International Journal of Scientific and Technical Advancements
IJSTA
ISSN: 2454-1532

A Survey on Classification of Different Techniques for Web Service Discovery

Anupama P. Nayak[1], E. Sruthi[2]
[1, 2]Computer Science & Engineering, VTU, Karnataka, India-574219
Email address: [1]anupamanayak92@gmail.com, [2]sruthiedayilum@gmail.com

Abstract— Web services allow application to communicate using standardized protocols with low cost. Finding most suitable web service from vast collection of web services is very crucial for successful execution of applications. As the number of web services increase rapidly and web services that provide the same functionality are developed , it is a major issue to consider not only the functional requirements but also the non functional requirement (NFR) during web service discovery process. Service discovery can be classified based on two criteria. One is syntax based or key word based service discovery and semantic service discovery. Another classification is one which considers the Quality of Service (QoS based service discovery). Various other approaches for discovering web services are also available. Having system for service discovery which can work automatically is also the concern of service discovery approaches. As these approaches are different, one solution may be better than another depending on requirements. Selecting a specific service discovery system is a hard task. The paper presents a study on various web service discovery approaches

Keywords— Web service; WSDL; QoS; semantics; ontology; wordnet; clustering.

I. INTRODUCTION

Web Services are modular, self-describing, and loosely coupled software applications that can be advertised, located, and used across the Internet using a set of standards such as SOAP, WSDL, and UDDI. Web services are an emerging Service Oriented Architecture technology. A web service is a software system identified by a URI, whose public interfaces and bindings are defined and described using WSDL. Web service is the interface designed to make the original isolated sites to communicate and share data with each other. It utilizes uniform protocols and data formats such as HTTP, XML, SOAP and WSDL. The great popularity of web services is due to the adoption of open standards and protocols such as HTTP and XML.

The role of web service discovery is very important in using the web service, as it is a core functionality to locate the desired services. Service discovery is a process of finding the desired service by matching service descriptions against service requests. Discovery is the most central task in the web service model because web services are useless if they are not discovered. Web services are described using the service description language (WSDL) which gives the information regarding the particular service. This information can be used to find the desired web service and locate the service. If multiple web services provide the same functionality then Quality of Service can be used as criteria for service selection. QoS can be a combination of different qualities or properties of a service that are non functional. Some of the non functional attributes used for service discovery are availability, reliability, capacity, performance and cost. WSDL provide a limited information regarding the web service where as other semantic service description languages like OWL-S and WSMO gives more detailed semantic description of the web service. These descriptions are used for semantic web service discovery which gives a more accurate result.

This paper discuss various approaches of web service discovery based on service semantics and quality of service and also mentioned the importance of QoS in web service.

Rest of the paper is organised as follows. Section 2 describes web service discovery methods, Section 3 discusses the role of QoS in web service discovery, Section 4 describes how semantic description is used for service discovery, Section 5 describes some other approaches of web service discovery and Section 6 concludes with summary.

II. WEB SERVICE DISCOVERY METHODS

With the increasing number of available web services, service discovery has become a tedious task. Several services that have similar functionality are developed which makes the selection more difficult. Web service discovery process is carried out in three steps. First step is advertisement of web service by web service provider. Second step is web service request by user. Final step is selection and invocation of appropriate web service. Web service discovery based on users request basically can be done using keyword based search of user query on the service name. This method was modified by considering the service quality and semantics. There are mainly two broad classification of web service discovery. One is syntactic based service discovery and the next is semantic based service discovery.

Web services have very brief syntactic descriptions (from WSDL files). The lack of textual information makes keyword-based search models unable to filter irrelevant search results, and therefore, become very primitive means for effectively discovering Web services [1]. Keyword based search may give a huge number of service list. More over keywords are insufficient in expressing the semantics of the service and thus may retrieve irrelevant services in the context of consumers request. A lot of human interference is also needed due to the

Anupama P. Nayak and E. Sruthi, "A survey on classification of different techniques for web service discovery," *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 1, pp. 265-269, 2016.

inability of automatic service discovery, which makes them unusable in complex business environments.

