Journal of Theoretical and Applied Information Technology
10th July 2014. Vol. 65 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A NOVEL APPROACH FOR RANKING ONTOLOGIES BASED ON THE STRUCTURE AND SEMANTICS

[1]R.SUBHASHINI, [2]DR.J.AKILANDESWARI
[1]Assistant Professor (SG), Department of Computer Science and Engineering, Sona College of Technology
[2]Professor and Head, Department of Information Technology, Sona College of Technology
E-mail: [1]subhageorg@gmail.com, [2]akila_rangabashyam@yahoo.com

ABSTRACT

Semantic web is based on knowledge representation which contains a large number of ontologies. The increasing demand for ontology had triggered a growing number of usable ontology in web. This is mainly used for ontology merging, ontology mapping, and for reusing purpose. In order to get a solution for this problem, ontology search results need to be ranked. This ranking method increases the scope of the knowledge searching in ontology-driven searches. At present many ranking techniques are available. By exploring the advantages and weakness of a AKTive ranking algorithm in the semantic web , this paper proposes a new ranking algorithm named Onto-DSB ranking based on the semantic web link and the internal structure of ontology , the way which is achieved by introducing new measures based on its relation set. Experimental results indicate that this algorithm is more effective and satisfies the needs of the user.

Keywords: *Semantic Web, Ontology Merging, Ranking Algorithm, Reusable Ontology, Onto-DSB Ranking Technique, Internal Structure, Semantic Web Link.*

1. INTRODUCTION

Semantic Web [1] is the extension of World Wide Web, aimed to ensure a better understanding of the information and provides a good knowledge representation. It contributes several mechanisms that can be used to classify information and its context retrieving information on web. It has a number of Resource Framework Description (RFD) which overcomes the difficulties in understanding and integrating the information.

Ontology is a formal, explicit specification of shared conceptualization [7]. An increasing number of ontologies are being developed, and their reuse and sharing offer several benefits. One important benefit is that heterogeneous systems and resources can interoperate seamlessly by sharing a common ontology [25]. A number of ontology libraries and search engines are in existence to facilitate retrieval of potentially relevant ontologies and provides a domain-related ontology to depict the real world applications. There is a set of standard web ontology language (OWL) [2] which is based on RDF model to describe the concepts explicitly with their relationship [1].

The main problem with the ontology construction is the requirement of high cost for building the ontologies. The time required for gathering complete knowledge about a specific domain is more. Hence, no guarantee is given for the resulting ontology to be better than the existing one. Therefore, an approach is proposed for reusing the existing ontologies in the construction of new ontologies.

The search engines play a vital role in retrieving the information required by the user. However, the retrieved web pages also contain ineffective or irrelevant information. The latest web architecture represented by semantic web overcomes this limitation by applying the ranking algorithms. The ranking algorithm extracts the information on the user queries from the semantic search engine and provides the desired result. However, in order to rank results, most of the existing solutions need to work on the whole annotated knowledge base.

This paper analyzes the internal structure of ontology and tries to overcome some of the disadvantages of AKTive Rank algorithm in the semantic web. The limitations of AKTive ranking is addressed by introducing new measures that result in building high quality ontologies. Hence this paper results in a better ranking algorithm based on the application.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The rest of the paper is organized as follows; section 2 gives brief study of related work. Section 3 gives the detailed discussion on the proposed system. The experimental result is presented in section 4. The last section provides conclusion and future work.

2. RELATED WORK

Swoogle is currently dominating the knowledge representation area of development indexing which leads to an increasing number of ontologies covering a wide range of domains [2]. Ontologies are the backbone of semantic web and it represents domain knowledge in the semantic web [26]. The suitable ontologies for the particular domain application are retrieved by ranking ontologies. The idea of ranking ontologies is not only considered as a solution for query expansion scenarios, but for a wider scope of knowledge searching in ontology-driven searches [5].

