SEWSIP：Semantic based Web services Integration in P2P

LI Juanzi XU Bin YANG Wenjun CHEN Dewei ZHANG Po WANG Kehong
Department of Computer Science and Technology, Tsinghua University
Beijing, China, 100084
{ljz, xb, ywj, chendw, zhangpo, wkh}@keg.cs.tsinghua.edu.cn

Abstract

Semantic based Web services integration aims to add valuable semantic information in Web services Description and to implement the semantic based Web services discovery and integration. This paper puts forward a semantic based Web service integration framework in P2P-SEWSIP. SEWSIP can collect Web services in distributed web environment, lift the WSDL files into OWL-S by the use of XML schema in WSDL and domain ontology, and then deploy them in P2P. It also can provide service discovery, evaluation and composition based on semantic information in OWL-S.

1. Introduction

Web service has been becoming an important technology, such as services on demand and SOA. But there still exist some problems including: (1) Web services are stored in distributed instead of in central location. Web service providers may not register their Web services in UDDI. They may only give their service description in company website or in other web pages. How to find these Web services to make use of them? (2) WSDL is lack of semantic information and it is difficult to implement the intelligent and automatic Web services. (3) Web services do not provide the ability of service composition when existing individual Web service can not meet the need of customer.

There have been a lot of efforts in these topics. METOR-S[2], which is developed by Georgia University, exploits semantic web, workflow and Web services to perform service discovery and composition. It establishes P2P network between UDDI servers so that they can share the Web services between different UDDIs. Semantic Discovery Service (SDS)[1], which is developed by Stanford University, is a union of the industry process-modeling standard Business Process Execution Language for Web services (BPEL4WS) with the OWL-based Web services Ontology (OWL-S) and associated Semantic Web reasoning machinery to perform dynamic service binding of BPEL4WS Web services compositions based on use's personal preferences and constraints and Semantic translation to enable interoperability between disparate services. Maryland University developed a service composition prototype that guides users in creating a workflow of services step by step. Users select services in the context of a composition step [3]. The system supports further, user-driven filtering of the compatible services based on other service features described against generally available OWL ontologies.

This paper puts forward a Web services integration platform based on semantic web and P2P. The platform can automatically find Web services based on specific domain ontology. Then these Web services can be annotated and transformed into OWL-S so that they have richer semantic information. Based on user's requirement, SEWSIP can perform Web service discovery, selection and composition in a semi-automatic may. The integrated Web services are transformed into BPEL for the execution of the services.

The paper is organized as follows. Section 2 describes the some terminologies throughout the paper. Section 3 presents the framework of the Web services integration platform based on the semantic web and P2P. Section 4 presented the partial implementation in SEWSIP in more details. Section 5 is the conclusion of the paper.

2. Terminology and Definition

This section introduces the basic definitions and terminologies in the research on the framework of Web services integration used throughout the paper.

2.1. Web services description Language-WSDL

Proceedings of the 2005 IEEE International Workshop on Service-Oriented System Engineering (SOSE'05)
0-7695-2438-9/05 $20.00 © 2005 IEEE

Currently, Web services are described in the format of WSDL. WSDL is an XML document that mainly consists of four elements <PortType>, <Message>, <Types> and <binding>. <portType> defines the operations performed by the Web services. <message> defines the messages used by the Web services, <types> defines the data types used by the Web services and <binding> define communication protocols used by the Web services.

2.2. Ontology and Domain ontology

Ontology is a formal, explicit specification of a shared conceptualization of a specific domain[5]. The main components of ontology are concepts, relations, instances and axioms. A concept represents a set of classes, entities or 'things' within a domain. The concepts can be organized in a hierarchy structure. Relations describe the interactions between concepts or properties of a concept. Instances are the "things" represented by a concept. From the definition, domain ontology defines the semantic of the information in a specific domain. Ontology also provides a lot of necessary knowledge to enable the implementation of semantic Web services.

2.3. OWL-S

How to describe the Web services accurately and how to make them understandable are two important problems to support the matching between user requirement and service description. Ontology is introduced to enhance the machine understandability to Web services and supports the inference ability. At present, there are many different kinds of semantic based Web services modeling languages such as OWL-S[6], WSMO[7] and WSDL-S[1]. OWL-S describes a Web services using ServiceProfile, ServiceModel and ServiceGrounding. ServiceProfile states what the service provides for users and what are provided by users. ServiceModel states how the service works. ServiceGrounding states how to use the service. In this paper, OWL-S is used to describe the semantic of Web services.

