Ontology-based Semantic Search Engine for Healthcare Services

Jotsna Molly Rajan
P.G. Scholar
PSN College of Engineering and Technology
Tirunelveli, Tamil Nadu,India
jotsna@gmx.com

M. Deepa Lakshmi
Associate Professor
PSN College of Engineering and Technology
Tirunelveli, Tamil Nadu,India
deepasuresh12@gmail.com

Abstract— With the development of Web Services, the retrieval of relevant services has become a challenge. The keyword-based discovery mechanism using UDDI and WSDL is insufficient due to the retrieval of a large amount of irrelevant information. Also, keywords are insufficient in expressing semantic concepts since a single concept can be referred using syntactically different terms. Hence, service capabilities need to be manually analyzed, which lead to the development of the Semantic Web for automatic service discovery and retrieval of relevant services and resources. This work proposes the incorporation of Semantic matching methodology in Semantic Web for improving the efficiency and accuracy of the discovery mechanism.

Keywords-Ontology, Semantic Matching, Web Service Discovery, Semantic Matchmaking

I. INTRODUCTION

Due to the rapid and dynamic growth of the Internet, it is quite challenging to search and retrieve information or resources accurately and efficiently. Current search mechanisms make use of syntax-based mechanisms in which the search and retrieval is based on the keywords entered by the user. The keyword-based search leads to retrieval of huge amount of information that is inaccurate and irrelevant along with relevant and accurate information. Hence, data analysis must be done manually. The semantic-based retrieval system can retrieve information based on the semantics or meaning of concept given by the user.

Search engines are Web services that can be used to search and retrieve information based on some criteria specified by the user. The information retrieved will satisfy the criteria, mainly based on keywords.

Web Services are application components that can be used by other services or users to communicate over the Internet. They are based on open standards like Extensible Markup Language (XML), Web Services Description Language (WSDL), Universal Description, Discovery and Integration (UDDI) and are hence widely accepted in the industry.

WSDL [1] provides description of the Web services, but does not support any transportation mechanism; which is in turn given by SOAP; XML is used to describe the data in a structured way using user defined tags, UDDI [5] acts as a registry for publishing, searching and retrieving Web services provided by a service producer.

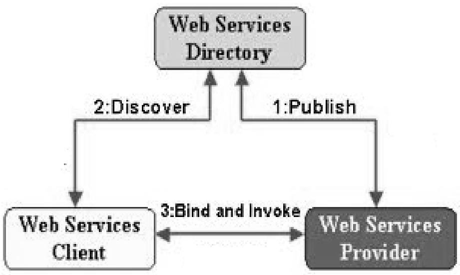

Figure 1: Web Service Model

Web service Discovery [11] is the process of finding Web services with a given capability. A service provider creates services that provide certain functional capabilities and publishes them in a registry. The registry is used to store the advertisements of the functional capabilities, and to perform a match between the request and the advertisements. The service provider unaware of the services available requests a service and checks the registry to find services satisfying his requirement. A matching process is required that compares the advertisements with the queries for the discovery of the Web services. If such a service is available in the registry, the service provider accesses and invokes the service.

The Web service components like WSDL and UDDI provides automatic service discovery but the search mechanism is keyword-based, hence inefficient. The efficiency of the matchmaking can be improved by incorporating semantics to the Web services. The concept of semantics has been introduced with OWL [3] which provides domain ontology and OWL-S [4] which provides Service annotations. The performance of service discovery in Semantic Web services can be increased by implementing the Semantic matchmaking methodology.

Semantics is the study of meanings of concepts, terms or words. Functional capabilities of a service can be expressed using input, output, preconditions and effects. Syntactically different terms can be used to represent the same concept using ontology. Ontology defines the entities in a domain, their relationships and properties. Matchmaking methods can be implemented to find the degree of match between an advertisement provided by the service provider and a request provided by the service consumer. When trying to implement matchmaking based on semantics, different degrees of matching can be achieved.

