ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 9(12), DOI: 10.17485/ijst/2016/v9i12/85910, March 2016* ISSN (Online) : 0974-5645

Semantic cluster based Search in UDDI for Health Care Domain

K. Tamilarasi[1*] and M. Ramakrishnan[2]

[1]Department of Computer Science and Engineering, Sathyabama University, Chennai - 600119, Tamil Nadu, India; thamizhanna@gmail.com
[2]Department of Computer Application, Chairperson - School of Information Technology, Madurai Kamaraj University, Madurai - 625002, Tamil Nadu, India; ramkrishod@gmail.com

Abstract

Objective: To improve relevancy and personalization of service discovery process of search enginesby introducing intelligence in representation of web service called semantic reputation. **Methodology:** The discovery and distribution of healthcare biological data through web services is being an important characteristic of today's search engine and this process should be efficiently handled. This paper proposes a semantic cluster based recommendation rules approach called semantic reputation for retrieving tagged biological web services. Textual description of biological web services is retrieved by using the semantic web services clustering approach and Ranking is done by recommendation rule based semantic matching. **Findings:** Retrieving meaningful information is difficult in today's available search engines. Semantic intelligent representation plays a major role in efficiently retrieving meaningful information intelligently. The proposed approach is implemented and compared with the other similar approaches and obtained the recall value of 97.6% and precision value of 98.5% which is better than the existing approaches comparatively by 10% to 19%. **Application/ Improvement:** Our system supports service requestors to retrieve the services relevantly without deviating them to some other fraudulent web sites which proves to be time consuming and trustworthy. Extended work of this paper is to implement these approaches in extending UDDI for efficient web service discovery.

Keywords: Cluster, Healthcare, Recommendation, Semantic, UDDI, Web services

1. Introduction

With the emergence of web services and advancements in the web service technologies, empowering users and multi core organizations have tremendous opportunities in enormous application domains including intelligence information gathering, electronic commerce, digital government, information analysis, travel, heath care, etc. Service oriented architecture presents an efficient solution for integrating autonomous, distributed, heterogeneous and interoperable information sources by employing the industrial standards through Web services platform elements[1].

To enhance knowledge sharing and collaboration between enterprises there is a need to integrate information sources into Web services and its emergence as a model for integrating heterogeneous web information leads to potential interaction and interoperability. Managing and maintaining an application protocol based on semantic level is forcefully needed rather than syntactic level as sequencing information and data structures are enhanced with semantic information[2]. The key element is ontology which enables the transformation from a purely syntactic to semantic interoperability.

The administrative tasks of heath care is practiced using electronic methods in long run and in the recent years health care has been adopting latest technologies for the betterment delivery of its services and to ensure quality of service. Health care domain must ensure the higher standards of effectiveness and accuracy while transforming into newer technologies[3].Health care web services include several life science domains: Bioinformatics, systems biology, computational biology, and clinical genomics and covers heterogeneous information categories like

* *Author for correspondence*

literature data mining, classification, disease, patient, pathway, interaction, experiment and annotation[4].

One important issue lies in finding a suitable web service by an end new user and it is mostly unclear that which web service and which operation satisfies the intended task with the expected constraints. There are two main strategies as the solution to the problem. The first method is building registries that gather a huge number of web services for a health care domain and the second method is adapting semantics for the automatic discovery of the required web service and its operations.

Web wide registry of web services is provided by UDDI but its search capability is limited to syntax which results in coarse nature of service discovery and data model description using XML schemas[5]. In UDDI, functionality of a web service can be described by schemas. For example, a Healthcare service can use the category bag as National Provider Identifier (NPI) as 10 digit number such as 345128901 to describe its functionality.

In the Interaction analysis and modelling scenario, persons or organizations play five main social roles: 1) the patient, 2) healthcare professionals in local healthcare institutions who provide on-site medical care for the patient, 3) other healthcare professionals who provide additional information or a second opinion about the patient, 4) external emergency assistance organizations, and 5) administrators of local healthcare institutions. Several possible interactions have been identified between this scenario's participants. We can easily extend our ontology to include new interaction types and roles, making it more complete and reusable[6].

