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.

Currently search techniques are mainly based on keyword matching. However, this technique has the drawbacks. As web users cannot express their search intention accurately using several keywords. Hence the exactly-matched results do not satisfy the web users. Another problem about current search engines is their ranking schemes. PageRank is the most popular ranking algorithm; however, it is based on the popularity of web documents, not the quality. Therefore, a newborn web document usually can-not get highly-ranked positions due to their freshness and thus little reputation. How to promote the new documents and maximize quality of search results seen by users is becoming a more and more challenging work.

In this paper, semantic analysis method to overcome the shortcomings of the current search techniques is used. In this work, when user searches query using proposed application, then the application processes the search text and the system check, compare and combine the search text entered by user and his preferences already saved in user profile. The search results are re-ranked and combined with previous users clicked links using semantics. The rest of the paper is organized as follows: - In section II literature survey from previous papers is presented. In Section III, discuss about ontologies for web personalization. In section IV, architecture of personalized search system is introduced. In section V, introduce the algorithm used to personalize search system, section VI gives implementation of proposed approach, VII contains the result and analysis and VIII gives the conclusion part.

2. Literature review

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. 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 in [13] and U. Lee, Z. Liu in [14], studied users' click-through behaviour, to understand the user's intentions. Clickthrough data plays an important role for tracking user actions on a search engine. The classification by U. Lee, Z. Liu [14], uses clickthrough data to identify the information need reflected by a query.

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.

Later, Agichitein in [17] proposed a method to learn users' clicking and browsing behaviours 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.

More recently, W. Ng, L. Deng, and D. L. Lee in [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.

Kenneth Wai-Ting Leung in [20] introduced an effective approach to predict users' conceptual preferences from clickthrough data for personalized query suggestions. 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.

3. Ontologies for web personalization

In this approach, the ontology is used for building the ontological user profile in the interest of a specific Web user. In fig.1, the information describes user's data such as name, age, address, interest, and email-id. The dynamic information describes person's data that is changing over the time. The dynamic information mainly includes user's interest and his preferences for searching information from the web. Each time when a user uses a web service, it will update the user ontology.

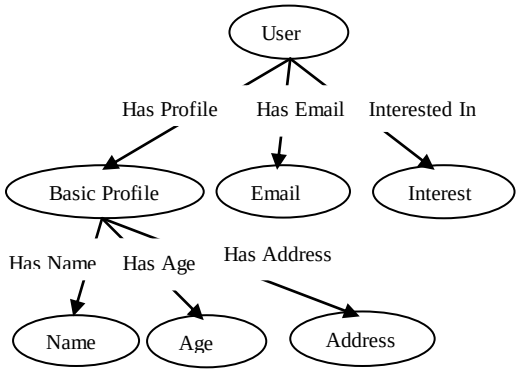

Fig1. The User Ontology

4. Architecture of personalized search model

Personalized search system's outline is as follows: User has to create his profile by signing up into the system. New user has to fill up the sign up form and submit it to the system to create username and password to access and login into the proposed system. Once the user login into the system, a home page will appear where the user has to enter a query into the search text box to get search results from the web. Then, the system offers two options to the user.

1. Search using re-ranking system with and without user's preference
2. Search using without re-ranking system with and without user's preference

When the user select search with re-ranking system with user's preference, then the system check, compare and combine the search text entered by user and his preferences already saved in user profile. Thus, a new set of search text is then sent to Google API to get the desired result set of URLs using semantic techniques. If the user select search without re-ranking system with user's preference option, then the system send the same search text to the Google API without considering the user preferences. When the user clicks on specific URL link which shows the area of interest of the user, the clicked URL is then saved into the application database for updating the user profile. The proposed system architecture is shown in fig.2. Figure 2 displays an approach for search personalization based on ontological user profiles. Semantic algorithm is used to re-rank the search results based on the interest and the semantic evidence in the user profile.