Some of the techniques used to rank ontology are followed with the description:

2.1 AKTive Rank Algorithm

AKTiveRank [2] is a technique for ranking ontology based on different analytical measures that assess the ontology in terms of depth of coverage. Users can use ontology search engine (e.g. Swoogle) for searching. The query submitted to the search engine is used by AKTive Rank to identify the concepts that match the user's request. The ranking measures applied by AKTive Rank will be based on the representation of those concepts and their neighborhoods. It increases the time complexity[24]

2.2 Content-based Ontology Rank Algorithm

The content-based ontology ranking algorithm [3] obtains a list of ontologies from a search engine. Based on the term given by the knowledge engineer the retrieved ontologies are ranked. The ranking is done according to the number of concept labels in those ontologies which matches a set of terms extracted from a Word Net. It is done related to the domain of knowledge identified by the knowledge engineer's original search terms.

2.3 Onto Rank Algorithm

The Onto Rank algorithm [8] applies the link analyze method. Here two concepts are considered as a reference relationship "if and only if" a relationship exists between the two classes in a relation set [1]. The reference relations are directional and transitive. It evaluates the importance of ontology in a static manner and doesn't consider the user query as an effective factor in ranking the results.

2.4 OS_Rank Algorithm

Ontology Structure Ranking (OS_RANK) [24] ranks the ontologies based on its semantic relation and structure. The overall ranking criteria are based on the three ranking scores:

- Ranking based on class name
- Ranking based on semantic relation
- Ranking based on ontology structure.

These measures are applied to retrieved ontology from search engine based on the user query and ranking is performed. The user can decide the weights of the ranking measure according to the needs and importance of their applications [24].

2.5 SIF Rank Algorithm

The Semantic-aware Importance Flooding (SIF RANK) [25] retrieves the OWL ontology and converts them into directed graph. The iteration fix point computation is done in each graph to calculate the importance of nodes. It is based on the nine kinds of patterns, semantically treated correct. This computation reaches the maximum number of iterations and the normalization is done to neglect the nodes which are not semantically linked [25].

3. PROPOSED WORK

3.1 System Architecture

In the proposed work, the ranking algorithm with the improved measures is introduced which enables reusing ontologies. The phases involved in this work are listed below:

a) Retrieval of URI's and Database management.
b) Computation of logical measures.
c) Combined ranking computation.

Journal of Theoretical and Applied Information Technology
10th July 2014. Vol. 65 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

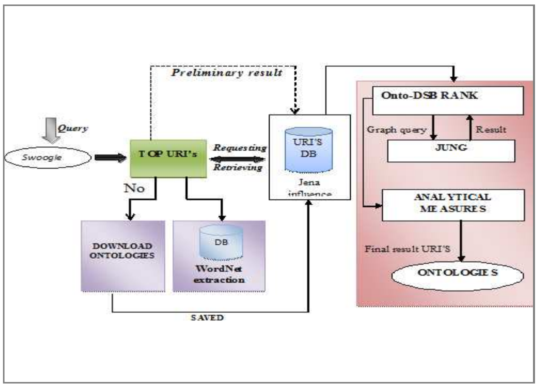

Figure 1. System Architecture

3.2 Module Description

3.2.1 Retrieval of URI's and Database management

The user post his/her query to the search engine SWOOGLE. Based on the content, it checks the URI's database to retrieve the related ontologies, if not found it will download the required ontologies and save it in the database. The URI's in the Database is given as input to the Inter-onto rank algorithm.

3.2.2 Computation of logical measures

3.2.2.1 Onto- DSB Ranking

Onto-DSB is a ranking algorithm based on the logical measures to calculate the concept coverage of particular keyword in ontology and to retrieve the best ontology for the users' based on their search. The steps involved in this algorithm are:

- The ontologies in the database are given as input to the graphical file, which converts RDF files into structural format.
- Three logical measures are applied to the structural format. The measures are:
  Depthness Measure (DTM)
  Semantic Informative Measure (SIM)
  Betweenness Measure (BEM)

3.2.2.2 Depthness Measure

Depthness measure has two main steps. In the first step it looks for the classes in the ontology that have terms matching the searching term either exactly or partially and count their total number, where the exact match is considered as better than the partial match. Next, the ontology classes are classified based on the three criteria listed below:

- Number of subclasses based on the search term.
- Length of the classification.
- Relation between the classes in the ontology class set.

