

# International Journal of Advance Research in Computer Science and Management Studies

Research Article / Paper / Case Study

Available online at: [www.ijarcsms.com](http://www.ijarcsms.com)

## *Personalization of Web Knowledge Using Ontology Model*

**Apoorva A. Andhare<sup>1</sup>**Department of Computer Science & Engg.  
G.H. Raisoni college of Engineering  
Nagpur – India**Nikita V. Mahajan<sup>2</sup>**Asst. Prof.  
Department of Computer Science & Engg.  
G.H. Raisoni college of Engineering  
Nagpur – India

*Abstract: There is an information deluge over the internet and this web information is increasing day by day with a high rate. Gathering of useful information from such a big deluge is a challenging issue. Many web information gathering systems already exists which try to gather useful information from the bulks of data over the web. One of the techniques is with the help of user profiles. The user profiles are created which gives the user background knowledge description. The concept models are implicitly generated from the user background knowledge. The integration of user background knowledge can be better identified and presented by global and local analysis. The ontology is an important concept that can be used to represent the knowledge from the user profiles. This paper reviews few personalized models for web information gathering is given and comparison is made between them and a system based on personalization of web data using ontology is introduced. This system will eventually try to remove the limitations of already existing models for the same purpose.*

*Keywords: user profiles, concept models, ontology, personalization, background knowledge.*

### I. INTRODUCTION

With the rapid development of computer technology and network technology, the internet is becoming an integral part of people's lives. In the past few decades, the web information has exploded rapidly. Thus the gathering of useful information from the web is becoming more difficult and challenging for the developers and researchers. In the attempt of solving this problem many information retrieval (IR) systems have been proposed. And also these systems have made a great difference in solving the problem. However there is still no solution for many issues such as irrelevant search results and ambiguity eventually. Users want that they should get the relevant pages for their searched queries. The user requirements can be satisfied by getting their information needs. For this purpose user profiles are created from the user background knowledge description [1].

A user profile is the description of the concept model possessed by the user [2], [4]. In the terms of web knowledge retrieval, a user profile is the semantic interpretation of the topic based on the user possessed concept model. For representing user profiles, user background knowledge must be gathered. This background knowledge can be taken from global analysis and local analysis. A lot of work has been done in web information gathering using either of the analysis types. But it has been observed that the generated result is not that much effective to get the desired and accurate analysis. The task of web knowledge retrieval can be made more efficient if the two knowledge bases used for analysis are combined and give collaborative result for the same. Global analysis uses existing global knowledge bases for user background knowledge representation and the local analysis uses user's local information [4].

Ontology is one of the concepts which are used to represent the user profiles. Ontology represents the knowledge as a set of concepts within a domain and the relation among those concepts [4]. It is a structural framework for organizing information. Basically, the role of ontology in the knowledge engineering process is to facilitate the construction of the domain model. This

paper presents preliminary work of an effective model that tries to analyze the user search requirements by creating user profiles.

## II. RELATED WORK

The main aspect of web information gathering is personalization. Personalization enables dynamic insertion customization or suggestion of content in any format that is relevant to the individual user based on the user's implicit behavior and preferences and explicitly given details. Jike Ge, Zuqin Chen and Taifu Li et al. [1] gave an approach for personalized recommendation system framework. This system is based on the ontology model to improve the quality of service of recommendation. But the information repository considered for this system is very small and also does not consider heterogeneous data. Another personalized ontology based meta search engine is introduced by Li Qing-Shan, Zou Yan-Xian and Sun Ying-Cheng et al. [2] In this a personalized meta search engine based on ontology with the use of Agent technology is put forward. The agent technology is used for information retrieval through mining user behavior. For the system in this paper there are issues related to the time complexity of the design and thus the system is not much suitable for web knowledge retrieval. Personalization is also used for creating XML based ontologies. S. Vigneshwari and Dr. M. Aramudhan et al. [3] introduced an approach to personalize web using XML based ontologies as XML plays an important role in information retrieval systems. But the developed system is much slow so fast algorithms are needed to make it more efficient. Xiaohui tao, Yuefeng Li, and Ning Zhong et al. [4] Introduced an ontology model for representing user background knowledge for personalized web information gathering. This system tried to combine global and local knowledge base but many important aspects like content based descriptors are assumed to be present but in real scenario it may not be present in web documents. Whereas, Raymond Y. K. Lau, Dawei Song, Yuefeng Li et al. [6] use the fuzzy domain ontology extraction algorithm for adaptive e-learning. Y. Li and N. Zong et al. [7] Also introduced an ontology model for personalization but this model considers only the queries entered by the user other aspects like download history etc. are ignored. Thus there is a lot of work done in the field of personalization for web knowledge retrieval. But there still remains a research gap in learning ontology to describe user background knowledge in information retrieval (IR).

