Application of Services Relation Tracing to Automated Web Service Composition

Qixuan Liang, Shizhan Chen* and Zhiyong Feng

School of Computer Science and Technology, Tianjin University, Tianjin 300072, China

Published online: 1 Feb. 2013

Abstract: Web service composition has received much interest due to its potential to achieve the visionary promise of Service-Oriented Computing paradigm in which services can be utilized as the basic constructs to create flexible dynamic business processes and agile applications that may span organizations and computing platforms with little effort even automatically. In this paper, we proposed a snowball method for automated Web service composition which is based on service relation tracing which starts off rather like a standard ego-centered network study. It discoveries one or several available Web services which can accept the input parameters in user requirement through semantic reasoning and determines whether its output parameters meet the user requirements. And then it fetches other candidate services following the Sequential-total and Sequential-partial relations, investigates their outputs, also asks further sequential relations, and so forth through a succession of "waves" of spreading. So the process is repeated and the composite service is constructed incrementally tracing sequential relations in Service Network just like a snowball. The experiment shows that the approach proposed in this paper can greatly improve the accuracy and efficiency of Web service composition.

Keywords: Web service composition, service relation tracing, Services-Oriented Computing, Web service

1. Introduction

Recently, Service-Oriented Computing (SOC)[1,2] has received much interest due to its potential to tackle many adaptive system architecture issues that were previously hard to overcome by other computing paradigms. It utilizes functionalities provided by business applications and encapsulated within services, to build low-cost and flexible dynamic business processes or service-aware application[3] within and across organizational boundaries and computing platforms, even in highly distributed, open and heterogeneous Internet environments.

With the characteristics of loose coupling, cross platform, self-describing and modularization and constructed on the basis of publishing, discovery and interaction, Web service[4] has become the best practice of realizing SOC, turning Software as a Service (SaaS)[5], and developing business process across organization boundary[6], and has drawn much attention. When no single Web service can satisfy the users request in functionality, there should be a possibility to combine existing services together to satisfy the request. That is Web Service Composition (WSC). Logically, composite service (result of WSC) can be considered as a services chain consisting of series of relevant Web services. Each component service on the services chain, controlled by control flow, receives output parameters of its immediate previous service (except initial service, in which, users request is directly input), and then output its result to its immediate following service. Thus repeatedly, till the component service at end of services chain outputs users expected messages.

Compared with static composition method provided by WS-BPEL and WS-CDL, automated Web service composition is the key of realizing the visionary promise of Service-Oriented Computing. The reason is that Web services can be created and updated on the fly and it may be beyond humans capability to analyze the required services and compose them manually. [7] improves BPEL to make automatically composition. Therefore, on the road of realizing Service-Oriented Computing, in addition to the challenge of composition automation level, other challenges are the efficiency of composition method and the flexibility of composition proposal.

Tianjin Key Laboratory of Cognitive Computing and Application “to Author Affiliation in a new line before”

* Corresponding author e-mail: shizhan@tju.edu.cn
School of Computer Science and Technology. Web service composition based on ontology and Semantic Web [8] technology is a very active hotspot of research in SOC. There are many activities and much effect mainly concentrated on the semantic model for services discovery and dynamic service composition methods applying ontology and AI planning techniques. In this field, there are various solutions, from recapitulative abstract method to framework that has potential of becoming industrial standard. And major research efforts in this area have summarized, classified and evaluated in [9–14] from several perspective. However, there are still many open research questions and challenges in how to accurately identify the services on demand and aggregate them into a complex business process (or composite services) automatically [15–17].