From the above steps, the Depthness measure is calculated which serves as the input for combined ranking [1] using the formula:

$$E_a[c] = \sum_{i=1}^{3} w_i |A_i| \qquad (1)$$

$$P_a[c] = \sum_{i=1}^{3} w_i |A_i| \qquad (2)$$

$$CMM[O] = \frac{1}{n+m}\left\{\sum_1^n E_a * \mu + \sum_1^m P_a * \beta\right\} \qquad (3)$$

$\mu = 0.6$; $\beta = 0.4$;

Where,
Let C[o] is the set of classes in the ontology o, and t is the set of search terms.
E_a (o, t) is the set contains the exact match classes in the particular ontology with the query term.
P_a (o, t) is the set contains the partial match classes in the particular ontology with the query term.
n is the number of exact match
m is the number of partial match.

3.2.2.3 Semantic Informative Measure

Semantic Informative measure is the summation of edge weight along the shortest path between them. The link strength is calculated through the conditional probability occurs between the two nodes along with the strength of union class, equivalent class in the RDF file of an ontology. This can be done using the formula:

$$wt(C_c, C_p) = \left(\frac{d(C_p)+1}{d(C_p)}\right)[IC(C_c) - IC(C_p)]T(C_c, C_p) \qquad (4)$$

$$sim(C_i, C_j) = \sum_{c \in \{path(c_i,c_j) - LSuper(c_i,c_j)\}} wt(c, parent(c)) \qquad (5)$$

Where,
P (C_p) & P (C_c) are the probability of instance parent concept and instance of child concept.
IC(c) – the information content of concept c.
LS – Link strength.
T (C_c, C_p) – link relation / type factor.

3.2.2.4 Betweenness Measure

Betweenness measure calculates the number of shortest path that pass through each node in the graph. The nodes that are inter- bonded with all the

Journal of Theoretical and Applied Information Technology
10th July 2014. Vol. 65 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

shortest paths will be scored higher than others. In the following, n is the number of matched class in the ontology 'o'. The formula used to calculate betwenness [1] is:

$$sbm(c) = \sum_{c_i \neq c_j \neq c \in C[o]} \frac{\sigma_{c_i c_j}(c)}{\sigma_{c_i c_j}} \quad (6)$$

$$bem(c) = \sum_{c_i \neq c_j \neq c \in C[o]} \frac{\sigma_{c_i c_j}(c)}{\sigma_{c_i c_j}} \quad (7)$$

$$BEM(o) = \frac{1}{n} \sum_{k=1}^{n} bem(c_k) \quad (8)$$

Where,

$\sigma_{c_i c_j}$ = shortest path from c_i to c_j

$\sigma_{c_i c_j}(c)$ = the number of sfhortest paths from c_i to c_j that passes through c

n is the number of routes exist in ontology o

BEM (o) is the average Betweenness value for ontology o.

3.2.2.5 Combined Ranking Computation

The overall ranking measure of Onto-DSB rank algorithm is calculated after applying all the three measures to the ontology. The final score [1] is calculated using the formula:

$$\text{Total score } (o \in O) = \sum_{i=1}^{3} w_i \frac{M[i]}{max_{1 \le j \le |o|} M[j]} \quad (9)$$

Let M = { M[1] , M[2] , M[3]}={DEM, SIM , BEM}

W_i – weight factor.

O – The set of ontology to rank

4. Experiment Result

In this section the result analysis is done for the domain "student". The running Inter-onto ranking technique retrieves the list of ontology for the query "student type: OWL". These ontologies are downloaded from the search engine Swoogle. Some of the ontologies retrieved from the Swoogle are duplicated which are neglected from the ranking process. The duplicate OWL files are noted as ("-") in the following list of ontologies which are ranked. Those duplicates are dropped from the experimental analysis. The list of owl files is mentioned in table 1.