## III. GENERAL APPROACHES

### A. Web Personalization

From the discussion so far it has been observed that the web data have issues related to relevancy and ambiguity. And these issues can be handled through personalization of user profiles. For personalization of web data Web Mining is used. Web mining is a mining of data related to the World Wide Web. This data may be data present in the web pages or data related to web transactions.

Web mining makes the task of Personalization of web pages by categorizing user profiles. Web data can be classified into the following classes [9]:

- Content of the actual web pages.
- The structure includes HTML or XML codes.
- The usage data include information about how the web pages are accessed by users.
- User profile contains demographics and registration information collected about users.

Web mining can be divided into three types: Web Content Mining, Web Structure Mining, and Web Usage Mining. The Fig.1 shows Web Mining Taxonomy.

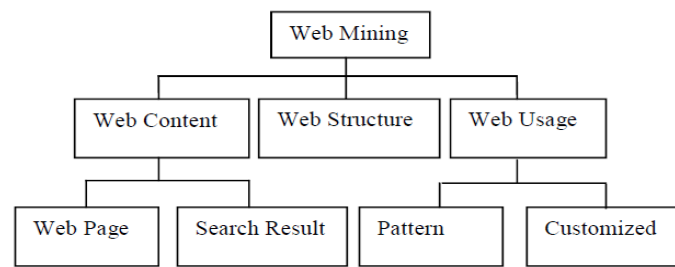


Fig. 1: Web Mining Taxonomy

Web mining can improve search techniques such as content hierarchies, user profiles, synonyms, and analyzing the links between pages.

### B. User Profiles

User profiles were defined as the interesting topics of a user's information need. The user profiles are generally categorized into two diagrams: the data diagram and the information diagram [10]. The former is the discovery of registration data and customer profile portfolios. The latter is the discovery of interesting topics for web user information needs. The data diagram user's profiles acquired by analyzing a database or a set of transactions. The information diagram user profiles acquired by using manual techniques, such as questionnaires and interviews or automatic techniques, such as information retrieval and machine learning. The user profiles are categorized into three groups: interviewing user profiles, semi-interviewing user profiles and non-interviewing user profiles [4].

The user profiles can be created with the help of global and local knowledge bases. The global knowledge base is also called as World Knowledge Base (WKB). The WKB is the knowledge possessed by the people acquired from experience and education. The WKB is the global ontology that formally describes and specifies world knowledge. The local knowledge base is nothing but user background knowledge. The user background knowledge can be retrieved from the user local information such as user's stored documents, browsed web pages, composed and received mails etc. these two knowledge bases can be combined together for more Web Mining Web Content Web Structure Web Usage Web Page Search Result Pattern Customized proper and accurate analysis with the help of text clustering and classification algorithms.

### C. Personalized Ontology Mining

To analyze these user profiles proper knowledge representation model is required which will make the task of analysis easy and visually understandable. Thus for this purpose many knowledge represents models are being introduced. Some of these models are Category model [11], Web model, TREC model [13] and Ontology model. The Table no.1 gives a brief comparison between the models.

Topics	TREC Model	Category Model	Web model	Ontology Model
User Profile	Manual acquiring methods	Non-interviewing	Typical Semi-interviewing	Semi-automatic and automatic interviewing
Topic Coverage	Limited	Depend on category	User profile are satisfactory	Contain less uncertainty and better
Precision	Good	Good	Weak	Moderate

Table No.1 Comparison of Benchmark Models

From Table No.1 it is seen that the Ontology model for knowledge representation is more efficient than others comparatively. Thus the system which is being planned to design will be using the ontology model for knowledge representation. Ontology model is also suitable for user profile representation. Ontology is a data model that represents

knowledge as a concept within a domain and the relationship between these concepts [5], [12]. Ontology provides a vocabulary of terms and relations with which to model the domain.

