A Survey on Ontology Model for Gathering Web Related Information

Sayali D. Behere
PG Student
Department of Computer Engineering
RCPIT Shirpur

R. B. Wagh
Asso. Professor
Department of Computer Engineering
RCPIT Shirpur

ABSTRACT
A model for knowledge description and formalization, these various ontologies are widely used to represent various user profiles in personalized web information gathering manner. Represent these user profiles, many of these models have evaluate only their knowledge from either a global knowledge base, and also it called as a user local information. In this paper, a personalized ontology model is introduced for knowledge representation and various reasons over user profiles. ontology model learns ontological user profiles from both a world knowledge base and also an user local instance repositories. The ontology model is also developed by comparing other models and in such way that the results shows that this ontology model is successful.

General Terms
Personalization, Knowledge base

Keywords
Ontology, Local instance repository, User profiles, Web information gathering

1. INTRODUCTION
The amount of web related information available and also has been increased dynamically. To achieve useful information from the web has become a challenging part of issue for various users. Currently added web related information gathering systems attempt to satisfy user requirements by providing user's information needs. For this reason user will create his profiles are created for user background knowledge description.

User profile to represent the concept models given by users when web information gathering process is run. A concept model is also implicitly given by users and is also generated from user background knowledge. Many web ontological user's have been observed it in user behavior. When users read through a document. They can also easily determine whether or not it is of their interest or convenience to them. User's concept model can be simulated, and then representation of user profiles can be created.

Simulation of user concept models are widely distributed in ontology. Knowledge description models are utilized in personalized web information manner to gathering some web related information. These ontology models are called ontological user profiles. To represent user profiles, many user's have research attempted to discover user background knowledge through global or local analysis.

Global analysis also uses existing global knowledge bases for user background knowledge representation. Commonly used various knowledge bases include ontologies and also various online knowledge bases. This global analysis technique also produce effective performance for user background knowledge representation.

Local analysis gives user local information and it also observes user behavior in ontological user profiles. Some ontological groups learned personalized ontologies repeatedly from user's browsing history. User background knowledge have been discovered from this feedback for user profiles. Local analysis techniques also rely on data mining. These classification techniques for knowledge discovery.

The world knowledge bases and a user's local instance repository (LIR) are also used in this ontology model. Local instance repository is a user's personal collection of information related items. The ontology model is developed by comparison against benchmark models through using a large standard data set. The evaluation results show that this ontology model is successful.

In this paper, this ontology model simulates user's concept models by using personalized ontology related information and it attempts to improve web information achieving or gathering performance by using ontological user profiles for gathering web information.

2. LITERATURE REVIEW
2.1 Ontology Learning
Global knowledge bases were used by many existing models to learn ontological user profiles for web information gathering. For example, Gauch et al. [1] and Sieg et al. [2] learned personalized ontology profiles from the Open Directory Project to specify user's requirements and interests in web search. On the basis of the Dewey Decimal Classification, King et al. [3] developed IntelliOnto to improve performance in web information in distributed manner. Wikipedia was also used by Downey et al.[4] to understand user interests in their requirements and queries. These works effectively performed by user background knowledge. Their performance was restricted by the quality of the global knowledge bases.

Learning personalized ontological profiles mined user background knowledge from user local information concept. Li and Zhong [5] used pattern recognition and also an association rule mining techniques to develop knowledge from user local documents for ontology profile construction. Tran et al. [6] gives translated keyword queries to give description logic's conjunctive queries and also used ontology model to represent user background knowledge. Zhong [7] gives a domain ontology learning approach that employed various data mining techniques and natural-language understanding processes. Navigli et al. [8] developed semantic concepts and relations from web related documents.

2.2 User Profiles

User profiles were used in web information gathering technique to interpret the semantic meanings of queries and also capture user information needs.

User profiles can be classified into three groups interviewing, semi-interviewing and non-interviewing user profiles can be perfect user profiles. They are acquired by using manual techniques, such as questioning user qureies, interviewing users and analyzing user classified training sets. The users read each and every document, and also gave a positive or negative judgment like answering to the document against a given topic. Because only these user's perfectly know their interests and also their requirements. Semi-interviewing user profiles are gained by semi automated techniques with restricted user involvement. These techniques usually provide users with a list of categories [3].

Non-interviewing techniques do not involve user's but as certain user who are interested instead. They acquire user profiles by observing user activity and performance and developing user background knowledge [9]. A typical model is OBIWAN, developed by Gauch et al. [1], which acquires user profiles based on user's online browsing history related to web information retrival. The interviewing, semi-interviewing, and non-interviewing user profiles can also be viewed as manual, semi automatic and automatic profiles respectively.

