ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 8, August 2016

Fuzzy Ontology Model for Gathering Web Information

[1] Sayali Dharmendra Behere [2] R. B. Wagh
[1] PG Student [2] Associate professor
[1][2]Department of Computer Engineering
R. C. Patel Institute of Technology Shirpur, Maharashtra, India

Abstract: **A model for knowledge description and formalization, these various on tologies 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 Fuzzy is a semi-automated collaborative tool for the construction of fuzzy ontology models. Fuzzy is an extension of the well known ontology model for which we have defined new meta classes to allow the definition of parameterized functions. Fuzzy also gives support to instantiate fuzzy concepts and roles. Fuzzy allows querying fuzzy ontologies based on fuzzy criteria. We present in this paper the Fuzzy Ontology Algorithm for gathering web related information we give some details on its implementation and also the way we use it to validate fuzzy ontologism.**

***General Terms*—Ontology, Local instance repository, User profiles**

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

I. 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 [1]. Simulation of user concept models is 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 [3].

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 produces effective performance for user background knowledge representation [4].

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 relyon data mining. These classification techniques for knowledge discovery [5].

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 [6].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[7].

A. The Concept of Ontology

An Ontology is the most important study of a nature of being existence, as well as the basic classification of being and also their relations. A model for knowledge

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 8, August 2016

description and formalization ontologism are widely used to represent user profiles in the form of personalized web information gathering. An ontology is defined as a set of representational primitives with which to model a domain of knowledge. The representational primitives are typically classes, attributes and relationships.

B. Fuzzy Ontology Model

The idea of fuzzy set and fuzzy logic theory was first proposed by Zadeh, as a mean of handling uncertainty. The appearance of a wide variety of methodologies for the construction of fuzzy algo motivates the application of fuzzy set theory in a great number of application fields. In this paper, we talk about the integration of fuzzy logic in ontology in order to define which we call fuzzy ontology. Ontology tools are most based on very accurate andcrisp logic and do not provide well-defined manner or means for expression fuzzyness. Ontology can be defined as a systematic description of part of relationships and entity dependencies. In other words, fuzzy ontology consists of a hierarchical description of important classes or concepts in a particular domain, along with the description of the properties of the particular instance of each concept. The Web Ontology Language (OWL) is a family of knowledge representation languages for authoring, ontologies and Description Logics (DL) are a family of knowledge representation languages which can be used to represent the terminological knowledge of an application domain in a structured and formally well understood way. Today description logic has become a corner stone of the Semantic Web for its use in the design of ontologies. The family of languages based on two semantics : OWL and DL. OWL lite semantics thar are based on Description Logics.

C. Architecture of ontology model

The architecture ontology model aims to discover user background knowledge and learns personalized ontologies to represent user profiles.

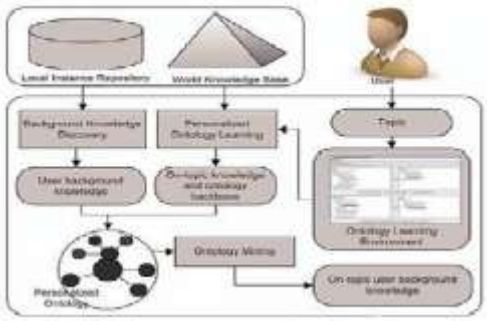

Fig. 1. Architecture of Ontology Model

Figure shows the architecture of the ontology model. A personalized ontology is constructed according to a given topic. Two knowledge resources the global world knowledge base and the users local instance repository are utilized by the model. The world knowledge base provides the taxonomic structure for the personalized ontology. The user background knowledge is discovered from the user local instance repository. Against the given topic, the specificity and exhaustivity of subjects are investigated for user background knowledge discovery [1].

II. LITERATURE SURVEY

Jiang and Tan proposes user information need acquisition many efforts have been involved to improve the accuracy and effectiveness. Closely related to work, user ontology consisting of both concepts and semantic relations are designed. Their goal is to represent and capture user's interests in target domain. Subsequently a method, they called Spreading Activation Theory (SAT) is employed for providing personalized services [1].

Li and Zhong proposes a term-based ontology leaning method for acquiring user information needs. Other work also realizes the importance of user information need, they treat user interest as implicit feedback and store in user profile. Also gives an automatic ontology learning method, in which a class is called a compound concept, assembled by primitive classes that are the smallest concepts and cannot be divided any further. Also used pattern recognition and association rule mining techniques to discover knowledge from user local documents for ontology construction [2].

