

Information Retrieval based on Semantic Matching Approach in Web Service Discovery

R.Kanthavel, PhD.

Professor
Dept. of ECE
Velammal Engg. College,
Chennai

K.Maheswari

PG Student
Dept. of CSE
Govt. College of Engg.
Tirunelveli

N.Padmanabhan

PG Student
Dept. of IT
Anna University Coimbatore

ABSTRACT

Information retrieval from web has been a tough task due to an increased amount of contents dynamically day by day. In that current information retrieval results are only based on syntactic description for web service which may not provide an accurate solution for all service requests. On the other hand semantic based information retrieval has been the better option for precise and accurate results for queries. This paper proposes a semantic based hybrid matching approach that comprises different kinds of matching strategies combined together with an aim of flexible and efficient service retrieval in terms of accuracy. Using semantic based web ontology language service (OWL-S) conversion of non semantic description into corresponding semantic description is proposed. We also extend prototype to support save-off pre-processed data. The proposed work showed an attractive result after experiments by considering the functions namely recall and precision.

Keywords

Information Retrieval, Web Service, OWL-S, Semantic, Save Off Pre-processed data.

1. INTRODUCTION

Because of the extraordinary growing popularity of Web services, the Information Retrieval is the preferable way to access content in the global web data. Currently web information retrieval system approach has been fulfilled by a keyword based search of UDDI (Universal Description Discovery and Integration). But the opinion from many web researchers conclude that the required efforts to forward in web information retrieval technology could be achieved by adding semantics to web contents and exploiting them during processing of the web search queries[2]. Semantic information retrieval mechanisms are introduced for providing accurate results on service requests. By taking the advantages of metadata in Semantic information retrieval technique results an attractive service discovery. Semantic web services have the semantic tagged description through WSDL (Web Services Description Language)[1]. Beyond, several limitations of semantic web search technique ie, difficulties to expect all web services to have semantic tagged descriptions and the service requestor may not aware of all the terms related to the service request[11].

To overcome the limitations of existing approaches, an integrated and hybrid matching approach needs to be developed for addressing the two major issues related to improving effectiveness of information retrieval[13]: First, Semantic based hybrid matching approach have different kinds of matching strategies that are combined together to provide an accurate, flexible and efficient service retrieval.

The keyword matching algorithm can quickly return the results but with a limitation of understanding user's goals [7]. On the other hand, a semantic matching algorithm exploits information retrieval metrics to measure the degree of match between user request and service publishing [4]. The proposed hybrid matching algorithm uses semantic metadata along with flexibility of keyword matching algorithm. Moreover, OWL-S is an upper ontology used to describe the semantics of the OWL and is grounded in WSDL[14]. The OWL-S ontology comprises three main components: the service profile for advertising and discovering services; the process model, which gives the detailed description of service's operation; and the grounding, which provides details on about the interoperability with a service, via messages. Save-off Pre-processed data a principle which is particularly useful when service repositories grow extremely large in size and Bipartite graph is used to organize the pre-processed data[6].

The rest of the paper is organized as follows: Chapter 2 presents the preliminary works and Chapter 3 renders a detailed discussion of proposed work. The implementation parts and evaluations are presented in chapter 4. Finally, conclusion and future work are depicted in chapter 5.

2. PRELIMINARY WORKS

A number of related research works that aimed on information retrieval and ontology for identifying suitable methods to retrieve web service in order to meet out a better classification, and matching performance. Semantic based automated service discovery to achieve functional level service categorization based on an ontology framework [15]. The parts of the initial WSDL set and its corresponding description and associated parameters are formed by the WSDL file forms. The WSDL document pre-processing includes the extraction of the associated operation parameters. Service categorization and service request enhancement are added as separate layers on top of the UDDI, addition of these layers affect the performance of the discovery process in increasing the time delays[8].

Using bipartite graph algorithm, degree of similarity between web services can be calculated [12]. Based on the user request, a list of candidate services will be provided to the client depending on the degree of matching with the search query from higher to lower order. The matching algorithm is based on singular value decomposition in linear algebra, which reveals semantic relationship among web services and to describe model in UDDI [5]. In addition, the weight is calculated based on the inverse document frequency (IDF) which is a measure of a word's importance.

The OWL-S is used for the description of web service to increase the ability of computer systems to find eligible services autonomously [10]. Owl-s is used to find matches

between requirements and advertisements according to their semantic description. Semantic web service could provide more powerful web service development tools that can enable, among other things, including automated simulation, verification of web service properties, consistency checking and debugging [3]. DAML-S is a language for describing classes and their interrelationship. The limitation of DAML-S is default grouping approach does not provide conjunction with semantically richer descriptions without OIL. In order to address the limitations of existing approaches, a need of an integrated approach to be developed for addressing the major issues related to effectiveness of information retrieval by combining hybrid qualities for the approach to be generic and should be tied to a specific description language[9].

