

Concept Prioritization in Ontology Matching- Concept Type Classification based Approach

Arthi J^a, Saruladha K^b

^aNational Institute of Technology Puducherry, Karaikal.
arthighp@gmail.com

^bPondicherry Engineering College, Puducherry.

The growth of semantic web has led to the development of multiple ontologies in a single domain giving rise to the heterogeneity problem. For effective utilization of these ontologies, ontology matching, a technique which helps to determine correspondences among the concepts in the ontologies was evolved. The constant growth in number and size of ontologies used in the semantic web enhances the need for search space optimization. This can be handled by prioritizing the concepts that are to be considered for matching. The existing approaches perform concept prioritization by determining concept importance, determining anchors and so on. In this paper a novel method for concept prioritization is being dealt. Initially, the concepts are being classified into four novel concept types which focus on the structural, semantic properties and attributes, the concepts possess. The concept importance determination is done for the type classified concepts using the proposed concept type classified concept importance metric which considers inter and intra ontology concept importance. Ontology matching determine correspondences based on similarities among concepts. The computation of similarities using a single matcher may not always identify all possible matches and hence several similarity matchers are employed. The similarity matchers that are used for matching in the this approach are chosen by comparing several string, linguistic and structural matchers and the best among them are determined by comparing the determined similarities with human judgments. The proposed system is evaluated using the OAEI data sets. The results obtained are compared with the existing approaches and a significant increase in the number of significant concepts retrieved and an increase in correlation co-efficient was observed for the obtained similarity matchers thus increasing the accuracy.

Keywords : Concept Importance, Concept Type, Ontologies, Ontology Matching, Similarity Matchers.

1. INTRODUCTION

Ontologies are widely being used to enrich the semantics of the web. Ontology is defined as "a formal, explicit specification of a shared conceptualization" [1]. Ontologies incorporate a collection of discrete entities such as classes, properties, individuals *etc.*, The usage of ontologies in the task of information retrieval apart from world wide web has led to the development of various ontologies for a single domain leading to the problem of heterogeneity. This is because of the different perspectives of various designers in modelling ontology. Ontology matching aims to handle the problem of heterogeneity by finding corre-

spondences among the components of the entities involved among these ontologies. Ontology matching has several problems which include search space optimization, increase in computational complexity, selection of similarity matchers for determining correspondences, uncertainty that arise due to the aggregation of similarity values *etc.*, .

Search space optimization has been a long handled issue in ontology matching. Several works like using anchors for refining the search space, partitioning ontologies to determine the matching candidates, selection of candidate matches based on various criteria, determination of importance among concepts to identify the con-

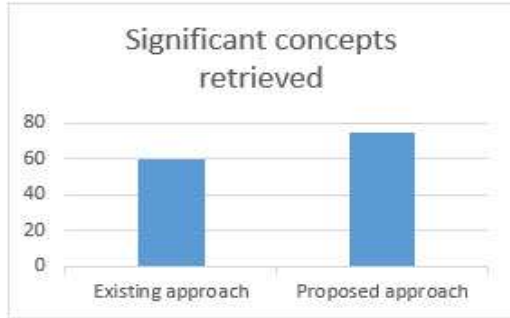


Figure 3. Comparison of Significant Concepts Retrieved



Figure 4. Comparison of Correlation Co-Efficient for String Similarity

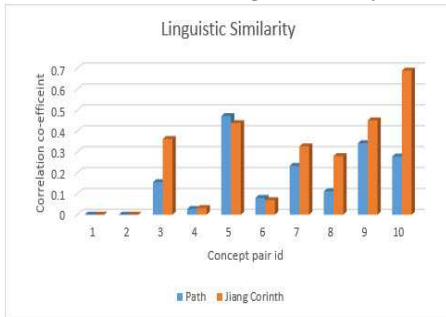


Figure 5. Comparison of Correlation Co-Efficient for Linguistic Similarity

9. CONCLUSIONS

The optimization of search space has been handled by several works. But the effective means of prioritizing concepts have not been effectively handled. The proposed system aim to handle the problem of search space optimization by classifying concepts into concept types and determining concept importance taking into account the synset based concept impor-

tance other than the inter ontology concept importance. This approach yields 15% increase in the number of significant concepts retrieved when compared to the existing system. The usage of the best similarity matches further increases the accuracy of the results obtained and thus enhancing the performance of the proposed system. The problem of uncertainty is to be handled in future.

REFERENCES

1. T R Gruber. A Translation Approach to Portable Ontology Specifications, in *Knowledge Acquisition*, 5(2):199–220, 1993.
2. Gang Wu, Juanzi Li, Ling Feng and Kehong Wang. Identifying Potentially Important Concepts and Relations in an Ontology, In *ISWC 2008, LNCS 5318*, pages 33–49, 2008.
3. T H Duong, G S Jo, J J Jung, N T Nguyen. Complexity Analysis of Ontology Integration Methodologies:A Comparative Study, In *Journal of Universe Computer Science*, 15(4):877–897, 2009.
4. Trong Hai Duong and Geun Sik Jo. Enhancing Performance and Accuracy of Ontology Integration by Propagating Priorly Matchable Concepts, In *Neurocomputing*, 88:3–12, 2012.
5. S Jan, I A Shah, I Khan, F Khan and M Usman. Similarity Measures and their Aggregation in Ontology Matching, In *International Journal of Computer Science and Telecommunications*, 3(5), 2012.
6. Lihua Zhao and Ryutaro Ichise. Aggregation of Similarity Measures in Ontology Matching, in *ISWC*, 2004.
7. Quang-Vinh Tran, Ryutaro Ichise and Bao-Quoc Ho. Cluster-based Similarity Aggregation for Ontology Matching, *6th ISWC Workshop on Ontology Matching*, 2011.
8. <http://jena.sourceforge.net/>.



Arthi J received her MTech in Distributed Computing Systems from Department of Computer Science and Engineering, Pondicherry Engineering College, Puducherry and B.Tech in Information Technology from Christ College of Engineering Technology, Puducherry. She started her career

as Program Analyst and worked as an Assistant Professor at National Institute of Technology Puducherry, Karaikal. Her research interests include Ontology Matching, Cloud Computing, Big data and Image Processing.



K Saruladha is an Assistant Professor in Department of Computer Science and Engineering, Pondicherry Engineering College, Puducherry. She completed her Under Graduation at Pondicherry Engineering College, Post Graduation and Ph.D in Computer Science and Engineering at Pondicherry University. She has 25 years of teaching experience and 10 years of research experience. Her areas of interests include Ontology Matching, Cloud Computing and Sentiment Mining.