Similar XML descriptions represent different context of web services and lack of semantic representation paves the barrier to efficient discovery in UDDI[7]. This problem is addressed by semantic web initiative of creating XML languages such as OWL and RDF where the content of the tags are explicitly specified. This paper proposes the discovery mechanism based on semantic cluster matching which returns the most relevant web services by performing logic inference to match with the requested capabilities. Proposed extended UDDI with text tagging and recommendation combines the features which results in minimum search time and relevant discovery[8].

Data providers are independent in the information space and they have credentials of their schematic representations and functionalities. For addressing this problem different approaches have been presented in the literature. Clustering based approaches using different methods have been proposed by many authors[9-11]. What set us apart from the proposed approaches is that we aim at extracting the features of the services using query based system. Identification and classification of web services is analyzed by means of related rules generation from the query and therelevant services are retrieved with minimum computation and access time.Healthcare case study is implemented and analyzed to illustrate the proposed work.

This paper proposes an integrated architecture for semantic cluster and recommendation rules formation for text tagged services and rest of the paper is organized as follows. Detailed description of methodology and modules is presented in section 2, Performance evaluation of the proposed architecture is enumerated in section 3. Conclusion and Future work is furnished in section 4.

2. Proposed Methodology

The proposed system retrieves the health care web services using two phases such as client phase and server phase. In the server phase, the semantic based cluster is formed in the web server after the service provider registers their services in the UDDI registry[12]. In the client phase, three modules are included to retrieve the web services such as i) Generating Recommendation Rules ii) Applying Recommendation Rules for searching process iii) Applying Recommendation rules for replying process.

To combine tagging, recommendation and UDDI, we need to embed a semantic profile description in a UDDI data structure and augmenting these components as extensions to UDDI has been done for processing semantic profile information[13]. The architecture of the extended semantic UDDI registry is shown in Figure 1. The recommendation component is tightly coupled with UDDI and relies on registry for all its operations to be successfully completed.

2.1 ServerService

The incoming request from the requestor has to be pre-processed by parsing, removal of unwanted words, function removal and stemming. Then the features or keywords are extracted from the pre-processed data. The keywords are grouped using the semantic based cluster[14]. The similarities between the keywords are identified using vector space model by using the following equation (1).

$$sim(q,s) = \frac{|q * s|}{||q||^2 * ||s||^2} \quad (1)$$

where q, s is represented as the different keywords for which the measurement of similarity is to made.

Algorithm for Cluster Formation

Step 1: Get the Biological Services
Step 2: Preprocess the service, or biological data
Step 3: Extract the keywords from the preprocessed data
Step 4: Find the similarity between the keywords
Step 5: Check the similarity with the threshold value (max)
Step 6: Form the cluster

2.2 Client Service

The healthcare data related services are efficiently retrieved in minimum time by generating the recommendation rules, applying them during searching and replying process[15]. The rules are generated based on the content i.e. description of the service, items and user preferences and past searching history[16]. By keyword importance and the user feedback, the weight factors are determined through which the rules for recommendation are generated by using the rule induction method[17].

The method proposed here is decision tree which forms the rule in the form of nodes such as root, end and chance nodes. For a particular query, the rules are contained in the root node, feedback changes are done in the chance node and weight factors based on rules are followed in end node[18]. Finally the weight is assigned as a rule used during the request or response. The matchmaking is done based on the rules during retrieval from several databases, repositories and registries and the transformed to the replying process[19]. The rules and related services are arranged in the sorting order for top k rules to be identified. Keyword weight or priority of the particular keyword is used for identification of top k rules which is used in retrieval of quality web services[20].

2.3 Implementation

The scalability and performance of a UDDI registry is done by preliminary evaluation by comparing the results of retrieval with and without component extension of an UDDI. An open source UDDI registry, jUDDI is extended with recommendation and tagging component[21]. WordNet 2.1 API is used for retrieving the synonyms of the keywords used by service consumers for retrieving the service candidates. Processing time of a service publishing is measured by calculating the interval between the request arrival and results delivery and the network latency time is eliminated[22].