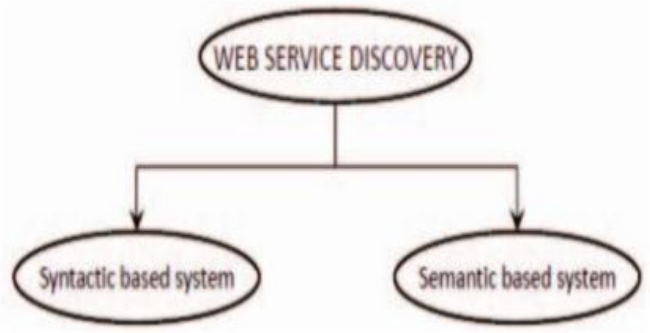

Fig. 1. Service discovery classification.

While original web service descriptions contained only information about the data types and bindings as a description of a Web service functionality, a number of different semantic languages have been created that allow describing the functionality of services in a machine interpretable form so as to promote the automation of web service discovery. Ontologies are used in the semantic Web service descriptions to describe the behaviour of a Web service. [10]. Ontologies are created by humans and there for contain natural language. NLP techniques can help overcome the ambiguity problems between different ontologies that are being used by semantic Web service descriptions.

QoS is a set of non functional attributes of a service provided by the web service. The QoS requirements for Web services are more important for both service providers and consumers since the number of Web services providing similar functionalities is increasing. The major issue in considering QoS for automatic service discovery is publishing the QoS and reliability of these published QoS. Different approaches were made to overcome these issues.

III. QOS BASED WEB SERVICE DISCOVERY

Ran proposed a service discovery model where quality of service is taken as constraints when searching for Web services [2]. This model proposed a QoS certifier and also proposed an extension to UDDI registry to add a new data structure type. This data structure type represents description of quality of service information about a particular service. The Web service provider first needs to communicate its QoS claim to the Web service QoS certifier. The certifier checks the claims and either certifies or down grade the claim. A consumer can retrieve the WSDL and QoS information of the services with desired quality by querying the UDDI registry and also can verify the QoS claim with certifier by using the certification id.

Ziqiang Xu proposed a Reputation enhanced QoS based service discovery model in which the advertised QoS information is expressed in XML style format and is stored using tModels in a UDDI registry [3]. This model contains a discovery agent and a reputation management system. Services that meet a customer's functionality and QoS requirements are ranked using the service reputation scores which are maintained by a reputation management system. The reputation management system is responsible for collecting and processing ratings of services from consumers, then updating the reputation score of the related service.

A similar model was proposed in paper [4] which has a new component service mediator. The aim was to refine the discovery process through designing a new framework that enhance retrieval algorithms by combining syntactic, semantic matching of service and also support QoS information exploiting the environment knowledge form user experience of service invocation and service initial run tests results. For every service or group of services there exist a service mediator agent which will handle all communication with registries, bindings, negotiations, voting, requests, responses for that service.

According to Kritikos, all QoS-based WS discovery algorithms fail to produce accurate results because they rely on either syntactic or semantically-poor QoS metric descriptions [5]. Hence, they cannot infer the equivalence of two QoS metrics based on descriptions provided by different parties. Provided that two QoS metric descriptions are expressed in OWL-Q, they developed a rule-based QoS metric matching algorithm that infers the equivalence of the two metrics. The algorithm used in this approach takes as input the QoS offers of all the WS advertisements and the QoS demand of the request in the form of OWL-Q specifications and returns four ordered lists of WS advertisements. The Alignment process used here aligns all offers and demand by finding their common QoS metrics by exploiting the QoS metric matching algorithm. The matchmaking process takes as input the CSPs of the offers and the demand and produces four types of results: super, exact, partial, fail with decreasing order of significance.

