A Survey on Web Service Discovery Approaches

Abdelghany Mosa[1] and Ahmed Abdelaziz[2]

[1]Faculty of Business, Economics and Law, Auckland University of Technology, Auckland, New Zealand
[2]Nova Information Management School, Universidade Nova de Lisboa, 1070-312, Lisboa, Portugal
body_1986@live.com, D20190535@novaims.unl.pt

Abstract:
Service Oriented Architecture (SOA) is an approach to build distributed systems that deliver application functionality as services that are language and platform-independent. Web service is one of the fundamental technologies in implementing SOA based applications. Web services are modular, self-describing, self-contained and loosely coupled applications that can be published, located, and invoked across the web. As the number of web services is increased, finding a set of suitable web service candidates with regard to a user's requirement becomes a challenge. Web service discovery is the process of finding the most suitable service by matching service descriptions against service requests. Various approaches for web service discovery have been proposed. In this paper, we present an overview of different approaches for web service discovery described in the literature and try to classify them into different categories. We also determine the advantages and disadvantages of each category. The goal is to help researchers to propose a new approach or to select the most appropriate existing approach for service discovery.

Keywords: Web Service, Web Service Discovery, Semantic Web, Ontology, WSDL, UDDI.

1. Introduction

Service Oriented Architecture (SOA) is an approach to build distributed systems that deliver application functionality as services that are language and platform-independent. One of the key techniques in implementing SOA based applications is the web service [1]. A web service, as it is defined by the World Wide Web Consortium (W3C), is a software system designed to support interoperable machine-to-machine interaction over a network [2]. It is a modular, self-describing, loosely-coupled, platform and programming language-agnostic application that can be published, located and invoked across the Internet [3]. Web services are widely used because of their simplicity and data interoperability provided by their components namely XML (eXtended Markup Language), SOAP (Simple Object Access Protocol), UDDI (Universal Description, Discovery and Integration) and WSDL (Web Service Description Language) [1].

With the increasing number of web services available on the Internet, discovering web services concerning a user's requirement has become the most urgent problem to be resolved. Discovery is the process of finding a machine-processable description of a web service that may have been previously unknown and that meets certain functional criteria [2]. Discovery is the most central task in the web service model because web services are useless if they are not discovered.

Web service discovery has become a hot topic in the past few years. The task of finding the right service based on user's requirements from a pool of web services is still a problem. This is due to several issues including the following:

- Large number of web services on the Internet.
- Most of the searching is based on keywords rather than semantic.
- No standard WSDL format that a web service description complies with.
- UDDI contains static information that is only updated when the web service is registered or updated.

In this paper, we survey different approaches that offer solutions to the web service discovery problem and classify them into different categories according to their research focus. Furthermore, we discuss the advantages and

disadvantages of each category. This work is based on the papers that were published from 2005 to 2015. We investigate the following research questions:

- RQ1: What are different approaches used in web service discovery?
- RQ2: What kind of future research directions have been reported in the present literature on web service discovery?

The remainder of this paper is organized as follows: section 2 gives a brief about of web service discovery process. Section 3 introduces a taxonomy for web service discovery approaches. Section 4 describes service discovery approaches. We conclude the paper in section 5.

2. Web Service Discovery Process

A web service discovery process typically includes three major steps as shown in Fig. 1. In the first step, a service provider creates a web service and its service definition and then publishes the service with a service registry. In the second step, a service requester sends requests specifying the requirement in predefined format to a web service repository. Web service matcher matches the user's request with available web services and finds a set of web service candidates. In the last step, the service requester selects and invokes one of the retrieved web services [4]. Three factors mainly affect the web service discovery process: 1) the ability of the service providers to describe their services, 2) the ability of the service requestors to describe their requirements, and 3) the effectiveness of the service matchmaking algorithm [5].
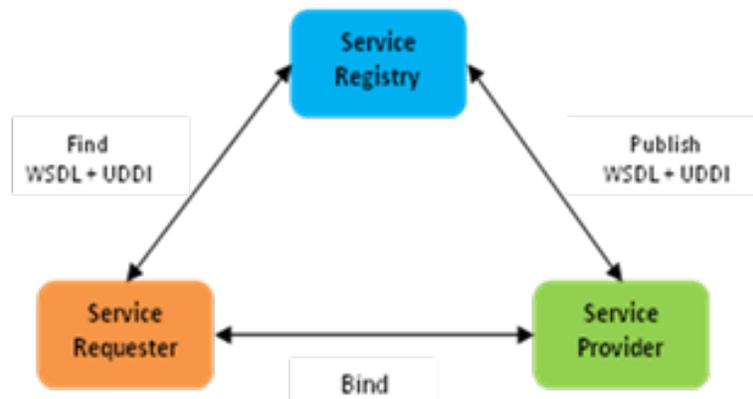