It is made up of classes and relationships. The classes are nouns. The relationships exist among the things for representation generally is: “part of” and “is a”. Ontology is a formal explicit specification of a shared conceptualization. It is a framework for organizing information and used in System Engineering, Software Engineering, Semantic Web, Artificial Intelligence. The global knowledge and a user’s local instance repository (LIR) are used in the ontology model [7]. Also the Ontology model is important because of following reasons:

- Unambiguous definition of things
- Reuse of the domain knowledge
- Can expand or narrow search terms
- Enable Activity based search
- Separates the domain knowledge from the operational databases
- Validate taxonomy membership
- Distributed and aggregated

The study shows that Ontology model has many advantages over the other three models. These three models are also referred as Benchmark models. The advantages of the ontology model over benchmark models can be listed as follows:

- The ontology model retrieves user background knowledge from user LIRs than the documents read and judged by users.
- Compared to the TREC model, Ontology model have better recall. But in case of precision performance it is relatively weak. Thus the ontology user profiles are not as precise as user profiles of TREC.
- The Ontology profiles topic coverage is broad.
- A large number of uncertainties are eliminated when user background knowledge is discovered. The user profiles acquired by the ontology model have better performance than the web model.
- As compared with data used by the web model from web, the LIRs used by the ontology model are controlled and contained less uncertainty.

#### IV. PROPOSED PLAN

From the discussion till now, it has been observed that that for web knowledge retrieval there is a need of an efficient system that will help to analyze the searching needs of the users and give the relevant web pages as per the requirements of the user. Thus, in this section plan is proposed for a system that will try to remove the problems related to irrelevancy and ambiguity of data present over the web. This system aims to gather the web information accessed by an individual user from the very past to the far future. This mechanism will help to keep track record of the user’s knowledge about particular or various topics the user has searched the web for. Also, it will evaluate the precision of the data resulted from the search till date. In addition to this, it will give an idea about the knowledge base of the user about both global and user’s local knowledge bases. Following fig.2 shows the phases in which the system in the project will be divided:

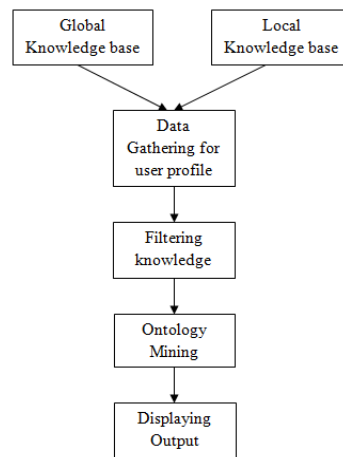


Fig. 2: Phases of the system

The phases in the fig.2 are as follows:

- Creating World knowledge base (WKB) and local knowledge base.
- Creating user profiles.
- Constructing a personalized concept model
- Ontology mining of concept model
- Categorization of gathered information

In the first phase of the system the *WKB* and user's local instance repositories are generated for a particular user. This is done in order to adapt to changes in need of the user search. At times this search may be local or global. "World knowledge is necessary for lexical and referential disambiguation, including establishing reference relations and resolving ellipsis as well as for establishing and maintaining connectivity of the discourse and adherence of the text to the text's producer's goal and plans" [14]. The second phase is about the creation of user profiles from the gathered *WKB* and local instance repositories. The generated user profile is of semi interviewing type as the global knowledge base is generated by taking information from the user. After that from these user profiles personalized concept models are formed for each user. These concept models are then mined and the information is represented using the ontology model. Thus finally with the help ontologies the calculations like precision, specificity and exhaustivity of the search results for a particular user are done.

## V. CONCLUSION

Thus, the paper reviews the various techniques which are used for the purpose of knowledge retrieval. In the case of web information a lot of techniques are used for analyzing the web data. For this purpose user profiles are used. But existing models do not completely handle the issues related to the user requirements. Hence in this paper, a system is introduced that will try to analyze the problems related to web information with the help of personalized ontology model and text categorization/classification algorithm. This analysis system will give the ontological representation of the user's profile as a result. And it will also calculate the parameters related to the precision of filtered data to check the accuracy of the result generated.

