ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

Ontology Based Personalized Mobile Search Engine With Location Preferences

Snehal D. Jadhav[1], Vaishali P. Suryawanshi[2]

Dept. of Information Technology, MIT College of Engineering, Pune, India[1]

Asst. Professor, Dept. of Information Technology, MIT College of Engineering, Pune, India[2]

Abstract: Proximity search is the most frequent activity people do in order to get information about nearby things. During this search most of the users make use of mobile devices for performing this kind of search. The small screen and default keypad of these mobile devices limit the interactions between user and search server. In such case, the search must not be irrelevant to the user search, we need an efficient way to give query and get response. In this paper, we propose a personalization approach by means of which we will be able to capture user's interests and preferences by maintaining history of their choices. The search keywords are arranged into ontology. The search engine works on client-server model. Heavy tasks such as creation of ontology, maintaining history, performing search are done by the server, client acts as an interface between user and server. We prototype mobile search engine on Google Android Platform. The approach is given an application interface: shopping and navigation. Two approaches are implemented to accomplish the search: Ontology search and Semantic Ontology Search. From the comparison of retrieval times of a normal search engine and an ontological search engine, it has been found that, the retrieval effectiveness of the proposed system is higher than that of a search engine.

Keywords: Ontology search, personalization, mobile search engine, location preference.

I. INTRODUCTION

As the number of products and amusements available in the market increases, it becomes more difficult for a user to search the product and see its particulars, same is the case with navigation. In such case, by using a mobile device, more constraints are added to the search since the size of screen, computational power and memory are the limiting factors. By considering this scenario, it is necessary to provide the user with relevant search results. To do this, the concept of "personalization" was introduced. Several methods are introduced till now to achieve this personalization of search results.

Ontology describes how something exists, and in a search context this refers to instances, classes, attributes and their relations. Ontology can be also viewed as a vocabulary that saves meanings of data unambiguously. Computers do not have a vocabulary like people do hence they are not able to relate various terms to each other. Ontology defines entities and their relations that can be used by queries and assertions for keyword mapping.

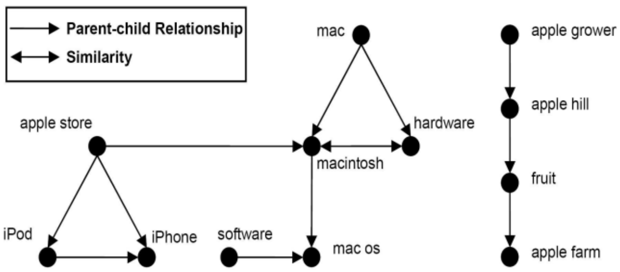

Fig. 1. Ontology for "apple" query

Most of the search engines search for keywords to answer the queries from users. The search engines usually search web pages for the required information. However they filter the pages from searching unnecessary pages by using advanced algorithms. These search engines can answer topic wise queries efficiently and effectively by developing state-of art algorithms. The main focus of these search engines is solving these queries with close to accurate results in small time using researched algorithms. In this paper, we propose a personalization technique that gives personalized search result to user queries along with the recommended results. The concept of ontology is used to provide more relevant search results and provide mobility to the search by providing location preferences.

II. LITERATURE SURVEY

Wilfred Ng, et al, [2] proposed a personalization approach which was based on query clustering. First, they developed online techniques that extract concepts from the web-snippets of the search result returned from a query and use the concepts to identify related queries for that query. Second, they proposed a new two phase personalized agglomerative clustering algorithm that was able to generate personalized query clusters. In this system, the time to start the agglomerative clustering was very crucial one, if started earlier, the desired level of personalized results are not obtained and if started late the clustering would go wrong and unnecessary elements would be clustered.

Wang-Chien Lee, et al [3], proposed a new web search personalization approach that captures the user's interests and preferences in the form of concepts by mining search results and their clickthroughs. Due to the important role location information plays in mobile search, they separated concepts into content concepts and location concepts, and organized them into ontologies to create an

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

ontology-based, multi-facet (OMF) profile to precisely capture the user's content and location interests and hence improve the search accuracy.