Table 1. List of Owl files from Swoogle

Search Result from Swoogle For the query "student type: OWL"(URI)	
a	http://annotation.semanticweb.org/iswc/iswc.owl
-	http://semweb.mcdonaldbradley.com/OWL/Cyc/FreeToGov /060704/FreeToGovCyc.owl
b	http://www.openmetadir.org/om2/prim-3.owl
-	http://counterterror.mindswap.org/2005/terrorism.owl
-	http://www.csd.abdn.ac.uk/~cmckenzi/playpen/rdf/ akt_ontology_LITE.owl
-	http://protege.stanford.edu/plugins/owl/owl-library/ka.owl
c	http://www.mindswap.org/2004/SSSW04/aktive-portal -Ontology-latest. Owl
d	http://www.tt.cs.titech.ac.jp/~fukatani/University/HU.owl
e	http://protege.stanford.edu/plugins/owl/owl-library/koala.owl
f	http://www.tt.cs.titech.ac.jp/~fukatani/University/TMDU.owl
g	http://www.tt.cs.titech.ac.jp/~fukatani/University/TITech.owl
h	http://www.informatik.uni-bremen.de/~shi/Lehre/lang-tech-bremen-05/student-work/semantic-rep-assignment/du-liang/semantic_analysis/robot -world.owl
i	http://www.mindswap.org/2004/multipleOnt/Factored Ontologies/ItalianUniversities/ita_partition1.owl
j	http://www.cs.toronto.edu/semanticweb/maponto /MapontoExamples/univ-cs.owl
-	http://www.lehigh.edu/~zhp2/2004/ 1028/bib-bench.owl
l	http://www.historiographus.org/owl/ histemm.owl
m	http://owl.cs.manchester.ac.uk/2007/07/sssw/ university2.owl

When the Onto-DSB ranking technique was applied for the ontologies in the table 1, it produces the normalized result of three measures for each ontology. In this result the duplicates are not considered for ranking. Finally the result for each ontology is weighed and the combined score is calculated and prioritized in figure 2. The normalized scores are displayed in table 2.

Journal of Theoretical and Applied Information Technology
10th July 2014. Vol. 65 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2. Logical Measure Scores

Ontology name	DEM	BEM	SIM
a	0.967	0.515	0.208
b	0.803	0.562	0.478
c	0.724	0	0.472
d	0.264	0.210	0.277
e	0.490	0.421	0.801
f	0.289	1	1
g	0.546	0.513	0.459
h	0.264	0.557	0.694
i	1	0.397	0.781
j	0.222	0	0.484

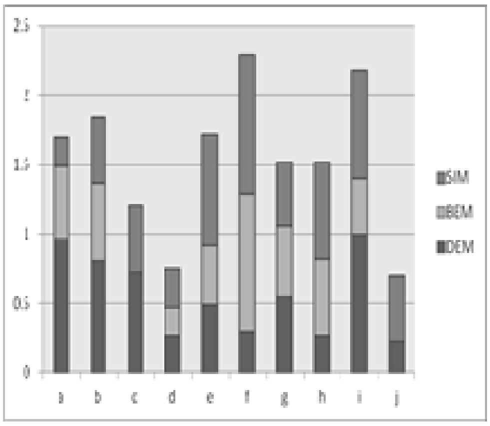

Figure 2. Comparison Of Logical Measure Score

The ranking weights for Depthness measure, semantic informative measure and Betweenness measure were set to 0.5, 0.2, 0.3 to calculate final ranking. The most importance is given for the exact match and the Depthness covered the user query in the particular owl file. Figure 3 represent the total scores for the top ten ontologies.