5. Algorithm

A web has become a largest available data set in public domain to the extent that now-a-days; we are using a term "Information Explosion" as the data indexed by the search engines is so huge. This explosion of information has brought some side effects of its own. A keyword based search algorithms used in search engines adds more and more confusion in indentifying requisite data. Hence, the scientists fear about the hiding of expected information in the large set of relevant and irrelevant information. This becomes worst when the keywords used for searching are unambiguous for example, "lotus" where lotus can be a name of a flower, name of hotel, or some individual etc. For carrying out the exact search, word sense disambiguates could be used. This process involves the use of other information present in a semantic analysis system.

A. Steps of semantic search Algorithm:

i. Input search text and get user interest from database
ii. Extract words from text
iii. Select all entries from database(contents and links)
iv. For each word
v. Check if interest is matching or if link is matching
vi. If link or interest matches
Get the output
Else
Exit.

B. Reranking Method

Semantic search method is use to overcome the short comings of the current search techniques. The search based on lexical semantics instead of keyword matching can be better to adapt the thinking pattern of human beings, and thus search results are more relevant to users' search intention. Meanwhile, using semantic factors can conciliate the freshness and make the high-relevant new pages get moderate rank promotion. In this work, we get the output from semantic search method returned by Google API for user query. We also get the number of clicks for the link from the database and arrange the links in descending order of clicks. Thus, we got the rank result set using ranking method.

Steps of Ranking Method:

I. Get output from semantic search
II. Get number of clicks for the link
III. Arrange the links in descending order of clicks

Now, by using the ranking method, the URLs have been ranked. These results of the ranking method are then semantically mapped with the saved contents of the user's profile. Thus, the results are further re-ranked on the basis of matching text between user profile database and the content of URL links. The new ranks are then allocated to the results on the basis of numbers of times text matched. Thus, we get the new ranks of the URL links in descending order of matched text.

Steps of Re-ranking method

I. Get output from ranking method
II. Check semantic scan
III. For each result set link, check no. of times text matches with database and url link
IV. Finally we get result in descending order of text matches

6. The proposed approach

The proposed approach constitutes following three main steps:

1. Creating the Google API for the proposed application.
2. Building the user profile and
3. Re-ranking the search results and update user profile automatically.

A. Creation of Google API

Since, API is not readily available as per the user requirements. Hence, Google API for the project was created. For this purpose, we used Google site to get search result from the web. The main function of API is to get search results of the user query.

B. Building User Profile

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. Wampserver is used to create web application with PHP and MySQL database. Localhost is used as database server. PhpMyAdmin allows managing database of user. Two tables i.e. user table and link table have been created for user profile. User profile has been created by sign up the proposed application and user gets username and password to login the proposed application. User enters the search query to get search results according user interest. Google API is used to search the query using the Google search engine. To fetch HTML documents from the web, it establishes connection to the servers to find all useful links that matches with the query and return them as a result set. By clicking on the desired result link, it opens the web page. The clicked URL links save into link table and the database is automatically updated.

C. Re-ranking of Search Results

When user searches query using re-ranking system, then the application processes the search text and the system check, compare and combine the search text entered by user and his preferences already saved in user profile. Thus, a new set of search text is then sent to Google API to get the desired result set of URLs using semantic techniques. This result sets semantically mapped with saved URL contents of the user's profile and re-ranked result sets are displayed in descending order by num_clicks. Then, another result set is further displayed for the same search text and user preferences using Google API. The time required to search information from web using this application is very much lesser than the time requires in searching using Google search engine. The clicked URL link is then mapped with the existing URL lists of user profile. If the same URL link is not present in the user profile then application adds it to user profile, and if the URL link is present in the user profile then the num_clicks increases by one in user profile.

The user profile has been created by signup to the system. New user fill up the signup form and submits to the system and gets username and password to login to the proposed system. Fig.3. shows login page of the application. Existing user should login to the system.