Let A and Q denote concept of advertisement and concept of query respectively. Then the four relationships [2] between A and Q are:

1. Exact (A, Q): A and Q represents the same concepts, i.e., both are equivalent.
2. Subsumes (A, Q): The concept Q is subsumed by the concept A, meaning that A denotes a more general concept than Q.
3. Plugin (A, Q): The concept A is subsumed by the concept Q. This relationship is also called inverted subsumption.
4. Fail (A, Q): The concepts A and Q do not match each other in anyway.

Using Semantic Web services, corresponding ontology, and an efficient matchmaking methodology, the accuracy of the search mechanisms can be vastly improved; knowledge overhead can be reduced and Web service discovery, without human intervention is possible.

II. BASIC APPROACH

Klusch [13] classifies the semantic matchmaking approaches into Logic-based, Non-logic based and Hybrid semantic matching, depending on the nature of reasoning used.

In logic-based matching, a semantic match between concepts is found out through logical reasoning. Such matching of services is written using languages like OWL-S, WSML or SAWSDL. Ontology-based and concept-based systems are some examples of such an approach. In Ontology-based approach, ontology of the domain is described using OWL or RDF. Concept- based approach deals with the semantics described in the Web services descriptions. In non logic-based matching, the degree of semantic matching is not done by means of inference. Such semantic matching can be done by measuring the semantic distance between the descriptions. Hybrid matching makes use of both logic and non-logic based matching methodologies, when the usage of either logic or non-logic matching will not succeed if applied individually.

Unlike existing search engines that make use of syntax or keyword-based search, the Semantic search engine can retrieve accurate and relevant information based on Ontology. The search engine makes use of semantic similarity of inputs, outputs, preconditions and effects between concepts, to retrieve relevant information.

Figure 2(a) provides the working of an existing system, in which the service provider advertises the services in a service registry, from which the service consumer searches the services. A matching engine matches the keywords in the query to return a set of results to the service consumer.

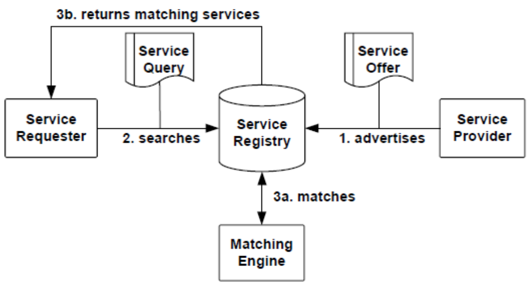

Figure 2(a) Existing system

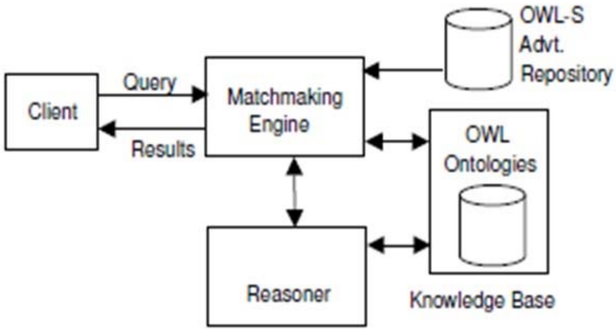

Figure 2(b) Framework of the semantic web service matchmaking system

Service providers publish their services to a registry storing the service information through a user interface. Advertisements are semantically annotated using OWL-S and loaded into a repository. Service providers can provide domain ontologies using OWL, which are then parsed and stored in Ontology repository. A service consumer can enter a request in the interface. The concepts in the request are matched against that of the advertised service using a reasoner and semantic similarity is computed. The matching advertisements are then displayed to the user.