2.4. CPSDL

CPSDL(Composition Process Semantic-based Description Language) is proposed in SEWSIP to present the process of semantic based Web services composition from two levels: semantic level and execution level. The semantic lever describes the syntax and semantic of inter activities in a process flow and execution level describes the information of each activity including its bounding and message links.

3. SEWSIP: Framework for Semantic Web services Integration

SEWSIP (SEmantic Web services Integration Platform) is a platform which uses domain ontology to implement the service discovery, evaluation, selection and semi-automatic composition. It is established using the semantic web and Peer-to-peer technologies. Figure 1 shows the diagram of the framework.

As shown in Figure 1, Web services for a specific domain are collected from internet or UDDI centre, and then they are extracted and annotated into OWL-S. These Web services are deployed in P2P environment. When a user wants to generate an application, he/she first inputs his/her requirement in a graphic user interface; then service selection is performed to select Web services according user requirement. When existing services can not meet user's requirement, user interface will uses domain ontology and CPSDL to prompt user to decompose his/her service into sub services, and then composite the Web services into an application.

Web services collection and classification: Web services collection and classification find Web services and filter the Web services of a specific domain. A Crawler finds the Web services and a classifier is employed to filter the services of a specific domain. The former should be able to collect enough Web services; the later should find Web services in specific domain correctly.

OWL-S annotator: Ontology is introduced to enhance the machine understandability of Web services and support the inference ability. OWL-S has been widely used to describe Web services by using ontology. At present, the services are described in WSDL format. So, in order to exploit the semantic of Web services, the WSDL files should be translated into OWL-S formats. Manual transformation is tedious and expensive. WSDL is in XML schema format, which include many useful cues to show the semantic of Web services, such as the name of complexType, element tags and so on. Main task of OWL-S annotator is to exploit the information in XML schema and domain ontology to translate the WSDL into OWL-S.

The service deployment in P2P: There would be no need for a central registry of Web services making the whole architecture more scalable in order to accommodate continuous addition of Web services. Services deployment is responsible to deploy the services in Peer-to-Peer environment properly. Peers are organized in peer groups and each group can store certain kinds of Web services. When a service is to be deployed, service deployment will find a proper peer group for it and insert it into the group by using routing table.

Service discovery engine: User requirement includes the description about domain, semantic, functions and key words of services etc. Specific domain ontology are employed to make up the understanding gap between machine and user. The service matching between user requirements and OWL-S based Web services is another important research topic. Using semantic information, service discovery can not only perform exact matching, but also can find the potential matching by using the inference engine so that to increase the possibility to find Web services.

Services evaluation and selection: The output of service discovery engine is a list of services discovered. How to evaluate the retrieved services and how to select the best services are the tasks of the component. When evaluate the services, many kinds of information can be exploited such as the service description information, service usage information and so on.

Service composition: Automatic Web services can not be accomplished due to the related Web services technology and specifications. At present, semi-automatic Web services composition is a reasonable solution. Service composition should have the functions that can navigate user to decompose service into proper sub services and combine them into an executable application. So, the key problem of service composition is to control the services decomposition and transformation of heterogenic messages.

Services validation and execution: The composed service in SEWSIP is a practical application which is presented in BPEL. In the execution, this component will validate the Web services to decide whether the service flow is reasonable.