[6] Proposed an algorithm named QSSAC. This algorithm makes use of clustering process for service selection. Service clustering method groups the atomic services with same QoS properties. The QSSAC algorithm filters the candidate services at every level of composition. The services with good utility values are considered and other services with less utility values are eliminated. Re-selection of service is done if QoS value of selected service has changed or the service is failed during invocation.

IV. SEMANTIC DESCRIPTION FOR SERVICE DISCOVERY

Web services are commonly being described via narrative Web pages. These web pages contain information about their operations in natural languages and plain text with no machine interpretable structure. Therefore machines cannot automatically process the descriptive information about a Web service. Semantic service descriptions use ontologies which in turn use natural languages. Therefore natural language processing techniques are used to interpret these service descriptions and automatic service discovery.

SWEE is an efficient and effective search engine. SWEE [16] uses the WSDL files by parsing them using XML parsing technology and extracting the operations defined in WSDL

Anupama P. Nayak and E. Sruthi, "A survey on classification of different techniques for web service discovery," *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 1, pp. 265-269, 2016.

files. Operations and their corresponding WSDL files are organized as a lot of documents then finally search with keywords is done. Appropriate string matching and semantic processing is also used in SWEE search engine. The system is mainly composed of five modules: data set acquisition, data extraction and processing, index establishment, search execution and interface interaction. SWEE is more efficient than Strict and NAÏVE methods.

Paulraj and Swamynathan [7] proposed a method for content-based semantic Web service discovery. While in general approach the user queries are matched against OWL-S inputs, outputs, preconditions, and effects (IOPE), the framework allows users to submit free text as input. This alleviates the restrictions put on user queries in that they must be of the same format as that of the IOPEs present in OWL-S. In their work, nouns are extracted from text that is initially unstructured. These nouns are subsequently used for service discovery, after a disambiguation process that makes use of the WordNet lexical database for determining the meaning of the nouns.

[8] Present a new way of enhancing web services semantically using WordNet concepts. It allows developers to enhance web services with semantic information without semantic annotation against ontology. This approach tries to inherit all the benefits of WordNet. This model presents the semantics of messages in terms of WordNet concepts or synsets. Different senses of a word are grouped in different synsets. The meaning of a synset is further clarified with short definitions and/or example sentences. Discovery phase involves calculating the similarity of all registered services to the submitted query followed by returning the most similar services as the result of discovery and selection phase.

Clustering web services based on function similarities would greatly boost the ability of web services search engines to retrieve the most relevant web services. The WSDL file is mined to cluster similar web services. This is a predecessor step to retrieving the relevant web services for a user request by search engines. WSDL provides a model for describing non-semantic web services. Mine the WSDL documents to extract features that describe the semantic and behavior of the web service, specifically the WSDL content, WSDL types, WSDL messages, WSDL ports and the Web service name. These features describe and reveal the functionality of a web service. Similar web services are clustered by integrating these features together. This approach [9] shows how to extract the five proposed features such as WSDL Content, WSDL Types, WSDL Messages, WSDL Ports, and Web Service Name from WSDL documents.

[10] Proposed a Semantic Web Service Discovery framework for finding Semantic Web services by making use of natural language processing techniques. This framework enables end users to search, using keywords, for existing Web services described by means of a Semantic Web language for service annotation. This process consists of several steps including:

- The context of a Web service was created by extracting the information from semantic descriptions;
- Natural Language processing for disambiguating words' meanings and then establishing a context for a set of words;
- Matching the users search context with a Web service context by means of a similarity measure.

When Paulraj and Swamynathan focus specifically on OWL-S, the focus of this approach is on WSMO and aim for a more universal approach that can be utilized in various semantic Web service description languages. This approach also provides a means for ranking the results, while Paulraj and Swamynathan (2012) do not implement any form of result ranking.