Table 3. Total Score And Their Rank

ontology name	Total score	Rank
a	0.680	2
b	0.666	3
c	0.456	7
d	0.250	9
e	0.532	5
f	0.644	4
g	0.519	6
h	0.438	8
i	0.775	1
j	0.207	10

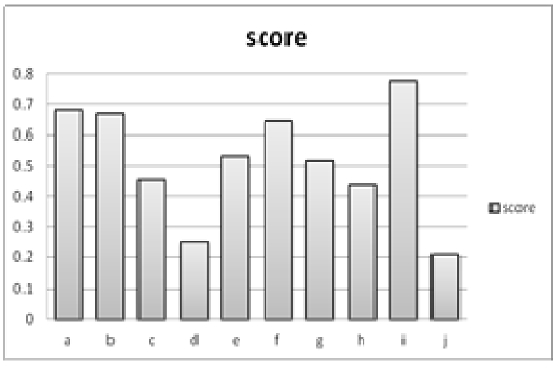

Figure 3. Total Score For Each Ontology

Finally the experimental results are evaluated by using the *recall rate* and *precision call*. The precision is evaluated at a given cut-off rank, considering only the topmost result returned by the system. Recall in information retrieval is the fraction of documents that are relevant to the query that are successfully retrieved. The preference measure for each ontology in the prescribed set is computed from both existing and proposed perspectives in figure 4. Precision-Recall measure with respect to the Swoogle and proposed ranking method are compared in figure 4.

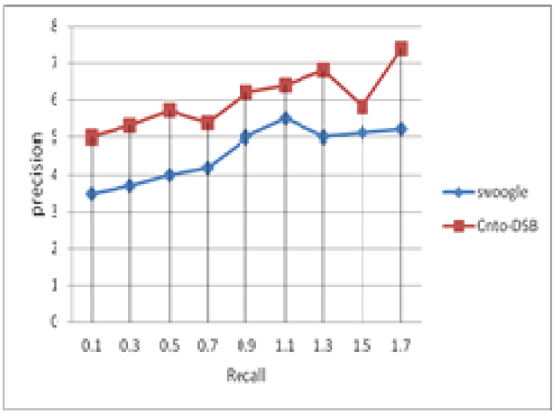

Figure 4. Precision And Recall

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. CONCLUSION AND FUTURE WORK

The set of ontology from a domain "student" based on the user query is retrieved from Swoogle search engine. The ranking algorithm is applied in those retrieved ontologies and the final result is obtained which has a set of top ranked ontologies. These resulted ontologies obtained are proved that Onto-DSB ranking method gives importance for both semantic web link and for its internal structure to get a better result than Swoogle and AKTive Ranking technique. The re-ranked ontologies are suggested for reusing purpose.

The logical measures applied in the Onto-DSB ranking algorithm can be enhanced by changing the factors in those measures.

REFERENCES

[1] Alani, H., and Brewster, C.,"Metrics for Ranking Ontologies", *Proceedings of the 4th Int. Workshop on Evaluation of Ontologies for the Web* (EON'06), at the 15th Int. World Wide Web Conference (WWW'06). Edinburgh, UK, 2006

[2] Alani.H and Brewster.C (2005), "Ontology Ranking based on the Analysis of Concept Structures". *In Proceedings of the 3rd International Conference on Knowledge Capture, Banff, Alberta, Canada.*, 2005

[3] Chintan Patel, Kaustubh Supekar, Yugyung Lee, E.K.Park, " A Semantic Portal for Ontology Searching, Ranking and Classification".

[4] Edward Thomas, Harith Alani, Derek Sleeman, Christopher Brewster,"Searching and Ranking Ontologies on the Semantic Web", 2007.