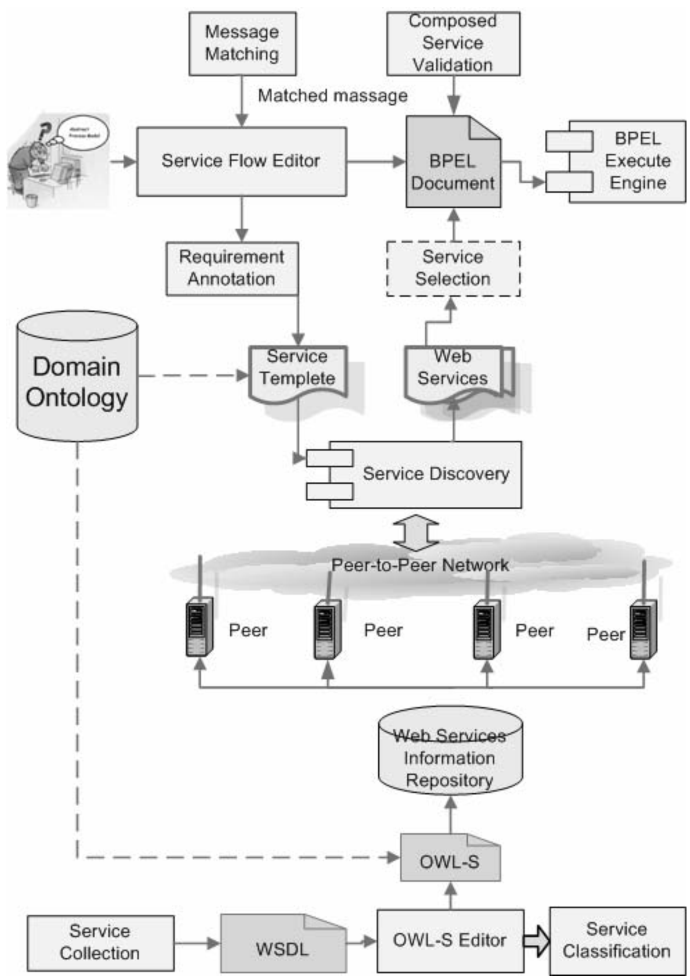

Figure 1. Diagram of Semantic service integration

4. The Partial Implementation of SEWSIP

The practical system of SEWSIP is consisted of service annotation, service discovery based on P2P, service composition and service exaction and monitoring. The relationships between these components are shown in figure 2.

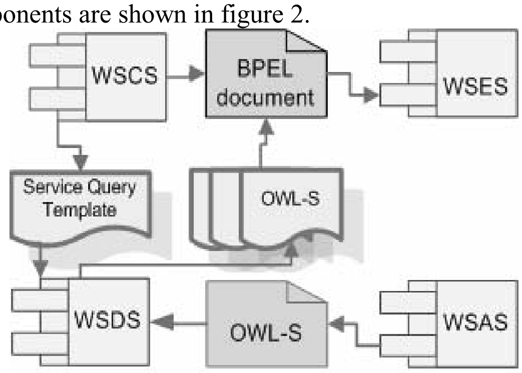

Figure 2. Components in SEWSIP

4.1. Web Service Annotation System-WSAS

In most service annotation methods, the input and output in WSDL are annotated according to the name of parameters. In fact, the XML schema in WSDL also provides the semantic information such as complexType and datatype. So, service annotation in SEWSIP uses two processes to perform the annotation. At first, we convert WSDL into an immediate ontology format. The semantic in WSDL are extracted and wrapped into a temporary ontology. For example, each complexType can be transformed into a concept, and elements and attributes are transformed into properties. In this way, a lot of semantic information can be contained in this temporary ontology. Secondly, a mapping between temporary ontology and domain ontology with OWL format is performed. When the mapping is finished, the WSDL is transformed into OWL-S.

4.2. Web Service Discovery System-WSDS

Service discovery component performs two functions because SEWSIP is a service integration framework in P2P. The first function is the peer routing based on keywords and the second is the precious matching based on semantic in OWL-S. The former can locate the services in proper peer group, and the latter can get the best matching result by inference in proper peers [9].

1. Routing in P2P based on keyword matching

This process exploits traditional keywords information retrieval. Keywords input by user are viewed as a query, and the vector of one peer group which integrated all peer vectors in the group is viewed as the document to be matched. Similarities are calculated between the query and vectors of each peer groups. The peer groups with large similarities are determined to be groups that may have the services user needed. The peer groups are organized as shown in figure 3.

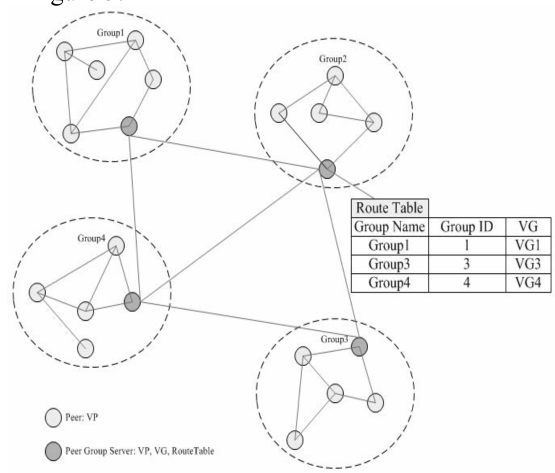

Figure 3. Organization of P2P

As Shown in Figure 3, each Web service is deployed in a peer group. Each peer group has several peers and contains Web services of certain kinds. Each group has a group server to manage the peers in the group and to join in the activities in P2P. The information about the group and Web services in the group are recorded in routing table.

2. Precious matching based on semantic information

In the first matching process, possible matched peers are constrained in a limited scope. The main task of the second process is to exploit the semantic in service description and a reasoner to perform precious matching. The matcher compares the service goal and capability of OWL with the user requirement by using FOL reasoning. When the OWL-S of services can meet the user requirement, the matched services are selected. Because the computing complexity of precious matching is much higher than that of keyword matching, the design of two phase match processes can save the computing cost in great deal.