3. PERSONALIZED ONTOLOGY CONSTRUCTION

Personalized ontology models are a conceptualization model that describes and specifies user background knowledge.

3.1 World Knowledge Representation

World knowledge base is important for information gathering. According to the definition ontology model provided by [6] world knowledge is commonsense knowledge possessed by people and acquired through their experience and education. world knowledge is necessary for lexical and referential disambiguation, including establishing relations and resolving ellipsis as well as for establishing and maintaining connectivity of the discourse and adherence of the text to the text producer's goal and plans. In this model, user background knowledge is extracted from a world knowledge base encoded from the Library of Congress Subject Headings (LCSH).

Table 1. Comparison of Different World Taxonomies

	LCSH	LCC	DDC	RC
# of Topics	394,070	4,214	18,462	100,000
Structure	Directed Acyclic Graph	Tree	Tree	Directed Acyclic Graph
Depth	37	7	23	10
Semantic Relations	Broader, Used-for, Related-to	Super- and Sub-class	Super- and Sub-class	Super- and Sub-class

Table 1 shows a comparison of the LCSH with the Library of Congress Classification (LCC) used by Frank and Paynter [10], the Dewey Decimal Classification (DDC) used by Wang and Lee [5] and King et al. [3], and the reference categorization (RC) developed by Gauch et al. [1] using online categorizations. As shown in Table 1, the LCSH covers more topics, has a more specific structure and more semantic relations.

The structure of the world knowledge base related web information used in this research is encoded from the LCSH references. The LCSH system contains three types of references Broader term (BT), Used-for (UF), and Related term (RT) [3].

3.2 Ontology Construction

The subject related to user interest are extracted from the world knowledge base via user interaction. This tool called Ontology Learning Environment (OLE) to assist user's with such interaction.

Fig. 1 shows a screen-shot of the ontology learning environment for the sample topic "Economic espionage". The subjects are listing on the top-left panel of the ontology learning environment are the candidate subjects presented in hierarchical form.

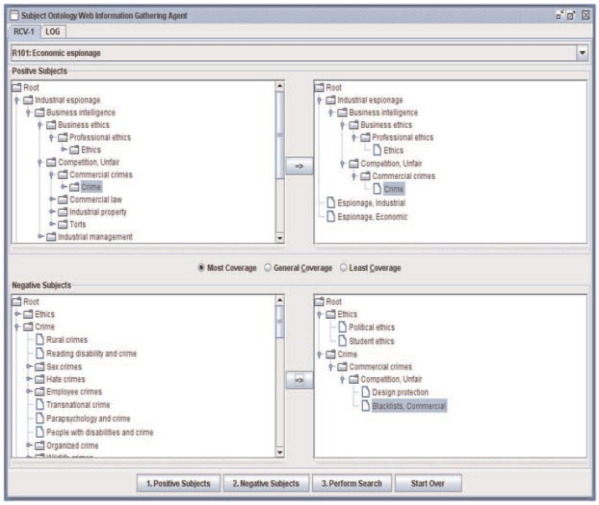

Fig 1: Ontology Learning Environment [4]

The user selects positive subjects for the topic. The user selected positive subjects are presented on the top-right panel in hierarchical form. The candidate negative subjects are the descendants of the user-selected positive subjects. They are shown on the bottom-left panel. From these negative candidates the user selects the negative subjects. These user-selected negative subjects are listed on the bottom-right panel.

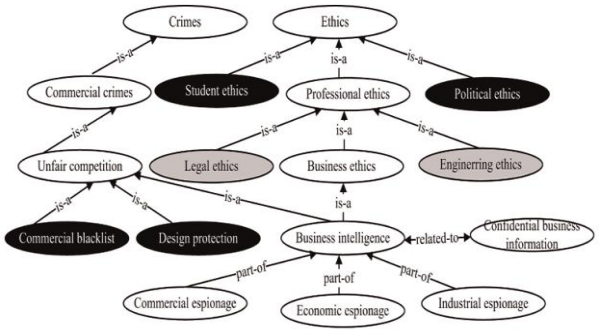

Fig 2: An Ontology Construction for topic Economic Espionage [5]

Fig. 2 shows the ontology constructed for the sample topic "Economic espionage" where the white nodes are positive, the dark nodes are negative and the gray nodes are neutral subjects.