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 8, August 2016

Zhong proposes a learning approach for task domain specific ontology, which employs various mining techniques and natural language understanding methods. It is an explicit specification of a conceptualization. Over the recent years, people who are mentioned below have often held the hypothesis that ontology-based approaches should perform better than traditional ones on IR, since ontologism are more discriminative and arguably carry more semantics. As a result, many researches concentrate on how to use ontology techniques. Also proposed a domain ontology learning approach that employed various data mining and natural-language understanding techniques [3]

Liu and Singh propose the Concept Net ontology and attempt to specify common sense knowledge. Concept Net does not count expert knowledge. Ontology is a collection of concept stands their interrelationships, which provide an abstract view of an application domain. An ontology called Onto Learn to mine the semantic relations among the concepts from web documents [4].

Trajkova and Gauch and Liu proposes a users profile from her/his browsing history, whereas they utilize ontological user profile on the basis of the users interaction with a concept hierarchy which captures the domain knowledge and also require the user to specify a profile manually. In short these work aim to enhance search performance through asking users explicit feedback such as preferences or collected implicit feedback, which are normally either expensive in extraction or inaccurate in description [5].

Jin et al. proposes integrated data mining and information retrieval techniques to further enhance knowledge discovery, and also categorized user profiles into two diagrams, the data diagram user 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 [6].

Navigli et al. use ontology references based on the categorization of online portals and propose to learn personalized ontology for users. It is built based on the Dewey Decimal Classification(DDC) system and attempt to describe the background knowledge. Unfortunately, the previous work on ontology learning covers only a small size of concepts, where mainly uses Is-A (super class, or sub-class) relation in the knowledge backbone. They don't consider mining and characterizing knowledge in a concept level rather than domains. To extend these methods, the backbone of personalized ontologism is been determined to build a real hierarchical structure by applying information in a world knowledge repository [7].

Sieg et al. proposes personalized ontologism from the Open Directory Project to specify user's preferences and interests in web search. On the basis of the Dewey decimal classification, also developed Intelli Onto to improve performance in distributed web information retrieval. Wikipedia was used to help understand underlying user interests in queries. These works effectively discovered user background knowledge however, their performance was limited by the quality of the global knowledge bases. Aiming at learning personalized ontologies, many works mined user background knowledge from user local information [8].

Shehata et al. proposes user information needs at the sentence level rather than the document level, and represented user profiles by the Conceptual Ontological Graph. The use of data mining techniques in these models leads to more user background knowledge being discovered. However, the knowledge discovered in these works contained noise and uncertainties. Additionally, ontologies were used in many works to improve the performance of knowledge discovery [9].

Lau et al. proposes a concept to construct maps based on the posts on online discussion forums. Also used ontologism to help data mining in biological databases. Also acquire user profiles by observing user activity and behavior and discovering user background knowledge. A typical model is OBIWAN, which acquires user profiles based on user's online browsing history. Also developed Onto Learn to discover semantic concepts and relations from web documents [10]

Doan et al. proposed a model called GLUE and used machine learning techniques to find similar concepts in different ontology's, and also proposes a framework for learning domain ontology's using pattern decomposition, classification, and association rules mining techniques.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 8, August 2016

These works attempted to explore a route to model world knowledge more efficiently [11].

Van der Sluijs and Huben proposed a method called the generic user model component to improve the quality and utilization of user modeling. Wikipedia was also used by to help discover user interests. The interviewing, semi interviewing, and non interviewing user profiles can also be viewed as manual, semiautomatic and automatic profiles respectively. Ontology Based Information Web Agent Navigation(OBIWAN) approach is to distribute the various information sources. The idea is similar to web rings. Websites are clustered intoregions [12].

Chirita et al. used a collection of user desktop text documents and emails, and cached web pages to explore user interests. Acquired user profiles by a ranked local set of categories, and then utilized web pages to personalize search results for a user. These works attempted to acquire user profiles in order to discover user background knowledge. User profiles can be classified into three groups interviewing, semi interviewing and non interviewing [13].

Chang proposes a similar non-invasive learning approach for constructing web user pro_les.A user problem consists of two components, a Web Access Graph(WAG) and a Page Interest Estimator(PIE). The WAG captures the Web page access patterns of a user. Based on the content of Web pages, a PIE learned from the users access behavior characterizes theinterests of the user [14].

Xiaohui Tao proposes a three-descriptor representation to monitor user interest dynamics. This model maintains a long-term interest descriptor to capture users general interests and a short-term interest descriptor to keep track of users more recent faster changing interests [15].