Figure.1 Web Service Discovery

3. Taxonomy of Web Service Discovery Approaches

This section presents a taxonomy of web service discovery. The selected papers were grouped into different categories according to their contents and research focus. The main classification of web service discovery is summarized in Fig. 2. The selected papers were initially classified into three broad categories namely: syntactic-based approaches, semantic-based approaches, and hybrid approaches. Additionally, the semantic-based approaches are further divided into domain-ontology approaches, general-ontology approaches, and agent-based approaches. Moreover, the selected papers were classified into other two categories: functional and non-functional categories as shown in Fig. 3.
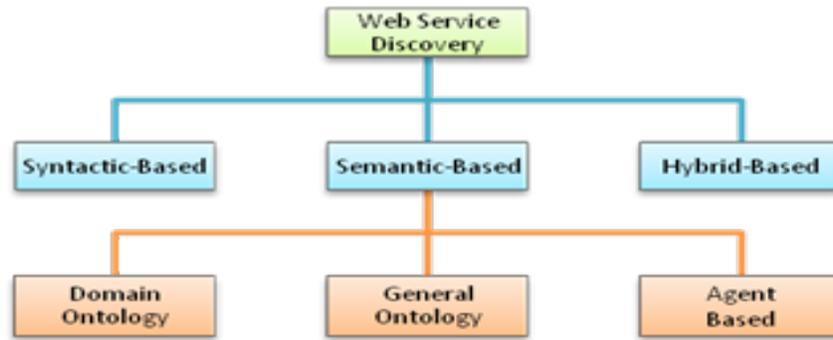
Figure.2 The Main Classification of Web Service Discovery
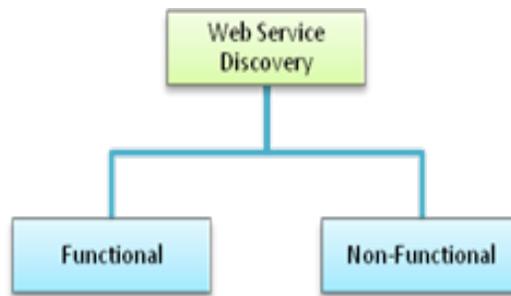
Figure. 3 Functional and Non-Functional Web Service Discovery Categorization.

3.1 Syntactic-based approaches

S. Huang, X. Wang and A. Zhou [6] presented effective service discovery and service composition strategies based on syntactical matching. Additionally, they introduced a software system to realize service discovery and composition algorithms. The architecture of this system consists of three main components. They are Index building, Web Service Discovery, and Web Service Composition. Index building Component includes three kinds of indices: IMIndex (Input Message Index), OMIndex (Output Message Index), and OpIndex (Operation Index). Web Service Discovery Component realizes service discovery corresponding to users' requirements. Web Service Composition Component realizes composition for corresponding queries.

Y. Badr, A. Abraham, F. Biennier and C. Grosan, [7] proposed a simple web services selection scheme based on the user's requirement of the various non-functional properties and interaction with the system. The proposed framework utilizes user preferences as an additional input to the selection engine and the system ranks the available services based on the requirement.

S. Hamza, K. Okba, B. Aïcha-Nabila and A. Youssef [8] proposed architecture based on mobile agents for web services discovery in a cloud computing environment. The architecture is composed of two areas of cloud. The first deals with Keywords based research and the second supports the filtering of the found web services. Additionally, they proposed a new algorithm of comparison between the request of the client and the description of the web services. However, the proposed algorithm doesn't take into account non-functional aspects such as quality of service, cost, etc.

