that the efficiency of SpyNB is comparable to existing simple preference mining algorithms.

Kenneth Wai-Ting Leung et al., [20] introduced an effective approach to predict users' conceptual preferences from clickthrough data for personalized query suggestions and suggested separating concepts into content concepts and location concepts, and organizing them into ontologies, to create an *ontology-based, multi-facet (OMF)* profile to capture the user's content and location interests for improving the search accuracy. The differences between their work and existing works are: Earlier works required the users' to manually enter their location preferences explicitly (with latitude-longitude pairs or text form). They suggested method that does not require users to explicitly enter their location interests manually. In it both of user's content and location preferences, are automatically learnt from the user's clickthrough data from the user's profile. The method studies entropies derived from a query's search results and a user's clickthroughs to estimate the query's content and location ambiguities. The approach consists of two major activities:-

- **Reranking**: The search results are obtained from the backend search engines (e.g., Google), when a user submits a search keyword. These search results are combined and reranked according to the user's profile trained from the user's previous search activities.

- **Profile Updating**: After the search results are obtained, the content and location concepts and their relationships are mined online from the search results and stored, as content ontology and location ontology. When the user clicks on a search result, the user's clickthrough data is updated. The content and location ontologies, along with the clickthrough data, are then employed in RSVM(ranked support vector machine) training to obtain a content weight vector and a location weight vector for reranking the search results for the user.

III PERSONALIZED USER MODEL

The proposed personalized user model consists of user profile and personalized search system. A user profile is a visual display of personal data associated with a specific user. It is common term for user models in information retrieval, filtering, and recommended system. A user's profile is a collection of information about the user of the system in which the system collects and maintains in order to improve the quality of information access. This information can be exploited by systems taking into account the persons' characteristics and preferences. User profiles ensure that personal preferences are used whenever search onto the web. A user profile is different from a user account, which use to log on to system or web sites. Each user account has at least one user profile associated with it.

User has to create his profile by sign in the application and the user profile is further updated implicitly while the users browse. The user profile is essentially a reference ontology in which each concept has a weight indicating the perceived user interest in that concept. Profiles are generated by clicking URL from the result set and saving those URL into the database server, no user feedback is necessary. Google search engine has been used to get the search result from the web.

In personalized search systems, user enters a query to search engine such as Google to search information according to the location of the user's interest. Result sets are displayed as URL Links, summary and titles that are relevant to the query. Ranking is employed to learn a personalized ranking function for rank adaptation of the search results according to the user content and location preferences. For a given query, a set of content concepts and a set of location concepts are extracted from the search results as the document features. The results are re-ranked and combined with previous user's activities, identify relevant documents and put them on top of the result list. The clicked URLs are saved into the database for future references. Profile update automatically when user clicks URLs, and saved into database server. The personalized user model is designed to support an effective execution usage on Web services and client applications.

IV. CONCLUSION

In the literature survey many Personalized Web Search approaches have been discussed in various environments. Personalized Web search is to carry out retrieval for each user incorporating his/her own information need. As the competition in search market increases, some search engines have offered the personalized search service. For example, Google's Personalized Search allows users to specify the Web page categories of interest. An auto completion engine has been built which helps the user to automatically complete his search query. LBS are studied that helps to know the geographical information related to the users. User's clickthrough data and page ranking algorithms play an important role in user's concept extraction.

Personalized Web browsing and search hope to provide Web information that matches a user's personal interests and thus provide more effective and efficient information access. A key feature in developing successful personalized Web applications is to build user profiles that accurately represent a user's interests.

REFERENCES

[1] Xiaohui Tao, Yuefeng Li, and Ning Zhong, "A Personalized Ontology Model for Web Information Gathering" IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 4, APRIL 2011, 1041-4347/11 _ 2011 IEEE Published by the IEEE Computer Society.

[2] X. Tao, Y. Li, N. Zhong, and R. Nayak, "Ontology Mining for Personalized Web Information Gathering," Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence, 0-7695-3026-5/07 © 2007 IEEE DOI 10.1109/WI.2007.88.

[3] Xingbang Liu, Niannian Liu, Hui Du, "Research on Personalized Service of Basic Data Resources Integration Platform Based on Ontology" IT in Medicine and Education (ITME), 2011 International Symposium on Volume: 1 978-1- 61284-704-7/11/ ©2011IEEE.