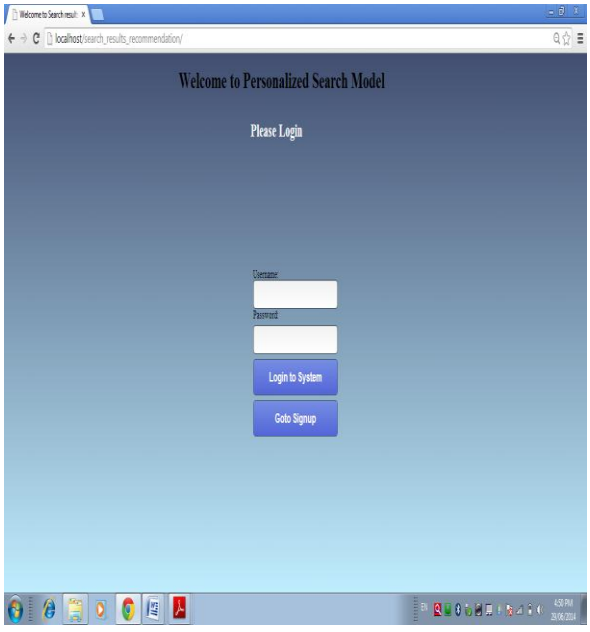

Fig.3. Login page of the application

Fig.4 shows the retrieval result of the user query using re-ranking method with user's preference.

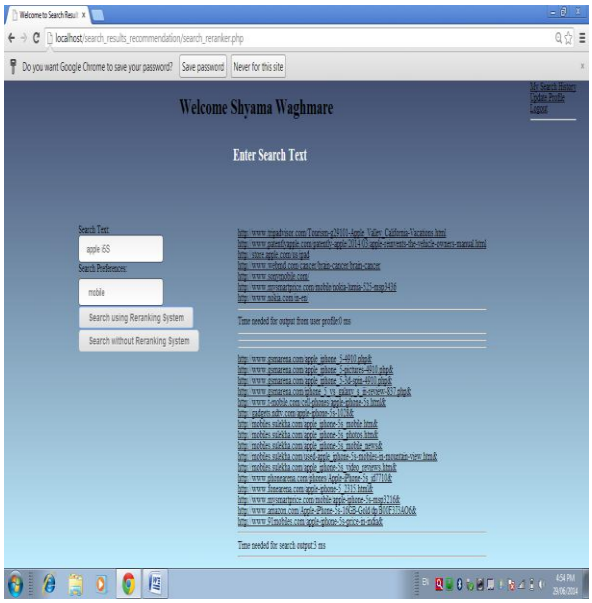

Fig.4.Retrieval result using re-ranking technique.

7. Result and analysis

A. Retrieval Performance measure

In this application, the query together with the user preference is being considered for execution, and then it is executed either by using re-ranking or without re-ranking method and thus, the system retrieves relevant URL links. The performance of this application has been evaluated against the result set generated by executing the queries and user's preference using re-ranking or without re-ranking methods. The measures such as **precision** and **recall** are being used to evaluate the performance of this application.

Precision, Pr is defined as the proportion of total number of relevant URL links and total number of retrieved URL links and is given in (1), where R is the relevant URL links calculated manually and A is the total number of retrieved URL links.

$$\text{Precision, (Pr)} = \frac{R}{A} \times 100 \quad \ldots. (1)$$

Recall, Rc is defined as the proportion of total number of retrieved URL links within the relevant URL links and the total relevant URL links.

$$\text{Recall, (Rc)} = \frac{Ra}{R} \times 100 \ldots. (2)$$

Where Ra is the total number of retrieved URL links from relevant URL links. A similar set of queries was considered for comparative performance measurement for different user. Also, different queries were considered for comparative performance measurement for different user.

B. Experimental Methodology and Results

In order to compare the search result with the personalized search result, we computed the average precision values and average recall values and retrieval time using re-ranking and without re-ranking method for different user using same queries.

A

Precision, Pr		
User	With R/R	Without R/R
U-1	76.4	72.69
U-2	76.89	70.96
U-3	75	67.68
U-4	74.11	66.8
U-5	73.16	70.14

B

Recall, Rc		
User	With R/R	Without R/R
U-1	51.46	49.64
U-2	51.04	50.84
U-3	46.66	45.46
U-4	52.01	50.17
U-5	56.16	50.44

Table 1: A. shows the average precision values with re-ranking and without re-ranking.

