Review Article Impact Factor: 0.621 ISSN: 2319-507X
PR Kharche, IJPRET, 2014; Volume 2 (9): 357-362 IJPRET

INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

A REVIEW ON ONTOLOGICAL USER PROFILE CONSTRUCTION FOR WEB INFORMATION GATHERING

MS. P. R. KHARCHE[1], PROF. D. R. DHOTRE[2]

1. Final Year ME, Department of CSE, SSGMCE, Shegaon (M.S.) India.
2. Associate Professor, Department of CSE, SSGMCE, Shegaon (M.S.) India.

Accepted Date: 27/02/2014 ; Published Date: 01/05/2014

Abstract: For web information gathering, the user profiles are created using ontologies which describe and formalize the knowledge. User profiles are concept models possessed by user. Concept models are generated from user background knowledge. Instead of using either global knowledge base or user local information, both are utilized. The subject of user interest is extracted from world knowledge base via user interaction. The extracted subjects consist of positive subject and negative subjects. Ontology is then constructed for given topic using users' feedback subjects.

Keywords: Ontology, personalization, world knowledge, local instance repository, user profiles, ontology mining.

Corresponding Author: MS. P. R. KHARCHE

Access Online On:

www.ijpret.com

How to Cite This Article:

PR Kharche, IJPRET, 2014; Volume 2 (9): 357-362

PAPER-QR CODE

INTRODUCTION

Gathering useful information from web became difficult as web information amount vastly increased over last decade. This difficulty is tried to be solved over years by capturing user need. For this user profiles are created from user background knowledge. User profiles are concept models possessed by user. These concept models are knowledge possessed by user which is used whenever he derive a judgment over any document. These concept models used to create user profiles. In personalized ontology model these concept models are stimulated to create ontological user profiles by knowledge representation and formalization model.

To create user profiles many researcher have used either global or local analysis for user background knowledge discovery. Global Analysis uses existing global knowledge bases like Word Net, thesaurus etc. It gives good performance but limited by the quality of knowledge base. Local analysis can discover user background knowledge from user browsing history, user local documents, query logs, user feedback etc. Local analysis suffers from ineffectiveness at capturing formal user knowledge. So the ontology model uses both local and global analysis.

1. RELATED WORK

Ontologies are means of knowledge sharing and reuse. Many existing knowledge bases are used by many models to learn ontologies. Gauch [1] et al. and Sieg et al. [2] learned personalized ontologies from the Open Directory Project to specify users‟ preferences and interests in web search. Doweney et al. [3] used Wikipedia which helps in understanding user interests in queries. Pattern reorganization and association rule mining technique to discover knowledge from user local information is used by Li and Zhong [4]. A domain ontology learning approach was proposed by Zhong [5] that uses various data mining and natural language understanding techniques to discover knowledge from user local documents for ontology construction. Doan developed a model called GLUE and used machine learning technique to find similar concepts in different ontologies[6]. In the web information gathering, user profiles were used to understand the semantic meanings of queries and capture user Information needs. Li and Zhong defined user profiles as the interesting topics of a user's information need. User profiles are categorized into three groups: interviewing, semi-interviewing, and non-interviewing.

Interviewing user profiles are acquired by using manual techniques, such as questionnaires, interviewing users [7] and analyzing user classified training sets e.g. TREC model. Semi-interviewing user profiles are acquired by semiautomated techniques which provide users with a list of categories and ask users for interesting or noninteresting categories. Noninterviewing

techniques do not involve users at all. They acquire user profiles by observing user activity and behavior and discovering user background knowledge.

2. ONTOLOGY CONSTRUCTION METHOD AND ALGORITHM

Personalized ontologies are a conceptualization model that formally describes and specifies user background knowledge. From observations in daily life, it is found that web users might have different expectations for the same search query. Sometimes even the same user may have different expectations for the same search query if applied in a different situation. A user's concept model may change according to different information needs. In this section, a model constructing personalized ontologies for web users's concept models is discussed.