Another framework for discovering web services with the help of functional and non-functional information was proposed by E. Kirubakaran, D. Ravindran and D. I. George [9]. The proposed framework consists of an interface module interacting with a service provider, service consumer and service registries. It supports two different levels of searches. The first one is based on keyword matching in which the functional data is checked and some services are identified. The second level uses the non-functional information to refine the identified services in level 1.

R. Karthiban [10] proposed a novel technique to mine Web Service Description Language (WSDL) documents and cluster them into QoS similar web service groups. In this model, the services are clustered based on the QoS values in order to achieve better service selection. K-means clustering algorithm is used to cluster the QoS values. The proposed model includes four kinds of steps. They are filtering the web service based on keywords; extracting the QoS values from WSDL document; forming the cluster using the extracted QoS values; and selecting the most suitable web service from the cluster. Its advantage is that it supports dynamic web service selection at run time. However, more QoS parameters need to be included.

There are some advantages and disadvantages of syntactic based approaches that we mention briefly below.

Advantages of syntactic-based approaches:

- They are Simple and widely used.
- Keyword-based search is more familiar to the user.
- Standard like UDDI exist.

Disadvantages of syntactic-based approaches:

- Syntactic discovery leads to low precision and low recall of the discovery results.
- Keywords are insufficient in expressing semantic concepts.
- They can't retrieve web services with similar functionality.
- They don't support automatic web service discovery.
- Different words can be used in different terms inside different domains, so some irrelevant information can be returned.

3.2 Semantic-based approaches

3.2.1 Domain-ontology approaches

K. Zamanifar, A. Zohali and N. Nematbakhsh [11] proposed a matching model that uses semantic description and quality description of web service based on ontology as the basis of service matching. Moreover, a user can specify weights for all parameters, so the users can tell which parameter of his/her request is premier and should provide and which one is unnecessary. The service matching method is composed of two phases: semantic similarity matching and qualitative filtering. Semantic similarity matching is used for matching algorithms in order to find the most proper services, and qualitative is used for selecting the best service among results. The performance of web service matching can be improved effectively with this approach.

M. Wang, X. Li and X. Qiao, [12] focused on how to reduce service discovery response time, without affecting discovery accuracy. They proposed a novel semantic web service discovery method based on the user preference cluster. They optimized the design of Unmixed Semantic UDDI Model. This model includes six components: Telecom Service Domain Ontology Library (TSDOL), Service Profiles Ontology Library (SPOL), User Preference Ontology Library (UPOL), Service Publish, Configuration Management, and Service Discovery. User Preference Cluster algorithm was introduced to preprocess the user preference from the standpoint of user demand before service discovery. The service discovery response time is minimized with this method.

A. Yousefipour, A. G. Neiat, M. Mohsenzadeh, and M. A. Seyyedi [3] proposed a new QoS-aware framework to improve the semantic web service discovery based on the broker by using ontology concepts. This framework is capable of supporting the QoS management, governance, monitoring and assurance in delivering SWS at runtime. It is based on the commonly used concept of QoS brokerage service. This QoS broker consists of the following components: Request Manger, Ontology Manager, QoS Database, Administration Manager, and Quality Manager. Ontologies are used to improve web service discovery and provide best selection of web services according to user preference. One of the problems of the proposed framework is to extend it by considering a federation of QoS brokers.

An approach for semantic web services discovery was presented by L. Zhou [13]. In this approach, the search is divided into three steps. A traditional vector-space model retrieves the most similar services according to their WSDL service descriptions at first. The second step includes matching web services with ontology hierarchically. The last step includes matching web services with QoS.

G. Wen-yue, Q. Hai-cheng and C. Hong [14] proposed a semantic web service discovery algorithm based on OWL-S. The proposed algorithm is composed of three layers: service category matching, service functionality matching and

quality of service matching. In the first layer, the service category matching degree is computed and the advertisements that satisfy the requirements are selected to enter the next layer. In the second layer, the service functionality matching degree is computed and the advertisements that satisfy the conditions are selected to enter the next level. In the third layer, the quality of service matching degree is computed. According to the service category matching degree, service functionality matching degree and quality of service matching degree, the service matching degree is calculated and the best advertisements are presented to the requesters.