[5] Gang Wu, Juanzi Li , "SW: An Approach for Ranking Semantic Web Reversely and Consistently" , *Third International conference on semantics , knowledge and grid,*2007

[6] Guangfei Yang, Kaoru Shimada, Shingo Mabu and Kotaro Hirasawa," A Personalized Association Rule Ranking Method Based on Semantic Similarity And Evolutionary Computation" , *IEEE conference,.2010*

[7] Harith Alani ,"Ontology Construction from Online Ontologies", *In proc. 15th International World Wide Web Conference,2006.*

[8] Harith Alani, Christoper Brewster and Nigel Shadbol ,"Ranking Ontologies with AkTive Rank",2007.

[9] Huang Xiao-xi, Zhou chang-le ,"An OWL-based WordNet lexical ontology" *journal of Zhejiang university science*,2007.

[10] Iván Cantador, Miriam Fernández, Pablo Castells, "Improving Ontology Recommendation and Reuse in Web CORE by Collaborative Assessments", 2007

[11] Jarunee Saelee, Veera Boonjing," A Metadata Search Approach with Branch and Bound Algorithm to Keyword Query in Relational Databases ", *Fourth International Conference on Computer Sciences and Convergence Information Technology*, 2009

[12] Matthew jones, Harith Alani , "Content-based Ontology Ranking",*9th International Protégé Conference*, 2006

[13] Medelyan O and Witten I.H, "Thesarus based automatic key phrase indexing", *Proc. Of the Joint Conf. on digital Libra*, 2006

[14] Neda alipanah , piyush Srivastava , Pallabi Parveen, Bhavani Thraisingham, " Ranking Ontologies using Verified Entities to Facilitate Federated Queries" , *International Conference on Web Intelligence and Intelligent Agent Technology*.

[15] Riad.A.M,Hamdy K.Elminir,Mohamed Abdu Elsoud,Sahar.F.Sabbeh,"An Architecture for Personalised Search engine",*International Journal on Advances in Information Sciences and Service Science, 2010*

[16] Sanjay Kumar Malik , Nupur Prakash , S.A.M. Rizvi, "Ontology Merging using Prompt plug-in of Protégé in Semantic Web" , *International Conference on Computational Intelligence and Communication Networks*, 2010

[17] Sarumathiy.C.K, R.Subhashini, Dr.J.Akilandeswari, "Ontology Integration from Existing Ontologies using the improved Ranking algorithm", *International Conference on Recent advances in Technology, Engineering, Management and Science*, 2010

[18] Tina Eliassi-Rad and Edmond Chow ,"Using Ontological Information to Accelerate Path-Finding in Large Semantic Graphs: A probabilistic Approach" American Association for Artificial Intellegence",2006

[19] Wei Yu, Jiaheng Cao, Junpeng Chen , "A Novel Approach for Ranking Ontologies on the Semantic Web" , *1st International Symposium on Pervasive Computing and Applications*, 2006

[20] Wei Yu, Junpeng Chen, "Ontology Ranking For the Semantic Web" , Third International Symposium on Intelligent Information Technology Application, 2009

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[21] Xiong Jianbin, LI Zhenkunl, FAN Luqiao, LIU Dongpingl, "Improved Depth First Algorithm and Its Application in Information Retrieval" *In Proc. Conference*, 2010

[22] Yixi Chen , Keting Yin and Xiaohu Yang , " A Semantic-Aware Algorithm to Rank Concepts in an OWL Ontology Graph" , *Information Technology Journal*, 2010

[23] Zhiguo Ding, Bing-ru Yang, De-quan ZhengjieDuan ,"Improved Ontology Ranking Algorithm Based on Semantic Web",*3rd IEEE International conference on Digital Object Identifier*,2010

[24] Subhashini.R, J.Akilandeswari, Sinthuja.V, "A review on ontology ranking algorithms", *International journal of Computer Application*, nov 2011.

[25] Jinsoo Park, Sunjoo Oh, Joongho Ahn, "Ontology Selection Ranking Model for Knowledge Reuse," *Expert Systems with Applications* 38 (2011) 5133–5144

[26] Vandana Dhingra, "Comparative Analysis of Ontology Ranking Algorithms," *International Journal of Information Technology and Web Engineering*, July -September 2012 7(3), 55-66