Dik Lun Lee, et al, propose a personalized mobile search engine (PMSE) that captures the users' preferences in the form of concepts by mining their clickthrough data [4]. Due to the importance of location information in mobile 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. Privacy preserving parameters in PMSE provide user with effective amount of privacy from the server that may be adversely affecting user profile.

Sinha, et al, [1] presented an algorithm in personalization of web search, called Decision making algorithm in order to classify the content in the user history. The segregated results are located into the corresponding directory. Extensive experiment demonstrates the efficiency and effectiveness of our construction.

Georges Gardarin et al. discussed a SEWISE [16] is an ontology-based Web information system to support Web information description and retrieval. According to domain ontology, SEWISE can map text information from various Web sources into one uniform XML structure and make hidden semantic in text accessible to program. The textual information of interest is automatically extracted by Web Wrappers from various Web sources and then text mining techniques such as categorization and summarization are used to process retrieved text information.

III. PROPOSED SYTEM

The proposed system can be explained as follows:

The proposed system works on client- server architecture, client is the user who is submitting the queries to the search engine. Server consists of a search engine and dataset, search engine is responsible for all the search related part of work and dataset consist of data related to our shopping and navigation domain. The server is also responsible for maintenance of user's search history and creation of ontology.

A. History Creation

The users are facilitate to create their own authentic account to the search server and perform the search. Whenever any user is performing the search activity, his/her searches and preferred results are kept recorded as a separate history file. By using the history file we would be able to build up an ontology file in the future. By the number of hits on a particular file, it is ranked as favorable for that particular search.

B. Ontology Creation

Ontology is nothing but the concept that states how the entities are related with each other. History file created for a particular session of a search, is the base for creation of the ontology file. For the navigational purpose, simple geometrical rules are followed to create the ontology, for example, India would be root and its states would be its child elements, districts would be states' child elements. For the shopping ontology, the categories and particulars would be classes enlisting the details of the products.

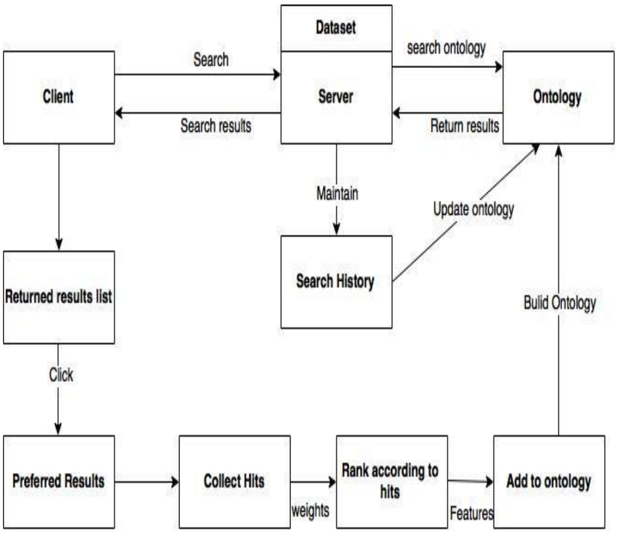

Fig. 2. Proposed System.

Here we propose algorithm that is deployed for the searching of keywords into the ontology.

Every search result, it contains information content that is content feature vector and the location part of result is location feature vector.