V. Oleshchuk [15] proposed an ontology-based approach for matching service descriptions and service requests which finds semantic similarity between service descriptions and service requests by using implicit knowledge from ontologies. The proposed approach utilizes knowledge from ontologies to enhance both user service requests and service descriptions by adding concepts that are not presented in the original descriptions, and use them in the comparison process. An algorithm called Art also was suggested that can be used to find semantic similarity between service request and service description with respect to a given ontology.

L. N. Kolhe, G. S. Mary and R. D. Pathari [5] proposed a novel approach for semantic-based automated service discovery. This approach focuses on two major aspects: semantic-based service categorization and semantic-based service selection. For semantic-based service categorization, they proposed an ontology guided categorization of web services into functional categories for service discovery. The semantic-based categorization is performed offline at the universal description discovery and integration (UDDI). For semantic-based service selection, they employed ontology-based request enhancement and LSI based service matching.

There are some advantages and disadvantages of domain ontology approaches that we mention briefly below.

Advantages of domain ontology approaches:

- Semantic web technology provides more precise and recallable results.
- Semantic web technology web services make web services more accessible to automated agents.
- SWSs support automatic discovery, automatic execution, automatic combination and automatic interaction.
- Using semantics can improve the relevancy of web service discovery.
- Effective and reliable approaches.

Disadvantages of domain ontology approaches:

- The cost of ontologies is relatively high and need special expertise for creation and maintenance.
- End users should have intimate knowledge of semantic web services which makes their usage difficult for end-users.
- Ontology mapping techniques are needed to overcome ontology heterogeneity.

3.2.2 General-ontology approaches

G. Ganapathy and C. Surianarayanan [16] suggested a two-stage filtering approach to identify candidate services with the required trust score during semantic web service discovery. These two stages are filtering by relevance and filtering by trust score. In the first stage, the candidate services are identified by finding relevant services using implicit semantics of service description and senses from WordNet. In the second stage, the trust score of candidate services is compared against a user-defined trust score using Trust Rank algorithmic tool. The candidate services with a greater trust score than the user-defined are chosen for semantic matching of services using a semantic reasoner.

Y. Peng [17] proposed a two levels semantic web service discovery method. At the first level, the service similarity degree is computed. At the second level, the interface similarity degree is computed. This method use a semantic word in WordNet to annotate service and service interface. The proposed method aims to locate target services quickly and precisely. A serial of experiments was conducted that showed that the proposed method can improve the precision of service discovery and have good scalability. The limitation of this approach is, it doesn't consider the non-functional constraint in the service discovery.

Y. Y. Du, Y. J. Zhang and X. L. Zhang [18] presented an approach for service discovery based on service clustering and refining service clusters. In this approach, WordNet is used as a public ontology to calculate the semantic similarity between ontology concepts to group services into functionally similar clusters. The time efficiency of the service discovery can be improved with this approach. However, many problems are needed to be resolved such as considering service quality and maintaining the service clusters.

There are some advantages and disadvantages of general ontology approaches that we mention briefly below.

Advantages of general ontology approaches:

- WordNet is a well known online dictionary that is an ideal ontology source.
- There are several kinds of relations between concepts in WordNet, such as synonymy, antonymy, and hypermymy.
- These approaches enhance Web Services with semantic information without semantic annotation against ontology.
- WordNet is not domain-specific and eliminates the semantic annotation cost of services.

Disadvantages of general ontology approaches:

- WordNet is too fine-granular for many purposes.
- There is no real multilingual WordNet.
- WordNet focuses on paratactic semantic relations between single words.

3.2.3 Agent-based approaches

A.G. Neiat, M. Mohsenzadeh, S.H. Shavalady and A. M. Rahmani [19], proposed an approach for semantic web service discovery and propagation based on semantic web services and FIPA multi-agents. The main part of the proposed approach is a broker. The broker provides semantic interoperability between semantic web service providers and agents by translating WSDL to DF (Directory Facilitator) description for semantic web services and vice versa. Also, ontology management in the broker was proposed, which creates the generalized ontology by merging user ontology with a general ontology (i.e. WordNet). However, this approach is facing some challenge such as inconsistencies during the merge and evolution of created WSDL based on the functional and non-functional requirements (i.e. quality of service).