3. Performance Evaluation

This semantic-service discovery mechanism considers relevant parts of the organizational context in which e-health services are used, to improve a service-discovery system's usability in medical emergencies. For evaluating the performance of the proposed architecture, the publishing time, the loading time and the querying time has to be taken into consideration[20]. Then the efficiency of the proposed system is measured with the help of the precision value and the recall value. From the performance results we can achieve the efficiency of the web service retrieval resulting in minimum search time and relevant results.Results of Precision and Recall values of different IR techniques have been depicted in Table2 and Figure 2.

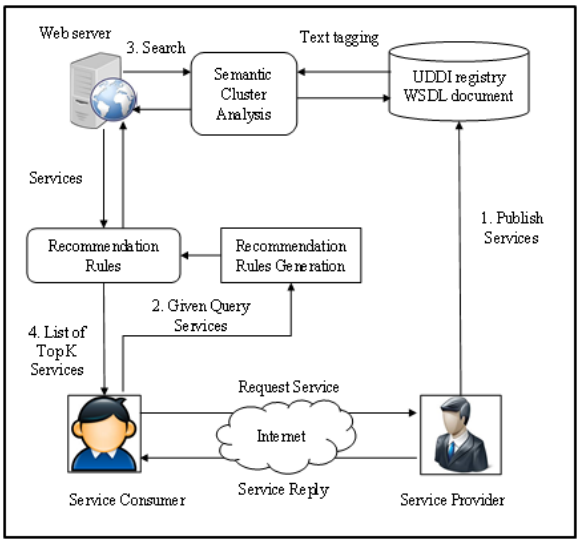

Figure 1. Proposed System Architecture.

Table 1. Schema of Healthcare Information Source

Patient (IdP, Name, Age, Tel)
Doctor (IdD, Name, Speciality, Hospital, IdS)
Hospital (IdH, Name, Localization)
Doctor_Hospital (IdD, IdH)
Patient_Hospital (IdP, IdH, IdD)
Diagnostic (IdP, IdD, DateT, Result)
Operation (IdP, IdD, DateO, Result)
Service (IdS, Speciality)

Table 2. Results of different IR Techniques

WSR Techniques	Precision	Recall
Vector Space Model	0.75	0.786
Latent Semantic Indexing	0.83	0.8463
Latent Dirichlet Allocation	0.861	0.8763
Semantic based Clustering and Recommended Rule	0.985	0.976

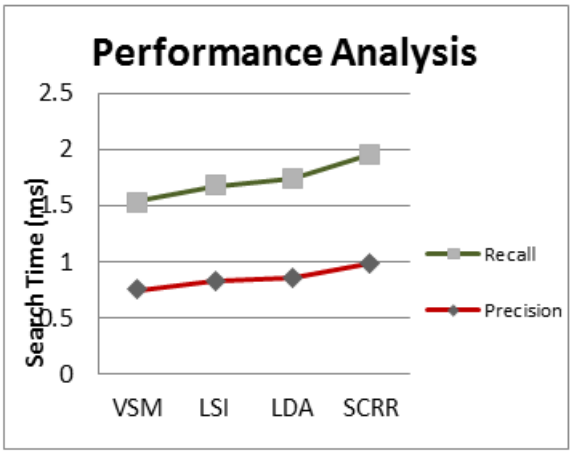

Figure 2. Comparison of Web service retrieval techniques.

4. Conclusions and Future Work

This paper elaborated the need of web service discovery and extending UDDI. UDDI data model has to be extended for particular domain and such architecture is very important for medical health care domains as it involves live and death issues in that no risk can be taken. Proposed extended UDDI performs semantic based cluster search based on recommendation rules generated and applied during search and retrieval process which leads to personalization solutions related to biological web services.

The limitation is thatsemantic matching lacks service parameters and categories. We are trying to integrate the type of service to the tagging so that a service client can express the service type literally rather than technically through preconditions, effects, input and output. Scalability of the proposed technique has to be evaluated with more number of web services.