For the product ontology, the categories and particulars would be classes enlisting the details of the products. For each product, feature vectors are extracted in order to store them onto the ontology, working in following way:

Content Feature Vector :

Incremented Content feature vector

$$\forall c_i \in s_k, \phi_C(q, d_k)[c_i] = \phi_C(q, d_k)[c_i] + 1.$$

Related concept

$$\forall c_i \in s_k, \phi_C(q, d_k)[c_j] = \phi_C(q, d_k)[c_j] + sim_R(c_i, c_j) + ancestor(c_i, c_j) + descendant(c_i, c_j) + sibling(c_i, c_j).$$

Location Feature Vector :

Incremented Location feature vector

$$\forall l_i \in d_k, \phi_L(q, d_k)[l_i] = \phi_L(q, d_k)[l_i] + 1.$$

Algorithm 1: Ontology Search

Require: searched keyword, ontology

1. for keyword=0 to n do
2. result = match(wordskeyword)
3. count = count_match(result)
4. if count > 1 then
5. for i=0 to count do
6. suggestionswords,i = resulti

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

8. end for
9. else
10. if count == 1 then
11. suggestionswords,i = null
12. end if
13. else
14. if count==0 then
15. // deal un-known keywords, in data set
16. end if
17. end if
18. end for
19. return results

Algorithm 2: Semantic Ontology Search
Require: searched keyword, ontology
1. Convert ontology to RDF
2. //According to classes ontology is created
3. for keyword=0 to n do
4. result = match(wordskeyword)
5. count = count_match(result)
6. if count > 1 then
7. for i=0 to count do
8. parent.i= resulti
9. end for
10. else
11. if count == 1 then
12. parent.i= null
13. end if
14. else
15. if count==0 then
16. // deal un-known keywords, in data set
17. end if
18. end if
19. end for
20. return results

The proposed system retains a list of ontologies from offline resources. The Ontology search requires the saved ontology list and path of the cache where ontologies are saved. The algorithm loops through the ontology list (from line 3 to line 19) to check if it is already cached with the updated copy. If the modified rank is greater than the original rank (i.e. the cached copy is older than the live one), it fetches the updated version from online repository (at line 9) and updates the indexing record (line 10). If the ontology is not already cached, it is fetched from the server (line 16) and generates an index for it (at line 17) and saves a copy in cache.

C. Location Parameter

A Location parameter is provided during the user search to specify in what radius of distance the searched entity must be present. GPS is incorporated in the mobile device in order to get the actual user location. The search results will then be displayed to the client with reference to the location parameter given.

IV. EXPERIMENTAL RESULTS

The system comprises of a server machine, running server, a mobile device running the client application. HTTP protocol is used to communicate between client and server. There is one to n relationship between server and clients. The datasets used are: There is primary custom dataset created for the testing purpose.

Amazon products(1390 instances)
Google products(4860 instances)
Superstore product data set(8932 instances)

The system can be experimented for the time parameter by:

- Keeping no. of records constant and varying no. of queries
- Keeping No. of queries constant and varying no. of records

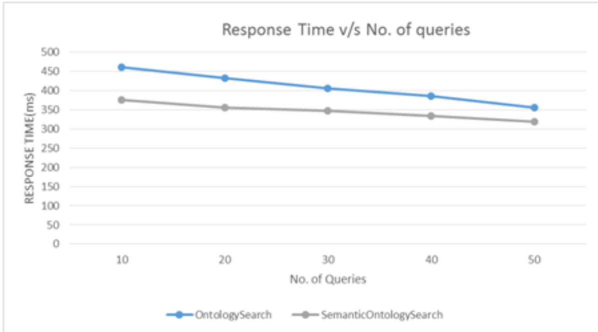

Fig. 3. Graph of response time versus number of queries

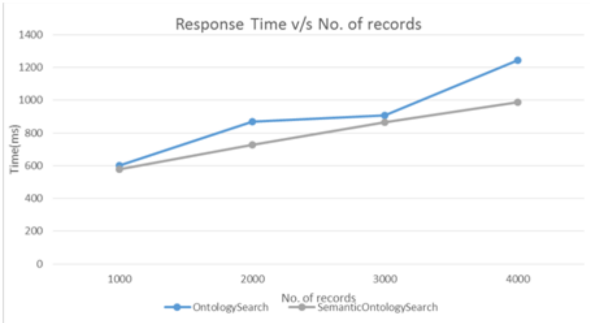

Fig. 4. Graph of response time versus number of records.