Some research thought that Web service composition is similar to work flow, thus directed graph [18] or Petri net [19] representing various basic service and execution sequence in the flow can be used to construct model of composite service; or Web service can be abstracted into an automaton that receives events via inputting arrays and decides how to respond (in form of outputting events) [20]. Other research considered the automated composition of Web services as planning issue of artificial intelligence (AI): given clients request and a set of component services (atomic services or composite service), find out a composition that can satisfy the target, and apply AI based planning, reasoning and other techniques to automatically Web service composition [21–24], validation [22, 25] and monitoring in run-time [26]. Aiming at the situation that Web service based WSDL is lacking of semantic information impedes the dynamic and adaptive interaction among services, the research focus of Semantic Web Services (SWS) [27, 28] under the background of Semantic Web is to depict services using Ontology technique (OWL-S, WSMO (Web Service Modeling Ontology), WSMO-Lite, SAWSDL, METEOR-S etc.), or to expand semantics of current solutions, thus web services can be understood consistently by different machines, thus promote automated service discovering [29–32], composition and interaction [33, 34].

Relevance among Web services and its positive effect on service computing process has also been attracting some attention. An early contribution to the Web services relation was that by Zhang et al [35, 36], in which the authors developed a business relationship description language, WSRL, to facilitate services discovery and integration. However, they only focused on static relevance information among services, and paid insufficient attention to dynamic relation and Quality of service (QoS) information. Barros et al [37] presented a collection of service interaction patterns, which allows emerging Web services functionality, especially that pertaining to choreography and orchestration, to be benchmarked against abstracted forms of representative scenarios from a business perspective. Combining social network and Semantic Web technique, Chen et al [38, 39] focused on the dependence, relevance, interaction and competition among Web services caused by attribute and capability, etc., and defined explicitly service relation of the three levels, parameters, operations and services, designed and constructed a service computing infrastructure, named Service Network (SN), to introduce service relation into service discovery, composition and management. On the basis of SN, Guo et al [40] realized a service composition method based on service relation and shortest path, thus Web services composition is remarkably improved in accuracy and efficiency. However, the composition method only considers single chain in composition; it cannot solve the complex circumstances such as branches and multiple merge, etc., in composite structure.

In this paper, we describe a snowball composition method for automated Web service composition which is based on SN and service relation tracing and reasoning. The composition process in our case is divided into two phases. First, a description of a possible solution (namely abstract service composition) fulfilled the user requirements, is constructed incrementally just like a snowball. Second, the non-functionality attributes are used to determine which concrete services would be bounding to the abstract business process to generate the composite service. This paper is focusing on the first step of the composition process. We assume that WSDL presentation of Web services has been fabricated SN by semantic annotation and service relation mining [41].

The paper made the following contribution technically.

a) On the basis of previous work [38–40], complete and partial subsequent relation is distinguished to realize automatically the modes of parallel split and synchronization in composition.

b) Through service relation tracing and reasoning, snowball composition method is designed to improve composite efficiency.

c) The process of Web service composition is divided into two stages, abstract service composition based on relatively stable function construction and dynamic service binding considering non-functional factor, thus flexibility of composition is improved.

d) Test on prototype system shows, for rather complex test scene, the response time of snowball composition based service relation tracing is controlled within seconds, and meanwhile it has better response capability.

The rest of this paper is organized as follows: Section 2 presents a motivating example and the reason behind it, Section 3 describes the details of snowball composition algorithm based on service relation tracing, Section 4 is involved in experiential study and discuss, and with a conclusion in Section 5 we will finish this paper.
2. Motivating Example