3.1 World Knowledge Representation

World knowledge is important for information gathering. According to the definition, world knowledge is commonsense knowledge possessed by people and acquired through their experience and education. In ontology model, user background knowledge is extracted from a world knowledge base encoded from the Library of Congress Subject Headings (LCSH) [8]. The structure of the world knowledge base used is encoded from the LCSH references. The LCSH system contains three types of references: Broader term (BT), Used-for (UF), and Related term (RT). The primitive knowledge unit in our world knowledge base is subjects. They are encoded from the subject headings in the LCSH.

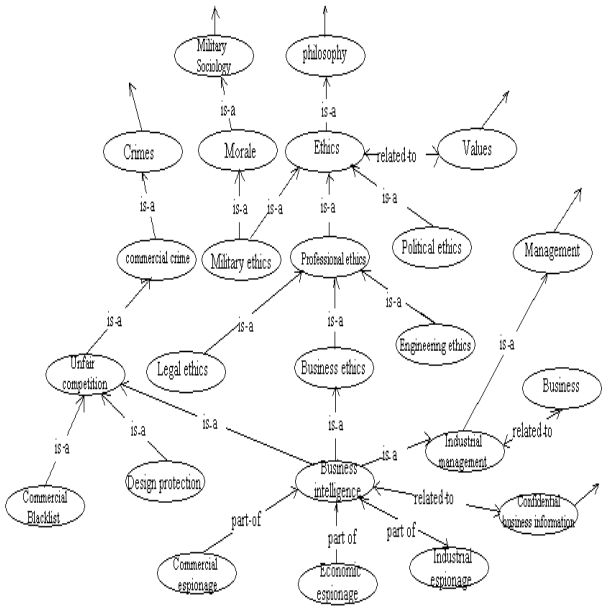

Fig. 1. A sample part of the world knowledge base. [9]

Definition. Let WKB be a world knowledge base, which is taxonomy constructed as a directed acyclic graph. The WKB consists of a set of subjects linked by their semantic relations, and can be formally defined as a 2-tuple WKB :=<S; R>,Where

- S is set of subjects S:= $<s_1; s_2; \ldots; s_m>$;
- R is a set of semantic relations R := $<r_1; r_2; \ldots; r_n>$ linking the subjects in S.

Fig. 1 illustrates a sample of the WKB dealing with topic "Economic espionage."

3.2 Ontology Construction

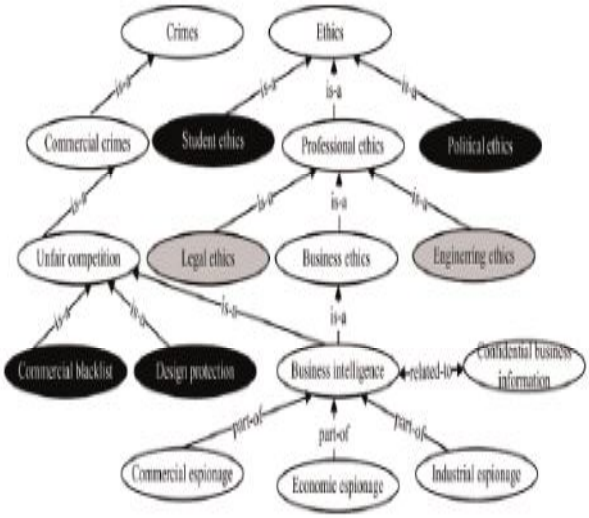

Fig.2. An ontology (partial) constructed for topic "Economic espionage" [9].

The subjects of user interest are extracted from the WKB via user interaction. A tool called Ontology Learning Environment (OLE) is developed to assist users with such interaction. Thus, for a given topic, the OLE provides users with a set of candidates to identify positive and negative subjects [9].

Fig. 2 illustrates the ontology (partially) 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.