4. ONTOLOGY BASED USER MODELING ARCHITECTURE

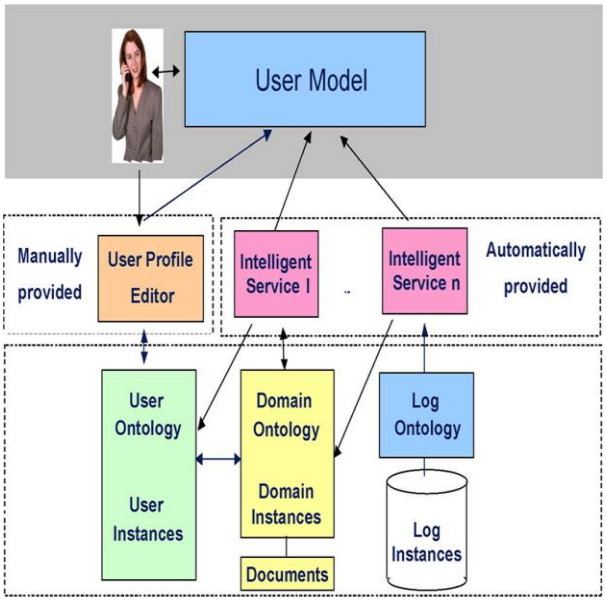

Fig 3: Ontology Based User Modeling Architecture [3]

Ontology Based User Modeling Architecture has been designed as a three-tiered application server to manage information about various user's.

Figure 3 shows the first is the user front end layer on top, second is a middleware layer or it also called as service layer and the third one is an data layer at the bottom[11].

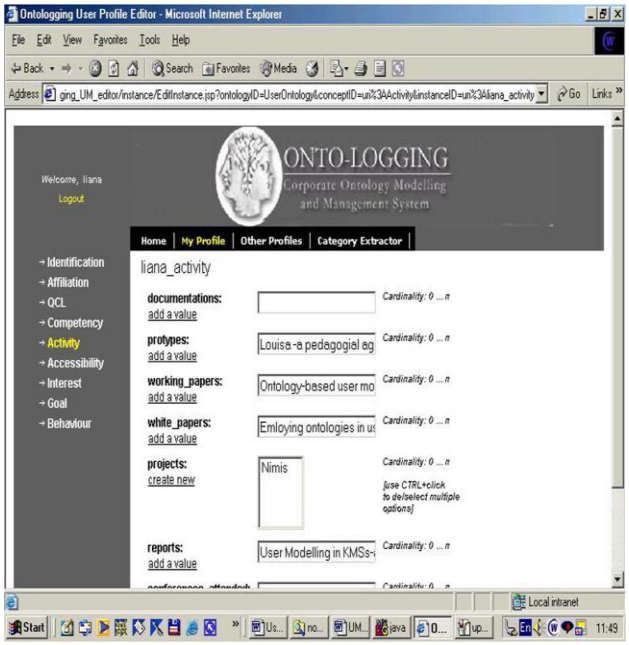

Fig 4: User Profile Editor, in the Edit Mode [6]

Figure 4 shows the user model,but it also enables user's to visualize and update it. The user editor model is an open user model which is expected to create awareness of the identified behavioral model to compare with other user's provide feedback [6].

4.1 Modeling User Behavior

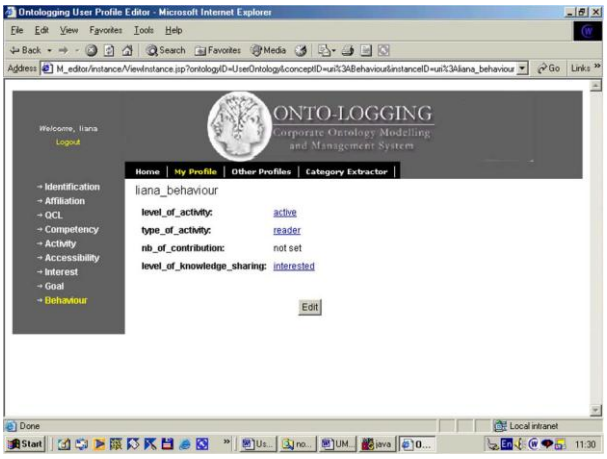

Fig 5: User Profile the Behavior Concept [7]

Figure Shows the behavior concept and it's subconcepts were introduced to model two processes that are important for the effectiveness of the knowledge management system.

The user modeling system classifies their users into three types, readers, writers and lurkers. These types are properties of the type of activity concept. Readers are categories of users accessing the resources of the system, on the other hand writers are accessing and contributing with resources, meta data to the system. And A lurker is defined as somebody who does not contribute and accesses very few knowledge in the system [11].

5. ARCHITECTURE OF THE ONTOLOGY MODEL

This ontology model aims to develop user background knowledge and also learns personalized ontology model to represent user profiles.

Figure 6 shows the architecture of the ontology model. A personalized ontology model is constructed, according to a given subject. Two knowledge resources, the global world knowledge base and the user's local instance repository are utilized by the ontology model. The world knowledge base provides the taxonomic structure for the personalized ontology. The user background knowledge is developed from the user local instance repository.