Paper [1] proposes a well formed functional semantics to describe an operation of a Web service. The extendible functional knowledge is designed so as to map the requested or published operation descriptions into an abstract operation. Here a web service is defined as a network accessible software interface having collection of operations that aim at providing some kind of value to the consumers of the Web service. Thus Web service operation is nothing but the execution of appropriate action on specific object to provide value to the requester. In a functional semantic approach a natural way is used for expressing the operations functional in a web service. That is, action, nouns and objects are used for describing the operations. The web service requestors as well as providers express the service functionality using a restricted natural form.

Clustering similar web services is one of the efficient ways for discovering web services. In [11] ontologies are automatically generated by using complex terms and their underling semantics. In order to create the ontologies, initially extract the relevant feature from the WSDL file by using data mining techniques. The most important key point of the ontology construction is identifying the semantically meaningful concepts and relationships that exist between the concepts. After the preprocessing, the next is to find the TF-IDF value of all the tokenized words. It is important to identify the most frequently used words for complex terms. After finding the average TFIDF value, words are arranged in ascending order according to the TF-IDF value and rank the words by giving lowest rank to the word with lowest TF-IDF value and highest rank to the word with highest TF-IDF value. Various pattern analysis techniques applied to capture latent patterns which are hiding in complex terms.

[12] Proposed a method based on service tags. In this paper two algorithm QEBT and QPBT are described for service discovery using service tags. Web service contains many tags. Since lots of tags are tagged, some tags are inextract to the service. The service discovery method based on service tags is influenced by precision of service tags. QEBT is used to implement the query expansion based tags. QEBT is correlated with the threshold. QEBT gives the web services which are semantically correlative. Semantic relations of some of these web services are poor. QPBT is used to get some other web services whose semantic relation is closer. QPBT takes the web services whose semantic relationship is

Anupama P. Nayak and E. Sruthi, "A survey on classification of different techniques for web service discovery," *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 1, pp. 265-269, 2016.

poor, so that other web services which are semantically close are taken.

[13] Proposes WSMO-M (Mobile), an enhancement of WSMO to describe NFPs as a context and Quality of Web Service (QoWS) information for mobile computing environment. The analysis of web services discovery based on NFPs in mobile computing is presented. The aim is to enhance the accuracy of relevant web service discovered based on the capability and users requirement. The context and QoWS models are specified by using Web Service Modeling Ontology (WSMO). The semantic algorithm and Degree of Match (DoM) calculation are presented in order to rank the web services results based on the users' requirement and service offered. The applicability of the concept is then presented in a simple case study in Smartphone requirements for downloading a song and ringtone. Semantic matchmaking and degree of match calculation are also presented to define the importance of non-functional properties in mobile computing during the discovery and selection of web services.

[14] Proposes an approach to searching for web services, called LS3 (Lexical and Semantic Service Search) architecture. LS3 is used for calculating the similarity between web service descriptions and service queries. LS3 employs domain ontology and WordNet to measure the similarity of capabilities of web service descriptions between a service and request based on WSDL and SAWSDL. the main contribution offered by this study is that the proposed query expansion mechanism can increase the possibility of locating relevant services that cannot be retrieved using traditional textual service retrieval or ontology-based service matching. This approach includes two main subprocesses, query expansion and service ranking to enable the retrieval of relevant web services during the discovery process by considering lexical similarity and semantic similarity.

In [15] a crawler based system proposed that conducts a specialized crawl of the web to find only service descriptions in the form of WSDL documents. The service descriptions are semantically analyzed in order to improve efficiency of the discovery process as per user context during keyword search. Since the framework is inherently dynamic in nature, incremental change management strategy is incorporated to better optimize the working of the crawler. The system also extracts structural and functional information from WSDL documents and utilizes it for automatic tagging of the service descriptions.