R. Benaboud, R. Maamri and Z. Sahnoun [20] presented a multi-agent framework based on ontologies for web services discovery. The proposed framework aims to simplify service discovery using semantics while satisfying QoS requirements. With the help of agents, information provided by web services can be made more efficient and more dynamic. With the use of OWL-S and domain ontologies, the most relevant services can be returned. This framework consists of two types of agents: Consumer Agent and Provider Agent.

There are some advantages and disadvantages of agent-based approaches that we mention briefly below.

Advantages of agent-based approaches:

- Software agent technologies will help enable new and advanced operational and usage modalities of Web services.
- With the help of software agents, the information provided by web services can be made more efficient and more dynamic.
- Agents are a necessary complement to web services to realize the vision of Semantic Web.
- An advantage of agent-based web services is that they combine the strengths of both web services and multi-agents.

Disadvantages of agent-based approaches:

- Agents and semantic web both use different service registries, service description languages and communication protocols.

3.3 Hybrid Approaches

Y. Tsai, S. Hwang and Y. Tang [21] proposed a hybrid approach that combines both text-based searching and ontology-based matchmaking for discovering web services that satisfy users' functional needs. The proposed approach uses multiple criteria decision-making techniques to determine the weights of different attributes. This approach considers service providers information, service descriptions by providers, service descriptions by users, operation descriptions by providers, tags and categories and also QoS attributes. However, this approach does not take into account the precondition and effect of web service operations.

Another approach for semantic web services discovery was proposed by R. Benaboud, R. Maamri and Z. Sahnoun [1]. This approach is based on an architecture composed of four layers: Web service and Request description layer,

Functional match layer, QoS computing layer and Reputation computing layer. In the first layer, the service providers register their web services and provide functional and non-functional information about the offered services. In the second layer, functional properties of request and web service descriptions are matched based on the syntactic and semantic match. In the third layer, the QoS score of each candidate web service is calculated. In the last layer, the reputation score of each candidate web service is calculated. Each layer uses the result of the previous layer and aims to decrease the number of candidate web services.

S. Pakari, E. Kheirkhah and M. Jalali [22] proposed a hybrid approach for service discovery in a service-oriented architecture. In this approach, three similarities (syntactic similarity, structural similarity, and semantic similarity) are computed, and then they are combined with a weighted averaging to have better results.

There are some advantages and disadvantages of hybrid-based approaches that we mention briefly below.

Advantages of hybrid approaches:

- They utilize the advantages of both the syntactic approaches and semantic-based methods for discovering Web services that satisfy users' functional needs.
- They can achieve more accurate results and discover more suitable services.

Disadvantages of hybrid approaches:

- More complex approaches.

4. Conclusion

This paper has given an overview of different approaches available for web service discovery. We have grouped them into different categories according to their research focus. We have also discussed the advantages and disadvantages of each category. Most of the approaches differ in the way web service matching is carried out. Some approaches focus on keyword-based search, while others are considering the concept of the semantic web. Syntactic discovery leads to low precision and low recall of the discovery results. Using semantics can support automatic discovery and improve the relevancy of web service discovery. Some approaches utilize the advantages of both syntactic-based and semantic-based approaches and hybrid them to achieve more accurate results. Therefore, we answered RQ2 presented in the introduction by showing these results. As the number of available web services that provide the same functionality is very large, it is a major issue to consider not only the functional requirements but also the non-functional requirements (e.g. reliability, availability, etc) during the web service discovery process. Concerning RQ2, some future research dictions that have been reported in the present literature include the following:

- Ontology mapping techniques are needed to overcome the ontology heterogeneity in semantic web service discovery.
- Automating web service discovery.
- Dynamic web service composition.
- Combining functional and non-functional properties for service discovery.

References

[1] R. Benaboud, R. Maamri, and Z. Sahnoun, "Agents and Owl-S Based Semantic Web Service Discovery with User Preference Support," International Journal of Web & Semantic Technology (IJWesT), vol. 4, no. 2, Apr 2013.

[2] D. Booth et al. (2002). Web Services Architecture. [Online]. Available: http://www.w3.org/TR/ws-arch/.

[3] A. Yousefipour, A. G. Neiat, M. Mohsenzadeh, and M. Seyyedi, "A New Broker-Based Semantic Web Service Discovery Framework for Selecting And Ranking Suggested Web Services," in IEEE int. Conf. Intelligent Computer Communication and Processing, Aug. 2010.