Fig. 2 and Fig 3 show the retrieval time required by both the algorithms and it can be seen that in both the cases semantic ontology search takes less time to retrieve search results and is improved by 13.70%.

For calculation of efficiency of the two algorithms, memory and CPU usage are also considered, both before and after firing the query.

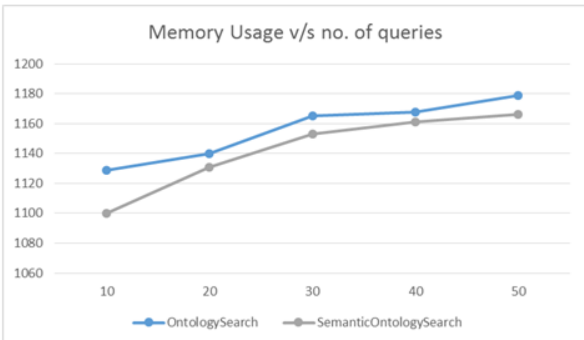

Fig. 5. Memory usage v/s no. of queries.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

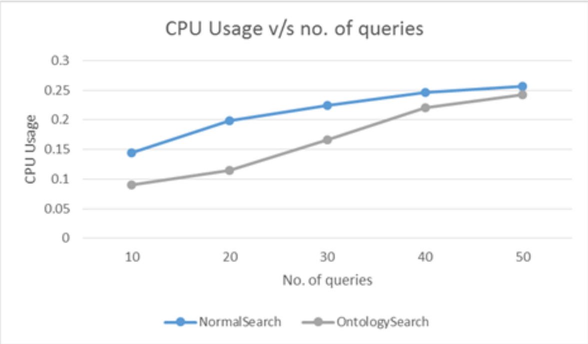

Fig. 6. CPU usage v/s no. of queries

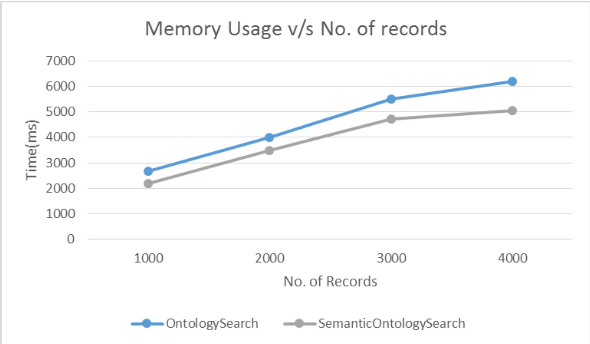

Fig. 7. Memory usage v/s no. of records

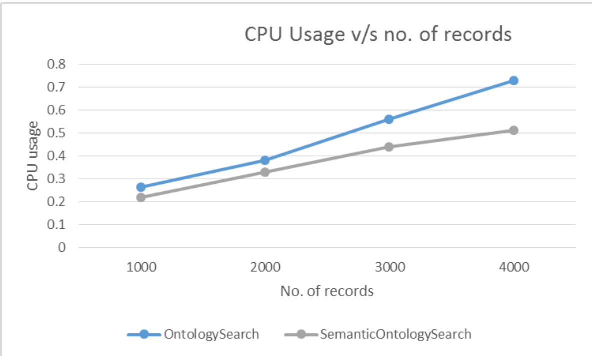

Fig. 8. CPU usage v/s no of records

Fig. 5 and Fig. 7 give the memory usage of both the algorithms, and it can be seen that semantic ontology search takes less memory for execution before and after execution.

Fig. 6 and Fig. 8 show the percentage CPU used by both algorithms while searching, and it can be seen that the semantic ontology search requires less CPU shares than ontology search.

Precision: fraction of retrieved docs that are relevant = P(relevant|retrieved)

Recall: fraction of relevant docs that are retrieved= P(retrieved|relevant)

	Relevant	Nonrelevant
Retrieved	tp	fp
Not Retrieved	fn	tn

Table 1. Table of confusion

Precision P = tp/(tp + fp)

Recall R = tp/(tp + fn)

The scenario for calculation of precision and recall was by judging the relevancy of the retrieved searched results. The value of precision and recall is incremented by the proposed system by 3.46% and 2.27% respectively.