In [17], a new method for extracting tags is proposed from WSDL files and mining semantic relationships between tags. It is foundation of building ontology, which can be used to annotate web service. Core tags of domain are extracted, and their semantic relationships are extracted from Wikipedia category graph. For a specific domain, the core tags should be mapped to articles of Wikipedia so that core tags can be transformed into concepts of the knowledge resource, and relationships between concepts can be extracted automatically. Yet disambiguation problem appears during mapping process. every core tag of a domain is mapped to a specific Wikipedia article. If no corresponding article exists, the core tag must be discarded. In practice, Wikipedia provides larger coverage concepts than other knowledge resources, such as WordNet.

V. Other Approaches for Web Service Discovery

The goal of [18] is to enhance discovery of web services through design of automated discovery algorithm. WSDL document for given domain name should be searched and parsed to obtain the invocation format. Then web service are automatically invoked. This algorithm makes use of several methods. First method makes use of Google web service API and developer kit to use the Google search engine for WSDL references. Second method makes use of Google web service API and developer kit to search Google for WSIL references. Third method makes use of web crawling to try and locate a WSDL document for a domain. The search is then implemented to query the registry. Then searching of a registry that is not associated with public UBR is done.

[19] Proposes connecting the isolated service islands into a global social service network to enhance the services' sociability on a global scale. First propose linked social service-specific principles based on linked data principles for publishing services on the open Web as linked social services. Then a new framework for constructing the global social service network should follow the linked social service-specific principles which are based on complex network theories. Next, an approach is proposed to enable the exploitation of the global social service network, providing Linked Social Services as a Service. Finally, experimental results show that this approach can solve the quality of service discovery problem, improving both the service discovering time and the success rate by exploring service-to-service based on the global social service network.

The WS-Discovery specification is sufficient only for single subnet networks. [20] Introduced a modification of the specification that allows for discovering services in different subnets without mandatory multicast. Target Services and Clients receive the IP address of the DP via a vendor specific option in DHCP and therefore do not need any additional static configurations. 2-layer discovery schema can be upgraded to a 3-layer discovery process by involving the Domain Name System (DNS) as an additional component. The information on the IP addresses of the DPs is stored in the global DNS instead of using inherently local DHCP services. Thus, if a user is looking for all services of a certain company, these can easily be browsed by looking up the DP using the DNS and then querying this DP for all available services. This would be an elegant way for realizing a global service directory that is managed in a completely decentralized way.

[21] Proposes design of a discovery cum publishing engine for web service discovery with refined searching mechanism which uses service rating techniques for efficient and effective web service discovery within optimum response time. Data mining techniques is used to narrow down the search space in UBRs. The proposed engine has an ability to publish or search web service across multiple UBRs. In addition to this an extended design of service registry is proposed to store service rating data along with the service information. The Engine and

Anupama P. Nayak and E. Sruthi, "A survey on classification of different techniques for web service discovery," *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 1, pp. 265-269, 2016.

rating have been utilized to help a user for selection of appropriate service. Publishes the web services in UBR by following a classification scheme and performs a validation test on discovered web services. Service reviews and rating have been utilized to help a user for selection of appropriate service.

VI. CONCLUSION

This paper aimed at giving an overview on different web service discovery techniques. More recent models are being proposed in this area which basically falls in any one of the classification that we discussed. New semantic languages are being introduced for giving a detailed semantic description of the web services. These descriptions are used to get a better service discovery based on the service semantics. And different approaches are used in storing, publishing, ranking and updating the service quality that results in a better QoS based service discovery. Natural language processing technique for is a recent approach for semantic web service discovery. A service crawler framework for similarity based web service discovery is a one more recent approach which uses crawler. The crawler gets the WSDL from the web and the functional structures are used for automatic tagging, resulting in a better discovery of web services.

REFERENCES

[1] D. A. D'Mello and V. S. Ananthanarayana, "Effective web service discovery based on functional semantics," *International Conference on Advances in Computing, Control, and Telecommunication Technologies*, 2009.