[4] Jike Ge, Zuqin Chen, Jun Peng, Taifu Li, "An Ontology-Based Method for Personalized Recommendation", Proc. 11th IEEE Int. Conf. on Cognitive Informatics & Cognitive Computing 978-1-4673-2795-4/12/ ©2012 IEEE.

[5] Qing Yan, Junli Sun ,Yan Li, Kaimin Cai, "Domain Ontology-based personalized recommendation research", *2010 2nd International Conference on Future Computer and Communication* 978-1-4244-5824-0/ _c 2010 IEEE.

[6] Yang Huai, "Study on Ontology-based Personalized User Modeling Techniques in Intelligent Information Retrieval", Digital Object Identifier, 978-1-61284-486-2/111- ©2011 IEEE.

[7] Harshit Kumar, Pil-Seong Park, Hong-Gee Kim," Using Folksonomy for Building User Preference List", Ninth IEEE International Symposium on Parallel and Distributed Processing with Applications Workshops 978-0-7695-4429-8/11©2011 IEEE DOI 10.1109/ISPAW.2011.61.

[8] Hyun-suk Hwang, Seong-hyun Shin, Ki-uk Kim, Seok-Cheol Lee, Chang-soo Kim," A Context-aware System Architecture using Personal Information based on Ontology", Fifth International Conference on Software Engineering Research, Management and Applications, 0-7695-2867-8/07 © 2007 IEEE DOI 10.1109/SERA.2007.100.

[9] Kenneth Wai-Ting Leung, Dik Lun Lee, and Wang-Chien Lee, "PMSE: A Personalized Mobile Search Engine" IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 4, APRIL 2013.

[10] Kumar H., Park S., and Kang, S., A Personalized URL Re-ranking Methodology Using User's Browsing Behavior, KES-AMSTA 2008, LNAI, pp.212-221.

[11] Kumar, H. and Kang, S., "Another Face of Search Engine: Web Search API's., IEA/AIE 2008, pp. 311-320.

[12] Susan Gauch, Mirco Speretta, Aravind Chandramouli, and Alessandro Micarelli, "User Profiles for Personalized Information Access", the Adaptive Web, LNCS 4321, pp. 54 – 89, 2007 © Springer-Verlag Berlin Heidelberg 2007.

[13] D. E. Rose and D. Levinson. Understanding user goals in Web search. In *Proc. Of the 13th Int. Conf. on the World Wide Web*, pages 13–19, 2004.

[14] U. Lee, Z. Liu, and J. Cho.," *Automatic identification of user goals in Web search.",* In *Proc. of the 14th Int. Conf. on the World Wide Web*, pages 391–400, 2005.

[15] Ma, Z., Pant, G., and Liu Sheng, " Interest-based personalized search", ACM Trans. Inform. Syst. 25, 1, Article 5 (February 2007), 38 pages. DOI = 10.1145/1198296.1198301.

[16] Ounas ASFARI, Bich-liên Doan, Yolaine Bourda Jean-Paul Sansonnet," Personalized Access to Information by Query Reformulation Based on the State of the Current Task and User Profile", 2009 Third International Conference on Advances in Semantic Processing, 978-0-7695-3833-4/09 © 2009 IEEE DOI 10.1109/SEMAPRO.2009.17.

[17] E. Agichtein, E. Brill, and S. Dumais, "Improving web search ranking by incorporating user behavior information", in *Proc. of ACM SIGIR Conference*, 2006.

[18] W. Ng, L. Deng, and D. L. Lee,"Mining user preference using spy voting for search engine personalization", *ACM TOIT*, vol. 7, no. 4, 2007.

[19] T. Joachims, "Optimizing search engines using clickthrough data", In *Proc. of ACM SIGKDD Conference*, 2002.

[20] K. W.-T. Leung, W. Ng, and D. L. Lee, "Personalized concept-based clustering of search engine queries", *IEEE TKDE*, vol. 20, no. 11, 2008.

[21] K. Leung, D. Lee, W. Lee ," Personalized Web Search with Location Preferences", ICDE Conference 2010, 978-1-4244-5446-4/10/@ 2010 IEEE.