B. shows the average recall values with re-ranking and without re-ranking.

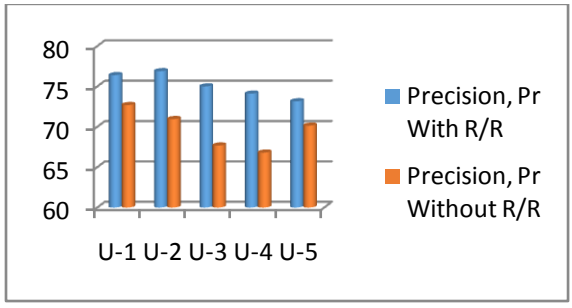

Fig. 5(a): Average Precision values in percentage for the different user and same queries.

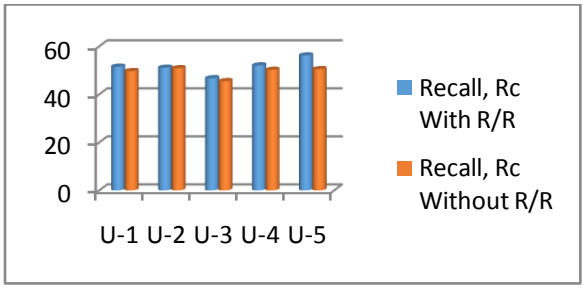

Fig. 5(b): Average Recall values in percentage for the different user and same queries.

To compare the search result with the personalized search result, we computed the average precision values and average recall values using re-ranking and without re-ranking method for different user using different queries.

A

Precision, Pr		
User	With R/R	Without R/R
U-1	76.1	71.68
U-2	77.49	69.75
U-3	75.09	70.18
U-4	78.12	73.82
U-5	74.21	70.97

B

Recall, Rc		
User	With R/R	Without R/R
U-1	62.1	55.58
U-2	62.79	56.09
U-3	60.84	57.52
U-4	60.24	58.57
U-5	65.12	54

Table 2: A. shows the average precision values with re-ranking and without re-ranking.

B. shows the average recall values with re-ranking and without re-ranking.

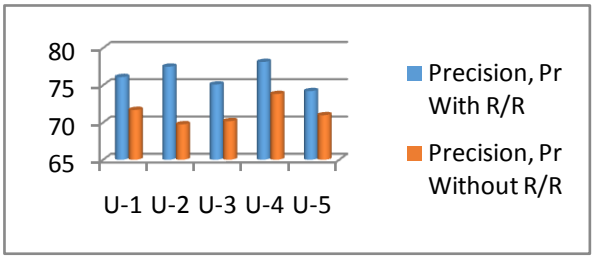

Fig. 6(a) Average Precision values in percentage for the different user and different queries.

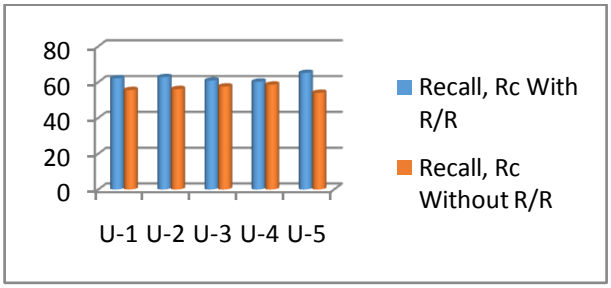

Fig. 6(b) Average Recall values in percentage for the different user and different queries.

8. Conclusion

We have presented a framework for personalized search system that can be used to retrieve the web information for individual user using ontological user profile based on user's interests and preferences. The evaluation results show significant improvement in retrieval time, recall and precision for re-ranking method using user's preference for different user for same queries and different user for different queries. The personalized search system help 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 ontological user profiles that accurately represent a user's interests. The personalized search system is designed to support an effective execution usage on Web services and client applications.

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.

[22] Fang Liu, Clement Yu, Weiyi Meng, "Personalized Web Search by Mapping User Queries to Categories", CIKM'02, November 4--9, 2002, McLean, Virginia, USA ACM 1-58113-492-4/02/0011.

[23] Ma, Z., Pant, G., and Liu Sheng,"Interest-based personalized search. ACM Trans. Inform. Syst. 25, 1, Article 5 (February 2007), DOI = 10.1145/1198296.1198301http://doi.acm.org/10.1145/1198296.1198301.