A. Fuzzy Ontology Generation

A fuzzy ontology is a quintuple F = <I,C,T,N,X> where, I is the set of individuals objects also, called instances of the concepts, C is a set of concepts or also classes. Each concept is a fuzzy set on the domain of instanes. The set of entities of the fuzzy ontology is defined by E = C Ū I. T denotes the fuzzy taxonomy relation among the set of concepts C. Also it organizes concepts into sub super concept tree structures. The taxonomic relationship T(i,j) indicates that the child j is a conceptual specification of the parent i with a certain degree. N denotes the set of non-taxonomy fuzzy associative relationship that relate entities across tree structure for example:

Naming relationship, describing the name of concepts Locating relationships, describing the relative location of concepts Functional relationships, describing the functions or properties of concepts

X is the set of axioms in a proper logical language, i.e. predicates that constrain the meaning of concepts, individuals, relationships and functions.

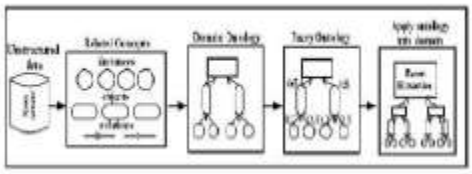

Fig. 2. The process of constructing a fuzzy ontology as extended domain ontology

Fuzzy ontology can be seen as an extended domain ontology , which makes use of the specific domain and fuzzy information processing as follows:

- The input is unstructured data
- (ii)The definition of related concepts in the domain, e.g. instances, objects and their relationships
- The generation of domain ontology
- The domain ontology extended as fuzzy ontology
- Applying the fuzzy ontology to the specific domain.

III. EXPERIMENTAL RESULTS

The experimental performance of the models was measured by three methods, achieved by Ontology Model

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 8, August 2016

and Trec Model. Performance measured by three methods namely,

- 11 standard Precision Recall (11SPR)
- Mean Average Precision (MAP)
- F1 – Measure

These modern methods are based on Precision and Recall methods which are basic methods to measure information gathering performance. Precision : Precision is the ability of a system to retrieve only relevant documents or words. Recall : Recall is the ability of a system to retrieve all documents or words.

11 Standard Precision Recalls

An 11SPR value is computed by summing or adding the precisions values at the specified recall cutoff points, and then dividing by the each number of topics or their related documents.

$\Sigma i=1 \wedge N$ Precisionλ/N ; $\lambda = \{0.0, 0.1, 0.2\ldots.., 1.0\}$

Where N = Number of Topics

λ = Indicates the topics which are linked with subtopics.

At each λ point average precision value over N topics are calculated, then perform Precision and Recall to measured performance of gathered information.

2. Mean Average Precision

Mean average precision method is also general purpose information gathering. The average precision for each topic is the mean of the precision obtained for each relevant document is retrieved.

3. F1 Measure

Measure averages the precision and recall and then calculates the F1 measure. Measure calculates the F1 measure for each returned result and averages the F1 measure to measure performance of gathered web related information.

Fig. 3. Results by Trec Model and Improved Ontology Model

Figure 3 shows the information gathering performance achieved by using Ontology Model to that achieved by using the Trec model which also called as Golden Model. To show that improved Ontology Model To achieve user's required information which gives relevant or proper information through ontology model.

Number of Topic Found = 2, Number of Subtopics Found = 6

Methods to Calculate Performance	Topic	Subtopic
Precision	0.8333	0.1666
Recall	0.5555	0.3333
F1 Measure	0.6666	0.6666

Table 1: Calculate Performance of Methods

Table 1 values shows the percentage change in performance is used to compute the difference in Mean average precision and F1 measure results obtained between the Ontology model and target model namely Trec Model gives results of gathered information which

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 8, August 2016

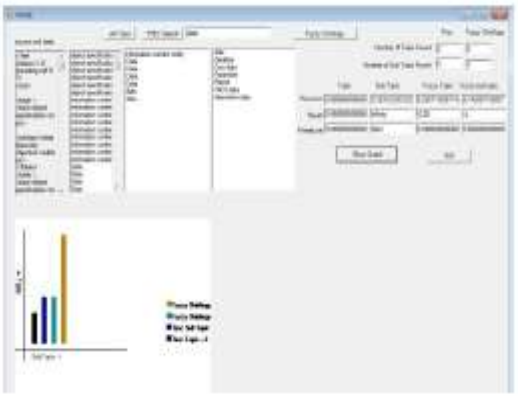

Fig. 4. Results by Fuzzy Ontology Model

Figure 4 shows the information gathering performance achieved by using Fuzzy Ontology Model to that achieved by using the Trec model which also called as Golden Model. To show that Fuzzy Ontology Model To achieve user's required information which gives relevant or proper information through ontology model, using threshold value keep consisting a value of threshold value take in between 10 to 100 percent threshold value. Number of Topic Found = 3, Number of Subtopics Found = 7