III. IMPLEMENTATION

The main aim of this paper is to develop an ontology-based search engine that makes use of semantic matching methodology and performs accurate retrieval of Web services in Vaccination of Healthcare domain. Figure 3 shows the System architecture. The Semantic annotations provided by OWL-S, the description of the Web and the Vaccine Ontology are incorporated into the UDDI registry. When a client makes a request, the request gets send to the registry from which the appropriate Web service gets selected and results displayed. Using Code-driven approach [6], the implementation can be divided into the following functional modules, namely, Service Description, Semantic Annotations and Semantic Matchmaking. The main aim of this paper is to develop an ontology-based search engine that makes use of semantic matching methodology and performs accurate retrieval of Web services in Vaccination of Healthcare domain.

Figure 3 shows the System architecture. The Semantic annotations provided by OWL-S, the description of the Web and the Vaccine Ontology are incorporated into the UDDI registry.

When a client makes a request, the request gets send to the registry from which the appropriate Web service gets selected and results displayed. Using Code-driven approach [6], the implementation can be divided into the following functional modules, namely, Service Description, Semantic Annotations and Semantic Matchmaking.

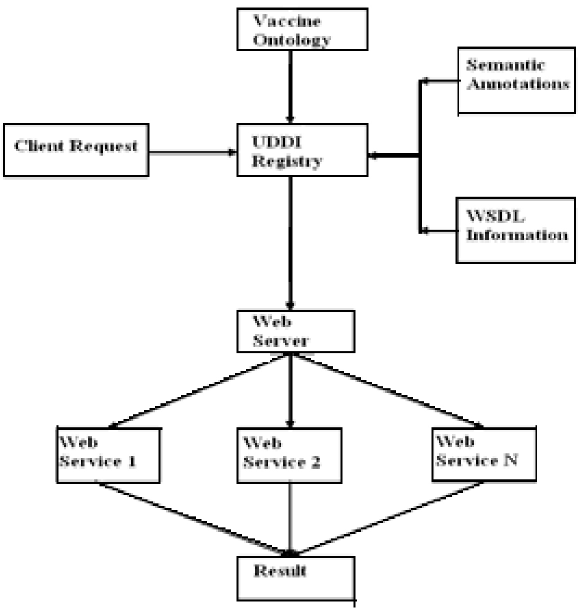

Figure 3 System Architecture

```
<?xml version='1.0' encoding='UTF-8'?>
<definitions .... targetNamespace="http://sch.pkg/" name="vaccineSchedule">
<types><xsd:schema><xsd:import namespace="http://sch.pkg/"
schemaLocation="http://localhost:8080/vaccineSchedule/vaccineSchedule?xsd=1" />
</xsd:schema></types>
<message name="vaccine_Schedule"><part name="parameters"
element="tns:vaccine_Schedule" />
</message><message name="vaccine_ScheduleResponse">
<part name="parameters" element="tns:vaccine_ScheduleResponse" /></message>
<portType name="vaccineSchedule"><operation name="vaccine_Schedule">
<input wsam:Action="http://sch.pkg/vaccineSchedule/vaccine_ScheduleRequest"
message="tns:vaccine_Schedule" />
<output wsam:Action="http://sch.pkg/vaccineSchedule/vaccine_ScheduleResponse"
 message="tns:vaccine_ScheduleResponse" />
</operation></portType>
<binding name="vaccineSchedulePortBinding" type="tns:vaccineSchedule">
<soap:binding transport="http://schemas.xmlsoap.org/soap/http" style="document" />
<operation name="vaccine_Schedule"><soap:operation soapAction="" />
<input><soap:body use="literal" /></input>
<output><soap:body use="literal" /></output>
</operation></binding>
<service name="vaccineSchedule"><port name="vaccineSchedulePort"
binding="tns:vaccineSchedulePortBinding">
<soap:address location="http://localhost:8080/vaccineSchedule/vaccineSchedule" />
</port></service></definitions>
```

Figure 4. A sample WSDL service interface document

A. Service Description

Service consumers are unaware of the various services provided by the Service producers. The consumers are provided with a way to access the available services through queries, by means of a user interface screen.