3. PROPOSED WORK

The objective of the proposed method is to address the semantics and descriptions of the web service as a whole which are not addressed by the existing methods. In this proposed approach, we extend the web service ontology framework to support an external database and implementing the matching algorithm to produce the accurate results for web services.

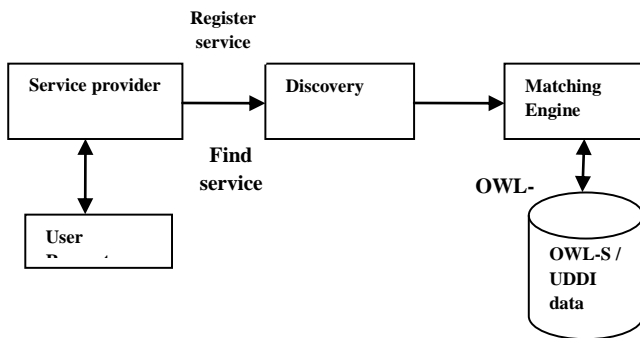


Figure 1: Architecture of semantic based information retrieval system

Figure 1 shows the proposed architecture of semantic based information retrieval system. In that, Matching engine contains the phenomenon as keyword matching, semantic and hybrid matching algorithm. Keyword matching algorithm can quickly return the results with a limitation of understanding user's goals. Semantic matching algorithm provide a way for matching on semantic level. Relations are computed between labels at nodes. The proposed Hybrid matching algorithm combines the flexibility of keyword-based retrieval and owing to the ability to query and reason on metadata typical of semantic search systems, the accuracy could be improved.

3.1 Semantic Services In Owl-S

OWL-S is an owl-based Web service ontology to supply a core set of mark-up language constructs for describing the properties and capabilities of Web services in unmistakable form. The overall ontology consists of three main components namely the service profile for advertising and discovering services, the process model to impart a detailed description of a service's operation and the grounding to provide details for interoperability.

The process elements of the document parse tree such as textDescription, and hasInput, are considered as text category with content in each owl-s service profile. Each word or concept in the content is canonically unfolded in the

underlying ontology; the resulting concept expression contains primitive components from the shared minimal basic vocabulary only. The OWL-S service retrieval metrics can be defined as follows [9].

$$OWL-S_{measure} = \frac{\overrightarrow{Req} \cdot \overrightarrow{Ser}}{\prod_{Req} \prod_2^2 \cdot \prod_{Ser} \prod_2^2} \quad (1)$$

The Weighting scheme and the unfolded concept expressions of request Req and service Ser are represented as n-dimensional weighted index term vectors namely Req and Ser respectively.

3.2 Semantic Based Hybrid Approach

The logic-based reasoning and content-based information retrieval techniques for owl-s service profile input and output matching are exploited by Semantic based hybrid matching using owl-s with the five variants for five different text similarity metrics. Exact, Plug-in and Subsumes filters are logic based only whereas the Subsumed by and nearest are neighbor hybrid due to the required additional computation of syntactic similarity values. The proposed Hybrid matching algorithm procedures are,

- Classifies the service user request.
- Determines the degree of semantic relation between pairs of service request and concept.
- Attached to each concept in the concept hierarchy. These lists of service identifiers are used to compute the set of relevant services.
- Pair wisely determines the degree of logical match.
- keyword similarity between the conjunctive input and output concept expression built by unfolding each of the query and service input concept in the local ontology.

4. EXPERIMENTAL WORK

The experimental results are presented in this section to analyse the hybrid matching approach in terms of recall and query response time. Precision and Recall parameters have been computed for the first 10 and 75 documents returned by each service request. The standard Precision and Recall measures are measured through the given equations represented in 2 and 3 respectively.

$$\text{Precision} = \frac{\text{CorrectSystemAnswers}}{\text{SystemAnswers}} \quad (2)$$

$$\text{Recall} = \frac{\text{Correct system Answers}}{\text{Expected Answers}} \quad (3)$$

Hybrid matching is not omitting relevant documents identified by the other ontology when intersecting the two sets. Table 1 show comparative evaluation of keyword, semantic and hybrid matching algorithm on 10 user requests and it is clearly found that an improvement in matching accuracy comparatively.

Table 1: Comparative Analysis of Matching Algorithm

User Request	Possible Correct Answer	Keyword Matching Algorithm [5]	Semantic Matching Algorithm [5]	Hybrid Matching Algorithm *
1	84	16	20	20
2	22	16	0	16
3	25	1	11	11
4	63	19	19	19
5	27	9	12	12
6	5	4	0	4
7	7	6	0	6
8	1	1	0	1
9	5	3	0	5
10	83	12	0	20

* Proposed hybrid matching algorithm

Figure2, shows the comparative analysis of the proposed Hybrid technique with existing matching algorithms as keyword and semantic matching algorithm and it is evaluated the proposed one overtakes the rest in Recall and Precision.