[2] S. Ran, "A model for web services discovery with QoS," *mSIGEcom Exchanges*, vol. 4, issue 1, pp. 1–10, 2004.

[3] Z. Xu, P. Martin, W. Powley, and F. Zulkernine, "Reputation-Enhanced QoS-based web services discovery," *IEEE International Conference on Web Services (ICWS)*, 2007

[4] M. Al Hunaity, Dr. A. El Sheikh, and Dr. B. Dudin, "A New Web Service Discovery Model Based On QoS," *ICTTA*, 2008.

[5] K. Kritikos and D. Plexousakis, "A semantic QoS-based web service discovery algorithm for over-constrained demands," *Third International Conference on Next Generation Web Services Practices*, 2007.

[6] Y. Xia, P. Chen, L. Bao, M. Wang, and J. Yang, "A QoS-Aware web service selection algorithm based on clustering," *IEEE International Conference on Web Services*, 2011.

[7] D. Paulraj and S. Swamynathan, "Content based service discovery in semantic web services using WordNet," *Advanced Computing, Networking and Security*, vol. 7135 of Lecture Notes in Computer Science. Springer, pp. 48–56, 2012.

[8] R. Karimpour and F. Taghiyareh, "Conceptual discovery of web services using WordNet," *IEEE Asia-Pacific Services Computing Conference*, 2009.

[9] K. Elgazzar, A. E. Hassan, and P. Martin," Clustering WSDL documents to bootstrap the discovery of web services," *IEEE International Conference on Web Services*, 2010.

[10] J. Sangers, F. Frasincar, F. Hogenboom, and V. Chepegin, "Semantic web service discovery using natural language processing techniques," *Preprint submitted to Expert Systems with Applications*, 2013.

[11] B. T. G. S. Kumara, I. Paik, and G. Lee, "Ontology learning method for web services clustering," *Computing and Convergence Technology (ICCCT) 7th International Conference*, 2012.

[12] Z. Ding, D. Lei, J. Yan, Z. Bin, and A. Lun, "A web service discovery method based on tag," *International Conference on Complex, Intelligent and Software Intensive Systems*, 2010.

[13] N. A. Saadon and R. Mohamad, "WSMO-M: NFP-aware web service discovery for mobile computing," *5th Malaysian Conference in Software Engineering (MySEC)*, 2011.

[14] S. P. Ma, C. H. Li, Y. Y. Tsai, and C. W. Lan, "Web service discovery using lexical and semantic query expansion," *IEEE 10th International Conference on e-Business Engineering*, 2013.

[15] S. S. Kamath, and V. S. Ananthanarayana, "A service crawler framework for similarity based web service discovery," *Annual IEEE India Conference*, 2014.

[16] Z. Qin, P. Li, Q. Zhu, and C. Tian, "SWEE: Approximately searching web service with keywords effectively and efficiently," *IEEE*, 2010.

[17] K. Xiao, K. Yang, and B. Li, "Extracting semantic relationships for web services based on wikipedia," *International Conference on Cloud and Service Computing*, 2011.

[18] J. Hicks, M. Govindaraju, and W. Meng, "Enhancing discovery of web services through optimized algorithms," *IEEE International Conference on Granular Computing*, 2007.

[19] W. Chen and I. Paik, "Constructing a global social service network for better quality of web service discovery," *IEEE Transactions on Services Computing*, 2013.

[20] S. Pöhlsen and C. Werner, "Robust web service discovery in large networks," *IEEE International Conference on Services Computing*, 2008.

[21] V. Tewari, N. Dagdee, and I. Singh, "An improved discovery engine for efficient and intelligent discovery of web service with publication facility," *World Conference on Services – II IEEE*, 2009.

Anupama P. Nayak and E. Sruthi, "A survey on classification of different techniques for web service discovery," *International Journal of Scientific and Technical Advancements*, Volume 2, Issue 1, pp. 265-269, 2016.