WSDLs (Figure 4) can be generated from the Web services developed in Netbeans IDE. Data that is to be retrieved can be stored within a database. The various Web services can be published in a UDDI registry. The registry also retrieves these published Web services, based on exact queries.

B. Semantic Annotations

The domain ontology described in this paper is the Vaccination Ontology, described using OWL, can be generated using the tool, Protégév3. Various subclasses like age of the patient, name of the vaccine, schedule of the vaccine, dosage, symptoms of the disease, and side-effects of the vaccine can be described along with their properties and relationship with each other.

Since the UDDI registry is syntax based, Semantic markups of the Web service must be provided. OWL-S provides Service annotations, which can be done with the help of OWL-S Editor, a plug-in facility available to both Eclipse IDE and Protégé.

OWL-S files can be classified into three, namely, Service Profile, Service Model and Service grounding. Service Profile describes the functionality of the services. Service Model describes the service as a process. Service Grounding described how to access the service. Service Model and Grounding together provided the means to invoke the service.

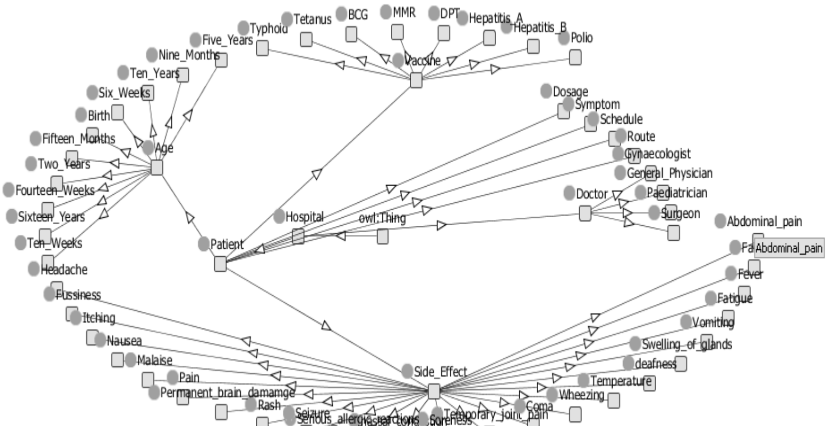

Figure 5. A sample of the vaccination ontology

```
<profile:Profile  rdf:ID="AddServiceName">
    <profile:serviceName>
      Add Service Name
    </profile:serviceName>
    <profile:textDescription>
      Add text Description
    </profile:textDescription>
    <profile:contactInformation>
      <profile:Actor rdf:ID="AddActorName">
      </profile:Actor>
    </profile:contactInformation>
    <profile:hasInput>
      <process:Input rdf:ID="DiseaseSymptom_vaccine_Schedule_disease_IN">
        <process:parameterType rdf:datatype="&xsd;#anyURI">
          &xsd;#string
        </process:parameterType>
      </process:Input>
    </profile:hasInput>
    <profile:hasOutput>
      <process:Output rdf:ID="DiseaseSymptom_vaccine_Schedule_vaccine_ScheduleReturn_OUT">
        <process:parameterType rdf:datatype="&xsd;#anyURI">
          &xsd;#string
        </process:parameterType>
      </process:Output>
    </profile:hasOutput>
  </profile:Profile>
```

Figure 6. A sample of the OWL-S profile generated

C. Semantic Matchmaking

An efficient matchmaking methodology is required for accurate retrieval of services. In [2], Paolucci et al, describes a greedy method that performs semantic matching on input and output terms of the concepts. The request from the service consumer is matched against all advertisements of services recorded in the registry. Whenever a match is found between the output concepts of query and advertisement, the match is recorded and the term is deleted from the list of query outputs called the candidate list. Hence, this approach depends on the order of the candidate list thereby producing false positives and false negatives in the results. False positives occur when queries and advertisements match, but returned results are irrelevant. False negatives occur when retrieval of relevant results cannot be done, even if such information is present. E.g., exact match occurs when the service name, input and output of a query is equivalent to the service name, input and output of an advertisement provided by the service producer.