[4] K. Gottschalk, S. Graham, H. Kreger, and J. Snell, "Introduction to Web Services Architecture," IBM Systems Journal, vol. 41, no. 2, pp. 170-177, 2002.

[5] L. N. Kolhe, G. S. Mary, and R. D. Pathari, "Semantics Based Automated Service Discovery," International Journal of Scientific Research Publications, vol. 4, Nov. 2014.

[6] S. Huang, X. Wang, and A. Zhou, "Efficient Web Service Composition Based on Syntactical Matching," in Int. Conf. e-Technology, e-Commerce and e-Service (EEE'05), 2005.

[7] Y. Badr, A. Abraham, F. Biennier, and C. Grosan, "Enhancing Web Service Selection by User Preferences of Non-functional Features," in 4th Int. Conf. Next Generation Web Services Practices, Oct. 2008, pp. 60-65.

[8] S. Hamza, K. Okba, B. Aïcha-Nabila, and A. Youssef, "A Cloud Computing Approach Based on Mobile Agents for Web Services Discovery," in Int. Conf. Innovative Computing Technology, Sep. 2012, pp. 297-304.

[9] E. Kirubakaran, D. Ravindran, D. I. George, "Service Discovery Framework with Functional and Non-Functional Information," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 2, Dec. 2012.

[10] Darthiban, R., "a QoS-Aware Web Service Selection Based on Clustering," International Journal of Science and Research Publication, vol. 3, Feb. 2014.

[11] K. Zamanifar, A. Zohali, and N. Nematbakhsh "Matching Model for Semantic Web Service Discovery," in Int. Conf. Advanced Information Networking and Applications Workshops, May. 2009.

[12] M. Wang, X. Li, and X. Qiao, "Semantic Web Service Discovery Based on User Preference Cluster," in 3rd IEEE Int. Conf. Broadband Network and Multimedia Technology, 2010.

[13] L. Zhou, "An Approach of Semantic Web Service Discovery," in Int. Con. Communications and Mobile Computing, Apr. 2010.

[14] G. Wen-yue, Q. Hai-cheng, and C. Hong, "Semantic Web Service Discovery Algorithm and Its Application on the Intelligent Automotive Manufacturing System," in 2nd IEEE Int. Conf. Information Management and Engineering, Apr. 2010, pp.601-604.

[15] V. Oleshchuk, "Ontology-Based Service Matching and Discovery," in 6th Int. Conf. Intelligent Data Acquisition and Advanced Computing Systems, Sep. 2011.

[16] G. Ganapathy, Surianarayanan, C., "An approach to Identify Candidate Services for Semantic Web Service Discovery," in IEE Int. Conf. Service-Oriented Computing and Applications, Dec. 2010, pp. 1-4.

[17] Y. Peng, "Two Levels Semantic Web Service Discovery," 7th Int. Conf. Fuzzy Systems and Knowledge Discovery, Aug. 2010, pp. 2523-2526.

[18] Y. Y. Du, Y. J. Zhang, X. L. Zhang, "A Semantic Approach of Service Clustering and Web Service Discovery," Information Technology Journal, vol. 12, 2013.

[19] A. G. Neiat, M. Mohsenzadeh, S. H. Shavalady, and A. M. Rahmani, "A New Approach for Semantic Web Service Discovery and Propagation Based on Agents," in 5th Int. Conf. Networking and Services, Apr. 2009,pp.37-42.

[20] R. Benaboud, R. Maamri, and Z. Sahnoun, "Semantic Web Service Discovery Based on Agents and Ontologies," International Journal of Innovation, Management and Technology, vol. 3, no. 4, Aug. 2012.

[21] Y. Tsai, S. Hwang, and Y. Tang, "A Hybrid Approach to Automatic Web Services Discovery," in Int. Joint Conf. Service Sciences, May. 2011, pp. 277-281.

[22] S. Pakari, E. Kheirkhah, and M. Jalali, "Novel Approach: A Hybrid Semantic Matchmaker For Service Discovery In Service Oriented Architecture," International Journal of Network Security & Its Applications, vol. 6, no. 1, Jan. 2014.