## References

1. Jike Ge, Zuqin Chen, Jun Peng, Taifu Li, "An Ontology-Based Method for Personalized Recommendation", 2012 Proc. 11th IEEE Int. Conf. on Cognitive Informatics & Cognitive Computing.
2. LI Qing-Shan, ZOU Yan-xin, SUN Yingcheng, "Ontology based User Personalization Mechanism in Meta Search Engine", 2012 International Conference on Uncertainty Reasoning and Knowledge Engineering.

3. S.Vigneshwari, Dr. M. Aramudhan, "An approach to personalize the web using XML based ontologies" 2012 World Congress on Information and Communication Technologies.
4. Xiaohui Tao, Yuefeng Li, and Ning Zhong, "A Personalized Ontology Model for Web Information Gathering", April 2011 IEEE Trans. On Knowledge And Data Engineering, Vol. 23, No. 4.
5. B. Umamaheswari Pramod Patil, "Personalized Ontology Model – A Survey", 2012 12<sup>th</sup> International Conference on Hybrid Intelligent Systems (HIS).
6. Raymond Y.K. Lau, Yuefeng Li, Terence C.H. Cheung, Dawie Song and Jin-Xing Hao, "Toward a Fuzzy Domain Ontology Extraction Method for Adaptive e-Learning", June 2009 IEEE Trans. On Knowledge And Data Engineering, Vol. 21, No. 6.
7. Y. Li and N. Zhong, "Mining Ontology for Automatically Acquiring Web User Information Needs," IEEE Trans. Knowledge and Data Eng., vol. 18, no. 4, pp. 554-568, Apr. 2009.
8. Xiaohui Tao, Yuefeng Li, Ning Zhong, Richi Nayak, "An Ontology-based Framework for Knowledge Retrieval", 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology.
9. Yuefeng Li, Ning Zhong, Xiaohui Tao, Richi Nayak, "Ontology Mining for Personalized Web Information Gathering", 2007 IEEE/WIC/ACM International Conference on Web Intelligence.
10. S. E. Middleton, N. R. Shadbolt, and D. C. D. Roure. Ontological user profiling in recommender systems. ACM Trans. Inf. Syst., 22(1):54–88, 2007.
11. C. Zhou, D. Frankowski, P. Ludford, S. Shekhar, and L. Terveen, "Discovering Personally Meaningful Places: An Interactive Clustering Approach," ACM Trans. Information Systems, vol. 25, no. 3, article no. 12, July 2007.
12. Ms. Amruta Y. Kokate, Mr. Bhushan D. Nighot, Ms. Chetana N. Sadare, "Personalized Ontology Model for Web Information Gathering", IJARCSSE Volume 3, Issue 2, February 2013.
13. Stephen Robertson , Ian Soboroff, "The TREC 2002 Filtering Track Report ", 2003 Text Retrieval Conference.
14. Shubhangi Shindikar, M.V. Nimbalkar, "A Personalized Ontology for Web Information Gathering by Domain Specific Search", July 2012, International Journal of Scientific & Engg. Research, Vol. 3, issue 7.
15. Yuzhong Qu and Gong Cheng, "Falcons Concept Search: A Practical Search Engine for Web Ontologies", July 2011, IEEE Trans. On Systems, Man, and Cybernetics- PartA. Vol.41, No.4.

#### AUTHOR(S) PROFILE



**Apoorva A. Andhare** received the Bachelor's degree in Information Technology from Guru Nanak Institute of Engineering and Technology, Nagpur in 2011. She has 1 year of teaching experience from Nagpur Polytechnic, Nagpur. Her main area of interest includes Web Mining and Data Mining. She is now pursuing Masters in Technology in Computer science and Engineering from Rasoni College of Engineering, Nagpur.



**Asst. Prof. Nikita V. Mahajan** is currently working as assistant professor in G.H. Rasoni College of Engineering. She has completed Masters in Technology in Computer science and Engineering from Rasoni College of Engineering, Nagpur. She has 3 years of teaching experience from engineering college. She has many national and international publications to her credit.