In addition to checking semantic matching between input and output concepts of advertisements and queries, effects and preconditions can also be taken into consideration. Preconditions are the criteria that must be satisfied before the Web service can be invoked and effects are the changes that occur once the invocation takes place.

Ranking of results can be introduced along with the semantic matching. In addition to this, the ontology of the domain is also incorporated using OWL. The query can be given as OWL-S files or as keywords whose semantics is taken into consideration.

The basic rules for degree of match, introduced by Paolucci et al[2], is shown in Figure 7.

```
degreeOfMatch(outR,outA):
if outA=outR then return exact
if outR subclassOf outA then return exact
if outA subsumes outR then return plugIn
if outR subsumes outA then return subsumes
otherwise fail
```

Figure 7. Basic rules for the degree of match assignment

IV. CONCLUSION

Current trends in the growth of Internet and Web related services require dynamic retrieval and invocation methods. Better search capabilities are also rapidly emerging. But performance of keyword-based search engines cannot be termed efficient. In order to handle such problems, the Semantic technology can be incorporated using Ontology and an efficient matchmaking methodology.

Retrieval of relevant information can be done automatically, efficiently and accurately, thereby, reducing knowledge overhead and hence, manual data analysis. The number of searches and the time of search can also be vastly reduced.

The current search is implemented using Vaccine Ontology in the Healthcare domain. This can be extended to the entire Healthcare domain. Also, other domains can acquire the capabilities of the Web service discovery by using the same methodology, by using other Ontology.

REFERENCES

[1] WSDL 1.1, http://www.w3.org/TR/wsdl, 2001

[2] M. Paolucci, T. Kawamura, T. R. Payne, and K. Sycara, "Semantic matching of Web services capabilities," in *Proc. 1st Int. Semantic Web Conf.*,Sardinia, Italy, Jun. 2002, pp. 333–347

[3] OWL, http://www.w3.org/2004/OWL/, 2004.

[4] OWL-S 1.1, Release: Examples, http://www.daml.org/ services/ owl-s/ 1.1/ examples.html, 2004.

[5] UDDI, www.oasis-open.org/ committees/uddi-spec, 2005.

[6] Naveen Srinivasan, Massimo Paolucci, Katia Sycara, "CODE: A Development Environment for OWL-S Web services", Oct 2005.

[7] Umesh Bellur, Roshan Kulkarni, "Improved Matchmaking Algorithm for Semantic Web Services Based on Bipartite Graph Matching," in *IEEE International Conference on Web Services*, July. 2007, pp. 86-93.

[8] Umesh Bellur, Harin Vadodaria and Amit Gupta, "Semantic Matchmaking Algorithms", Greedy Algorithms, Witold Bednorz (Ed.), 2008

[9] Debajyoti Mukhopadhyay, Aritra Banik, Sreemoyee Mukherjee, Jhilik Bhattacharya, Young-Chon Kim, "A Domain Specific Ontology Based Semantic Web Search Engine," in Feb 2011.

[10] M. Cai, W. Y. Zhang, and K. Zhang, "ManuHub: A Semantic Web System for Ontology-Based Service Management in Distributed Manufacturing Environments", IEEE Transactions on Systems, man, and Cybernetics PART A: Systems and Humans, VOL. 41, No. 3, May 2011

[11] Vaneet Sharma, Mukesh Kumar, "Web Service Discovery: A Study of Existing Approaches", Int. J. on Recent Trends in Engineering & Technology, Vol. 05, No. 01, Mar 2011

[12] Yang Zhang, Fagui Liu, Nan Zhang, "Toward Fine Grained Matchmaking of Semantic Web Services Based on Concept Similarity," Journal of Information & Computational Science 8: 2 (2011) 377-384.

[13] Semantic Web Service Coordination, Matthias Klusch.

[14] www.oracle.com