The complete automation of WSC is still a hard problem and requires further research. One reason is that composite structure is complex net/graph instead of simple chain. For example, a tourist plans to enjoy vocation at some tourism site. After successful booking of air ticket, he wants some information of destination, like weather, exchange, airport or railway station, etc. A solution and the required services is represented in Figure 2.1. It is important to mention that the flow mode is more than a simple chain of services. In particular, it is a complex graph which consists of ten standalone Web services, which can be developed by different organizations and described in different semantic models, and two workflow patterns, which are Parallel Split pattern and Synchronization pattern in BPEL4WS[42]. The interaction among Web services originates from the fact that in a composition, services can interact with each other to exchange parameters, for example a service’s output parameter could be another service’s input parameter. Service relation is constructed on the basis of parameter semantics and represents the semantic relevance among parameters of Web services. In WSDL archive, a complete parameter may be a simple type or a composite type with many simple types or composed by complex types. Naturally, there is difference in semantics matching level among parameters. For example, subsequent relation can be further divided into two types, total subsequent relation and partial subsequent relation. Sequential-total WS1 output parameter completely includes WS2 input parameter, e.g., WS1 may drive WS2 independently, and no service is required between the two services to realize functional interaction, as the relation between WS630 and WS640 shown in Figure 2.1. Sequential-partial WS1 output parameter is included in WS2 input parameter, e.g., WS1 can only provide partial parameters WS2 input requires, and cooperation with other service output parameter is relied on to drive WS2 to realize interaction. For example, WS3 and WS18 together can drive WS628. Figure 2.2 shows the four forms of composition flow constructed by total subsequent relation and partial subsequent relation. Service chain (Figure 2.2-a) depicted the simplest services composition form. That is, multiple Web services connected between previous and subsequent services form single chain structure. Service tree depicted a combination form of tree type. If each service in service tree only has one previous service, named positive service tree (Figure 2.2-b). If each service in service tree has only one subsequent service, named reverse service tree (Figure 2.2-c). Service graph (Figure 2.2-d) describes the most
complex combination forms among services composition, and each service may satisfy many subsequent services, and meanwhile, there are many previous services satisfying such service at the same time. In particular, simple service chain is rare, and more composition is a rooted-tree, a leafed-tree, or a complex graph.

3. Two-phase Service Composition Framework

With the primary purpose of snowball composition based on service relation tracing is to construct a possible solution to meet the given goal, our method contains two stages, abstract service composition and specific service binding.

The former stage only considers the semantic relevance between clients function demand and services, and it’s also the stage to specially embodying the characteristics of snowball composition based on service relation tracing. The result of abstract services composition is an abstract business process that can satisfy clients demand functionally, in which, each abstract service represents a function, and all specific services relevant with the abstract service can realize the function Abstract business process is a key factor affecting Web service composition.

The responsibility of the later stage is selecting a specific service for each abstract service in the abstract business process according to QoS-aware selection as executor of the function.

3.1. Snowball composition based on service relation tracing

Based on semantic relevance among Web services, we define a service collaboration graph (SCG). As a connection sub-graph of SN, each node in SCG expresses an abstract service, and the directed edge between nodes means a service can be totally or partially driven by another service, e.g., there is possibility of combination. Figure 3.1 gives an example of SCG. Such abstract service composition thus is developed into the problem of how to efficiently, accurately and incrementally construct SCG corresponding to clients demand from SN utilizing, as well as search one or many path(s) connecting client’s input parameters and output parameters from it.

The typical snowball composition based on service relation tracing starts off rather like a standard ego-centered network study. First, finding one or several available Web services through the input parameters in user requirement to initialize the SCG and determining whether its output parameters meet the user requirements. Then, upon the Sequential-total and Sequential-partial relationships, we find other candidates of the target with whom they are sequenced. Then we update the SCG using these sequences and investigate their output parameters, and also ask further sequential relations, and so forth through a succession of ”waves” of spreading. The process is repeated with any whose outputs mismatch, tracing their sequential relations as well, and so forth, until the requirements are matched or all leads have been exhausted. Its schematic diagram is as shown in Figure 3.2. In the whole searching and composition process, similar with LevelOrderTraverse, subsequent service of current service is stepwise added into SCG. Meanwhile, with level by level searching, SCG and parameters reviewed are become more gradually, just like snowballing. Detailed Algorithm is as follows:

Algorithm 1 Snowball composition based on SR tracing.