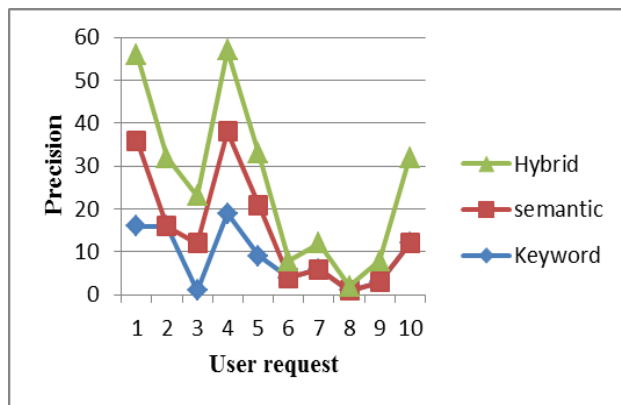


Figure 2. Result of evaluation of Hybrid matching algorithm with Existing algorithms

5. CONCLUSION AND FUTURE WORK

This paper proposed the hybrid matching approach with the objectives to improve the effectiveness of information retrieval. Since Hybrid semantic Web service matching utilizes logic based reasoning and Information retrieval techniques for semantic Web services in owl-s, it has the capacity to filter the matching more precisely achieve effective and accurate result. Moreover, Semantic services in owl-s are used to achieve the functional level service based on an ontology framework. The multiple services discovery along with a creation of an interactive, intelligent service composer that are to be semantically guided to locate the target service components will be the future work.

6. REFERENCES

[1] Cohen, W.; Ravikumar, P.; and Fienberg, S. 2003. A comparison of string distance metrics for name-matching tasks, In Proceedings of IJCAI-03 Workshop on Information Integration on the Web(IJWeb-03). DBLP at <http://dblp.uni-trier.de>.

[2] Heb.A and Kushmerick.N.2003, Automatically Attaching Semantic Metadata to Web Services, IJCAI Workshop Information Integration on the Web, 1-7

[3] Martin.D and McIlraith.S.2003,Bringing Semantics to Web Services, IEEE Intelligent Systems, Vol. 18, No. 1, 90-93

[4] Stojanovic.N, Stojanovic.L, and Studer.R.2003, An Approach for the Ranking of Query Results in the Semantic Web," International Semantic Web Conference (ISWC '03), 500-516.

[5] Hou.J, Sajjanhar.A and Zhang.Y.2004, Algorithm for Web Services Matching, Asia-Pacific Web Conference (APWeb), pp. 665-670.

[6] Anyanwu.K, Maduko.A, and Sheth.A. 2005, SemRank: Ranking Complex Relationship Search Results on the Semantic Web, Proc.14th Interenational Conference on World Wide Web (WWW '05), 117-127.

[7] Papakonstantinou.Y and Xu.Y.2005, Efficient keyword search for smallest lcas in XML databases. In *SIGMOD*, ACM,527-538.

[8] Bhuvan Bamba and Sougata Mukherjea, 2005, Utilizing Resource Importance for Ranking Semantic Web Query Results, Semantic Web and Databases Lecture Notes in Computer Science Volume 3372, 185-198.

[9] Adam N, Bornhoevd.C, Paliwal.A.V, and Xiong.H.2006, Web Service Discovery via Semantic Association Ranking and Hyperclique Pattern Discovery, Proc. IEEE/WIC/ACM Interenational Conference on Web Intelligence, 649-652.

[10] Kaufer and Klusch.M.2006, Wsmo-mx: A logic programming based hybrid service matchmaker, IEEE European Conference on Web Services , IEEE CS Press,161-170.

[11] Cimiano.P, Rudolph.R, Studer.R, and Tran.T.2007, Ontology-based Interpretation of Keywords for Semantic Search, Proceedings of the International Semantic Web Conference and the Asian Semantic Web Conference, 1-14.

[12] Klusch.M.2008, Semantic web service coordination., in: In Schumacher.M, Helin.H (Eds.): CASCOM – Intelligent Service Coordination in the Semantic Web. Chapter 4. Birkh" auser Verlag, Springer, 2008.

[13] Klusch.M and Xing.Z.2008, Deployed semantic services for the common user of the web: A reality check,IEEE International Conference on Semantic Computing (ICSC), IEEE Press, 347-353.

[14] Evangelos and Sakkopoulos. 2009, Semantic technologies for mobile Web and personalized ranking of mobile Web search results, Lecturer Notes on Metadata and Semantics,299-308.

[15] Aabhas V.Paliwal, Basit Shafiq, Jaideep vaidya, Hui Xiong and Nabil Adam. 2012, Semantics-Based Automated Service Discovery, IEEE Transaction on Service Computing, Vol 5, No.2, 260-275.