3.3 Ontology Mining

Ontology mining discovers interesting and on-topic knowledge from concepts, semantic relations, and instances in ontology. A 2D ontology mining method is used: Specificity and Exhaustivity. Specificity (denoted *spe*) describes a subject's focus on a given topic. Exhaustivity (denoted *exh*) restricts a subject's semantic space dealing with the topic. This method aims to

investigate the subjects and strength of their associations in an ontology. Subject's specificity has two focuses: 1) on the referring-to concepts (called semantic specificity), and 2) on the given topic (called topic specificity) [9]. The determination of a subject's spe_a is described in Algorithm 1 [9].

Algorithm 1. Analyzing semantic relations for specificity :

Input: a ontology $O(T) = (tax^S, rel)$; a coefficient θ between (0,1)

Output: $spe_a(s)$ applied to specificity

1. set $k=1$, get the set of leaves S0 from tax^S, for $(s_0 \in S_0)$ assign $spe_a(s_0) = k$
2. get S_1 which is set of leaves in case we remove nodes S_0 and related edges from tax^S.
3. if $(S_1 = \emptyset)$ then return
4. for each $s_1 \in S_1$ do {
5. if $(isA(s_1) = \emptyset)$ then $spe^1_a(s_1) = k$
6. else $spe^1_a(s_1) = \theta \times \min\{ spe_a(s) / s \in isA(s_1)\}$
7. if $(part\ of(s_1) = \emptyset)$ then $spe^2_a(s_1) = k$
8. else $spe^2_a(s_1) = \frac{\sum_{s \epsilon part\ Of(s1)} spea(s)}{|partOf(s1)|}$
9. $spe_a(s_1) = \min(spe^1_a(s_1), spe^2_a(s_1))$; }
10. $k = k \times \theta$, $S_0 = S_0 \cup S_1$, go to step 2.

The topic specificity of a subject is investigated, based on the user background knowledge discovered from user local information. The strength of *i* to a subject s is determined by

$$str(i) = \frac{1}{priority(s,i) \times n(i)} \qquad (1)$$

The exhaustivity of a subject refers to the extent of its concept space dealing with a given topic. This space extends if a subject has more positive descendants regarding the topic. The ontological user profile is then used for web information gathering to make searching easier and relevant to user need.

3. CONCLUSION

Every user has a distinct background and a specific goal when searching for information on the Web. The goal of Web search personalization is to tailor search results to a particular user based on that user's interests and preferences. The ontology model constructs user

personalized ontologies by extracting world knowledge from the LCSH system and discovering user background knowledge from user local instance repositories. The ontological user profile can be used to the design of web information gathering systems. It can also be utilized in the fields of Information Retrieval, web Intelligence, Recommendation Systems.

4. REFERENCES

1. S. Gauch, J. Chaffee, and A. Pretschner, "Ontology-Based Personalized Search and Browsing" Web Intelligence and Agent Systems, vol. 1, 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. D. Downey, S. Dumais, D. Liebling, and E. Horvitz, "Understanding the Relationship between Searchers" Queries and Information Goals," Proc. 17th ACM Conf. Information and Knowledge Management (CIKM "08), pp. 449-458, 2008.

4. 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.

5. 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.

6. A. Doan, J. Madhavan, P. Domingos, and A. Halevy, "Learning to Map between Ontologies on the Semantic Web," Proc. 11th Int"l Conf , World wide Web(WWW"02),pp.662-673,2002.

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

8. E. Frank and G.W. Paynter, " Predicting Library of Congress Classification from Library of Congress Subject Heading", Information Science and Technology,vol.55,no.3 pp.214-227,2004.

9. Xiaohui Tao, Yuefeng Li, and Ning Zhong."A Personalized Ontology Model for Web Information Gathering" IEEE Transaction on knowledge and data Engineering, vol-23,no-4,,pp-496-509,2011.