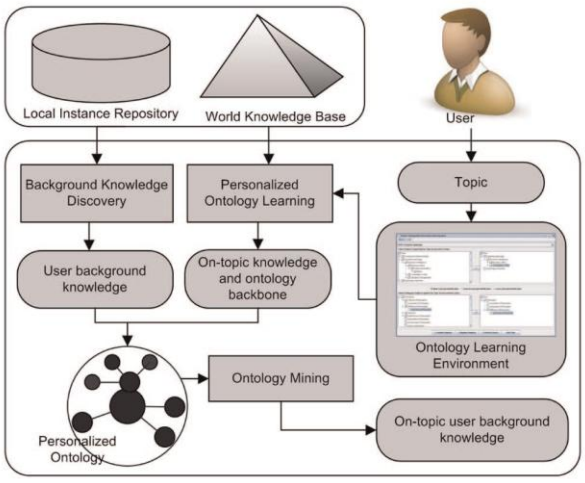

Fig 4: Architecture of Ontology Model [9]

6. METHODOLOGY

The performance of the model was measured by three methods the precision averages at 11 standard recall levels (11SPR), the mean average precision (MAP) and the F1 Measure. These are modern methods based on precision and recall the standard methods for information gathering [1], [3]. Precision is the ability of a system to retrieve only relevant documents. Recall is the ability to retrieve all relevant documents.

An 11SPR value is computed by the interpolated precisions at the specified recall cutoff and then dividing by the number of topics.

The MAP is a discriminating choice and recommended for general-purpose information gathering evaluation [3]. The average precision points for each topic is the mean of the precision obtained after each relevant document is retrieved.

7. CONCLUSION

In this paper, an ontology model is developed for representing user background knowledge for personalized web information gathering system. This ontology model constructs user personalized ontological profiles.

8. REFERENCES

[1] S. Gauch, J. Chaffee and A. Pretschner, "Ontology Based Personalized Search and Browsing", Web Intelligence and Agent Systems, vol. 1, nos. 3/4, pp. 219-234, 2003.

[2] A. Sieg, B. Mobasher and R. Burke, "Web Search Personalization with Ontological User Profiles", Proc. 16th ACM Conf. Information and Knowledge Management (CIKM '07), pp. 525-534, 2007.

[3] J.D. King, Y. Li, X. Tao and R. Nayak, "Mining World Knowledge for Analysis of Search Engine Content", Web Intelligence and Agent Systems, vol. 5, no. 3, pp. 233-253, 2007.

[4] D. Downey, S. Dumais, D. Liebling and E. Horvitz, "Understanding the Relationship between Searcher's Queries and Information Goals", Proc. 17th ACM Conf. Information and Knowledge Management (CIKM '08), pp. 449-458, 2008.

[5] 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. 2006.

[6] T. Tran, P. Cimiano, S. Rudolph, and R. Studer, "Ontology-Based Interpretation of Keywords for Semantic Search", Proc. Sixth Int'l Semantic Web and Second Asian Semantic Web Conf. (ISWC '07/ ASWC '07), pp. 523-536, 2007.

[7] N. Zhong, "Representation and Construction of Ontologies for Web Intelligence", Int'l J. Foundation of Computer Science, vol. 13, no. 4, pp. 555-570, 2002.

[8] R. Navigli, P. Velardi, and A. Gangemi, "Ontology Learning and Its Application to Automated Terminology Translation", IEEE Intelligent Systems, vol. 18, no. 1, pp. 22-31, Jan./Feb. 2003.

[9] J. Trajkova and S. Gauch, "Improving Ontology-Based User Profiles", Proc. Conf. Recherche d'Information Assistee par Ordinateur (RIAO '04), pp. 380-389, 2004.

[10] E. Frank and G.W. Paynter, "Predicting Library of Congress Classifications from Library of Congress Subject Headings", J. Am. Soc. Information Science and Technology, vol. 55, no. 3, pp. 214-227, 2004.

[11] Liana Razmerita, "An Ontology-Based Framework for Modeling User Behavior—A Case Study in Knowledge Management", IEEE Transactions on systems, man and cybernetics part A: systems and humans, vol. 41, no. 4, pp. 772-783, July 2011.

[12] M. S. Ahamad and S. N. Raju, "Web personalization using e_cient ontology relations", International Journal Of Computational Engineering Research (ijceronline.com), vol. 2, pp. 1329-1332, Sept 2012.

[13] S. Natarajan and Mr.J.Sethuraman, "A personalized ontology model for web information gathering using local instance repository", Journal of Theoretical and Applied Information Technology, vol. 38, pp. 145-151, April 2012.