F1-measure is the indicator for both recall and precision with relative importance set by weighing constant β.

$$F1 = \frac{2 * P * R}{P + R}$$

The F1 measure was 2% incremented by using our proposed method to that of ontology search.

V. CONCLUSION

We proposed an ontological personalized mobile search engine to provide relevant results according to user preferences. The personalization of search results can be improved by location preference provided during search. We observed that the location preferences and ranking of the search results aid to better retrieval effectiveness. The results of the proposed system indicate increased values of precision and recall thus improved retrieval. Since ranked results are provided to the users, there are more chances to get desired results in top searches. The future study may include to study the query patterns of users and to study frequent travel patterns to improve personalization.

REFERENCES

[1] Dr. C. R .Rene Robin and R Divya, “Onto-Search: An Ontology Based Personalized Mobile Search Engine”, 2014
[2] K.W.-T. Leung, W. Ng, and D.L. Lee, “Personalized Concept-Based Clustering of Search Engine Queries,” IEEE Trans. Knowledge and Data Eng., vol. 20, no. 11, pp. 1505-1518, Nov. 2008.
[3] K.W.-T. Leung, D.L. Lee, and W.-C. Lee, “Personalized Web Search with Location Preferences,” Proc. IEEE Int’l Conf. Data Mining (ICDE), 2010.
[4] K.W.-T. Leung, D.L. Lee, and W.-C. Lee, “PMSE: A Personalized Mobile Search Engine”, IEEE Trans. Knowledge and Data Eng., vol. 25, no. 4, April 2013
[5] E. Agichtein, E. Brill, and S. Dumais, “Improving Web Search Ranking by Incorporating User Behavior Information,” Proc. 29th Ann. Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2006
[6] E. Agichtein, E. Brill, S. Dumais, and R. Ragno, “Learning User Interaction Models for Predicting Web Search Result Preferences,” Proc. Ann. Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2006.
[7] Y.-Y. Chen, T. Suel, and A. Markowetz, “Efficient Query Processing in Geographic Web Search Engines,” Proc. Int’l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2006.
[8] T. Joachims, “Optimizing Search Engines Using Clickthrough Data,” Proc. ACM SIGKDD Int’l Conf. Knowledge Discovery and Data Mining, 2002
[9] Y. Xu, K. Wang, B. Zhang, and Z. Chen, “Privacy-Enhancing Personalized Web Search,” Proc. Int’l Conf. World Wide Web (WWW), 2007
[10] S. Yokoji, “Kokono Search: A Location Based Search Engine,” Proc. Int’l Conf. World Wide Web (WWW), 2001.
[11] Arooj Fatima, Cristina Luca, “User Experience and Efficiency for Semantic Search Engine”, 978-1-4799-Reformulation, Internation Conference on Computer, Communication, and Control Technology (I4CT 2014), Sept 2-4, 2014.
[12] Azilawati Azizan, Zainab Abu Bakar, “Analysis of Retrieval Result on Ontology-Based Query, Sreekrishna. M, “Ontology Based Classification of User History in Obscured Web Search”, 2nd International Conference on Current Trends in Engineering and Technology, ICCTET’14
[13] Y.-Y. Chen, T. Suel, and A. Markowetz, “Efficient QueryProcessing in Geographic Web Search Engines,” Proc. Int’l ACMSIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2006.
[14] Q. Gan, J. Attenberg, A. Markowetz, and T. Suel, “Analysis of Geographic Queries in a Search Engine Log,” Proc. First Int’l Workshop Location and the Web (LocWeb), 2008.
[15] T. Joachims, “Optimizing Search Engines Using Clickthrough Data,” Proc. ACM SIGKDD Int’l Conf. Knowledge Discovery and Data Mining, 2002.
[16] www.georges.gardarin.free.fr/Articles/Sewise_NLDB2003.pdf.