5. Reference

1. Maheswari S, Karpagam GR. Comparative Analysis of Semantic Web Service Selection Methods. Indian Journal of Science and Technology. 2015 Feb; 8(S3):27–30. Doi: 10.17485/ijst/2015/v8i1/60499.
2. Karthikeyan K, Karthikeyani V. Ontology Based Concept Hierarchy Extraction of Web Data. Indian Journal of Science and Technology. 2015 Mar; 8(6):536–47.Doi:10.17485/ijst/2015/v8i6/61070.
3. Suchithra M, Ramakrishnan M. A Survey on Different Web Service Discovery Techniques. Indian Journal of Science and Technology. 2015 Jul; 8(15):1–5. Doi:10.17485/ijst/2015/v8i15/70773
4. Abarna R, Pradeepa S. A Hybrid Approach for Extracting Web Information. Indian Journal of Science and Technology. 2015 Aug; 8(17):1–6. Doi:10.17485/ijst/2015/v8i17/61595.
5. Mustafa AS, Kumaraswamy YS. Performance Evaluation of Web-services Classification. Indian Journal of Science and Technology. 2014 Jan; 7(10):1674–81. Doi: 10.17485/ijst/2014/v7i10/50859.
6. Chesson P. Species Competition and Predation. Encyclopedia of Sustainability Science and Technology. 2015; 1–28.
7. Souza C, Canhos G. Global Biodiversity Informatics: Setting the scene for a "New World" of ecological modeling. Biodiversity Informatics. 2004; 1:1–13.
8. Kim SM, RosuMC. A Survey of Public Web Services. Available from:http://www.ambuehler.ethz.ch/CDstore/www2004/docs/2p312.pdf.Date Accessed: 22/07/2015.
9. Priyadharshini G, Gunasri R, SaravanaBalaji B. A survey on semantic web service discovery methods. International Journal of Computer Applications (0975 – 8887).2013 Nov; 82(11):8–11.
10. PatilN, Gopal A. Ranking Web-services based on QoS for Best-fit Search. International Journal of Computer Science and Communication. 2010 Jul-Dec; 1(2):345–49.
11. Hasan MH, Jaafar J, Hassan MF. Fuzzy-based clustering of web services' quality of service: A review. Journal of Communications. 2014 Jan; 9(1):81–90.
12. RostamiNH, KheirkhahE, Jalali M. An optimized semantic web service composition method based on clustering and ant colony algorithm.2014; 8.
13. Skoutas D,Sacharidis D, Simitsis A, Sellis T. Ranking and clustering web services using multicriteria dominance relationships. IEEE Transactions on Services Computing. 2010 Jul-Sep; 3(3):163–77.
14. Al-MasriE, Mahmoud QH.Discovering Web Services in Search Engines. University of Guelph. 2008; 12(3):74–7.
15. Wang X, Vitvar T, Kerrigan M, Toma L. A QoS-aware selection model for semantic web services. 2006; 4294:390–401.
16. Li H, Du X, Tian X. Towards Semantic Web Services Discovery with QoS Support using Specific Ontologies. 2007 Oct 29-31; 358–61.

17. Al-Masri E, Mahmoud QH.Investigating Web Services on the World Wide Web. 2008.
18. Gang YE, Chanle WU, Jun YE, Cheng S, Chanle WU.A QoS-aware Model for Web Services Discovery. 2009 Mar 7-8; 3:740–44.
19. Al-MasriE, Mahmoud QH. WSCE: A Crawler Engine for Large-Scale Discovery of Web Services. 2007 Jul 9-13; 1104–11.
20. Tamilarasi K, Ramakrishnan M. Design and Development of an Enhanced UDDI for Efficient Discovery of Web Services. Advances in Communication, Network, and Computing Lecture Notes of the Institute for Computer Sciences. Social Informatics and Telecommunications Engineering. 2012; 108:109–14.
21. Tamilarasi K, Ramakrishnan M, Design of an intelligent search engine-based UDDI for web service discovery.International Conference on Recent Trends in Information Technology.2012;520–25.
22. Juric MB, Sasa A, Brumen B, Rozman I. WSDL and UDDI extensions for version support in web services. J of Systems and Software. 2009 Aug; 82(8):1326–43.