4.3. Web Service Evaluation System-WSES

The accuracy of service evaluation is the key to selecting the best Web services from many available ones which provide overlapping and identical functionality. In SEWSIP, this component presents a new service evaluation model using ontology technology to describe model concepts, which supports

Proceedings of the 2005 IEEE International Workshop on Service-Oriented System Engineering (SOSE'05)
0-7695-2438-9/05 $20.00 © 2005 IEEE

to customize evaluation factors dynamically. It combines the machine learning algorithm and transcendent knowledge to compute the weight of each evaluation factor. Experiments show that this model gets more accurate evaluation results in comparison with others. This model can be used in Web services management system, Web services selection and composition engine, etc. More detail about evaluation model can refer the paper [8].

4.4. Web Services Composition System-WSCS

This component provides the input of the service discovery and accomplishes service composition based on the service discovery and selection components. The service composition in SEWSIP is responsible for the design of service flow, the binding of services and the message transforming. User interface lets user input his/her requirements and uses domain ontology to decompose his/her services into sub services. When the service discovery component finds the proper services, these services are composed into the service process flow. Otherwise, the component will continue to decompose the services until the user requirement can be matched.

Different from other composition component, this component also provides a transformer to enable the matching of messages between services. In most composition methods, the composed services can not execute because of the mismatch between service input and output. In SEWSIP, we use a XML transformer to perform the matching in a semi-automatically way.

All information of service flow is described in CPSDL. After composition, CPSL are translated into BPEL and become an executable application.

5. Conclusion

This paper discusses the semantic based Web service integration in P2P and presents a framework named SEWSIP. SEWSIP has following features: (1) can collect and filter Web services of a specific domain; (2) uses a lot of semantic information in WSDL to annotate the WSDL to improve the correctness of OWL-S files; (3) uses two phase service discovery to find the proper services, whose algorithm not only can save the discovery cost but also can perform precious matching based on semantic inference; (4) semi automatic service composition exploit the domain ontology and graphic user interface to aid user decompose his/her requirement into sub services or recommend user possible composition, which can save user workload; (5) A XML transformer is employed to perform the matching between the messages of input and output so that the composed application is a executable application.

In our further work, the main researches about SEWSIP includes: (1) collect more WSDL on Internet and develop high precious annotator of WSDL; (2) make full use of semantic and composition pattern to alleviate the manual work in service composition; (3) improve the accuracy of semantic based service match algorithm.

Reference

[1] Katia Sycara, Massimo Paolucci, Anupriya Ankolekar and Naveen Srinivasan, Automated Discovery, Interaction and Composition of Semantic Web services, Journal of Web Semantics, Volume 1, Issue 1, September 2003, pp. 27-46.

[2] Kunal Verma, Kaarthik Sivashanmugam, and Amit Sheth et al, Journal of Information Technology and Management, Special Issue on Universal Global Integration, Vol. 6, No. 1 (2005) pp.17-39. Kluwer Academic Publishers.

[3] Evren Sirin, Bijan Parsia, and James Hendler, Filtering and Selecting Semantic Web services with Interactive Composition Techniques, IEEE Intelligent Systems, 19(4): 42-49, 2004.

[4] E. Christensen et al., Web services Description Language (WSDL) 1.1, World Wide Web Consortium (W3C) note, 2001; www.w3.org/ TR/2001/NOTE-wsdl-20010315.

[5] T. Berners-Lee, J. Hendler, and O. Lassila, The Semantic Web, Scientific American, vol. 284, no. 5, 2001, pp. 34–43.

[6] Sean Bechhofer, Frank van Harmelen, et al. OWL Web Ontology Language, World Wide Web Consortium (W3C) note, 2004, http://www.w3.org/TR/owl-ref/

[7] Michael Stollberg, et al. D3.3 WSMO Use Case "Virtual Travel Agency" v0.1, WSMO Working Draft 08 October 2004. http://www.wsmo.org/2004/d3/d3.3/v0.1/20041008/

[8] Yang Wen-Jun, Li Juan-Zi,Wang Ke-Hong, Domain adaptive service evaluation model. Chinese Journal of Computers, Vol. 28 , No. 4. pp.514- 523.

[9] CHEN Dewei, XU Bin, CAI Yueru, LI Juanzi, Deployment and Publication Bound and P2P Based Web services Discovery Mechanism, Chinese Journal of Computers, Vol. 28, No. 4. pp.615- 626.

IEEE COMPUTER SOCIETY