We have to take threshold value upto 80% for fuzzy ontology model

Methods to Calculate Performance	Topic	Subtopic
Precision	0.8571	0.1428
Recall	0.25	0.5454
F1 Measure	0.6666	0.6666

Table 2 values shows the percentage change in performance is used to compute the difference in Mean average precision and F1 measure results obtained fuzzy ontology model.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, Ontology Model is successfully implemented; Searching Techniques of Ontology Model gives relevant information which is required for all user's. An ontology model is proposed for representing user background knowledge for personalized web information gathering. The proposed ontology model in this paper provides web information which is required by user's in their personal interests. The improved ontology model also has contributions to give the fields of Information Retrieval systems, web information gathering system and also Recommendation Systems and Information Systems. In our future work, we will investigate the methods that generate user local instance repositories means it will find the better searching technique using the same our ontology model using slight change in algorithm namely Fuzzy Algorithm to match there presentation of a global knowledge base or to match exact word which is in user's personal interests. The present work assumes that all user local instance repositories have content-based descriptors referring to the subjects or user's related documents, however a large volume of documents or user's related information which is in user's personal interests existing on the web may not have such documents. These all strategies will be investigated in future work to solve this problem or drawback. The investigation will extend the applicability of the ontology model to the majority of the existing web documents or words which are user's related and increase the contribution and significance or importance of the present work.

REFERENCES

[1]. X. Jiang and A.-H. Tan, "Mining Ontological Knowledge from Domain-Specific Text Documents", in the proceedings of Fifth IEEE Int'l Conf. Data Mining (ICDM '05), pp. 665-668, 2005.

[2]. 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.

[3]. N. Zhong, "Representation and Construction of Ontologies for Web Intelligence", Int'l J Foundation of Computer Science, vol. 13, No. 4, pp. 1-14, 2003.

[4]. Y. Li and N. Zhong, "Web Mining Model and its Application for Information Gathering", Knowledge Based Systems, Vol. 17, pp. 207-211, 2004.

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

[6]. W. Jin, R.K. Srihari, H.H. Ho, and X. Wu, "Improving knowledge Discovery in Document Collections through Combining Text Retrieval and Link Analysis Techniques", in Proc. Seventh IEEE Int'l Conf. Data Mining(ICDM'07), pp. 193-202, 2007.

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

[8]. A. Sieg, B. Mobasher, and R. Burke, "Web Search Personalization with Ontological User profiles", in Proc. Of the 16th ACM conf. Information and knowledge Management (CIKM '07), pp. 525-534, 2007.

[9]. S. Shehata, F. Karray, and M. Kamel, "Enhancing Search Engine quality Using Concept-Based Text Retrieval", in Proc. IEEE/WIC/ACM Int'l Conf. Web Intelligence (WI '07), pp. 26-32,2007.

[10]. R.Y.K. Lau, D. Song, Y. Li, C.H. Cheung, and J. X. Hao, "Towards a Fuzzy Domain Ontology Extraction Method for Adaptive e-Learning", IEEE Trans. Knowledge and Data Eng., Vol. 21, No. 6, pp. 800-813, June 2009.

[11]. 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

[12]. K. van der Sluijs and G.J. Huben, "Towards a Generic User Model Component", Proc. Workshop Personalization on the Semantic Web Conf. (PerSWeb '05), 10th Int'l Conf. User Modeling (UM '05), pp. 43-52, 2005.

[13]. P.A. Chirita, C.S. Firan, and W. Nejdl,"Personalized Query Expansion for the Web", Proc. ACM SIGIR ('07), pp. 7-14, 2007.

[14]. J. Han and K.C.-C. Chang, "Data Mining for Web Intelligence", Computer, Vol. 35, No. 11, pp. 64-70, Nov. 2002.

[15]. Y. L. Xiaohui Tao and S. M. Ning Zhong, "A Personalized Ontology Model for Web Information Gahtering, in the proceedings of IEEE Transactions on Knowledge and Data Engineering, Vol. 23, pp. 496-511, April. 2011.