Input: the array of request input items by user (requestInputList), the array of request output items (requestOutputList)

Output: roadList

BEGIN: //Initializing module. parameter_5 is a hash table storing parameter and all abstract services
whose output semantically match the parameter in \(<\text{Parameter},\text{List} >\text{of AbstractService}>\) form. 
part_meet_list, a list of abstract services storing those services whose input cannot be provided fully. The road and roadList are used to store solution. All of them are initialized as NULL.
1: Parameter AS Map < Parameter, List > AbstractService > parameter as;
2: Part_Meet AS List part_meet_list;
3: Road road;
4: List< Road > roadList;
5: Queue< AbstractService > queue ; // SCG in queue from Creating virtual start node on the basis of requestInputList which has been extended through semantic reasoning, then it includes all parameters in SN whose semantics is less than the user input parameters.
6: AbstractService start;
7: Start.outputList = SemanticReasoning( requestInputList );
8: queue.push(start); //Initializing SCG with the virtual start node. Loop until queue is empty or parameter_as includes all the user request outputs (the given goal is satisfied!)
9: Snowball: While(queue is not empty and not parameter_as.containAllKey(requestOutputList)){
10: currentService = queue.pop();
11: // dealing with the totally sequential service
12: List< AbstractService > totalSequentialServiceList = currentService.totalSequentialServiceList
13: for AbstractService si in totalSequentialServiceList 
14: for Parameter pi in si.outputList do
15: parameter_as.put\(<\pi, si >\) 
16: end for
17: Si.FatherList = currentService
18: queue.push(Si)
19: end for
20: // dealing with the partially sequential service
21: List< AbstractService > partlSequentialServiceList = getPartSequentialServiceListByHighMatchDegree();
22: part_meet_list.addAll(partlSequentialServiceList);
23: for AbstractService si in part_meet_list do
24: if parameter_as.containAllKey(si.inputList) then
25: for Parameter pi in si.outputList do
26: parameter_as.<\pi, si > 
27: end for
28: si.FatherList = parameter_as.get(pi)
29: queue.push(si)
30: end if
31: end for//Building a solution through traversing the output parameters required with the FatherList pointer. When there are several possible solutions, our method only chooses the one who is composed of the fewest nodes (The complexity will be relatively small.).
32: for Parameter pi in requestOutputList do
33: AbstractService si = getNodeByOutputParameter (parameter_as, pi)
34: while si is not NULL do
35: road.add(si)
36: si = si.FatherList.get(0)
37: end while
38: roadList.add(road)
39: end for
40: END

When there is possibility of various abstract business processes, only solution with the least nodes (whose complex degree will also be less) is selected and returned to client. Pretty soon the process "snowball" and we have a possible solution to user requirement. Clearly this is a better way of finding a composition based on sequential service relation tracing than random-walk search in service repository, since each WS we had inspected must has potentially collaboration contacts to the former. Essentially, the snowball composition based on service relation tracing is a typical form of snowball sampling[43] applied to SOC.

3.2. Dynamic service binding

On the basis of abstract business process generated at the first stage, the second stage is to dynamically bind specific WS as the executor of the functional unit in abstract business process.

As the selection of the WS with multiple QoS attributes has been proved NP-hard[44]. Considering that in Internet environment the consumers are unlikely to care about all aspects of the QoS when they use Web services, we only choose three QoS properties (response time, successability and reputation) specified by the Web Services Quality Modeling to made simplicity of QoS-aware selection. Based on the QoS metrics preprocessing by normalization method, our method simply select the most adequate specific services from the set of concrete services who attach to an abstract one, which can be proved the algorithm complexity is O(n).

3.3. Optimization

In snowball composition based on service relation tracing the sample size grows exponentially with the number of relationship tracing waves while in random walk discovering the sample size grows only linearly. Measures taken in optimizing performance include:

a) When adding subsequent service nodes incrementally, total subsequent relation is considered first, and then followed by partial subsequent relation. Meanwhile, in partial subsequent relation, partial subsequent service with high parameters is considered first. That is, those with less missing parameters (can be driven easily) are considered with priority.
b) Hash table is used to store parameters. All abstract services that can output the parameter are guaranteed to quickly search and output all the abstract services of the parameter within $O(1)$ time complexity.

It can be proved the snowball algorithm complexity is $O(n^2)$

4. Experiment

4.1. Experimental design

SN system prototype is a typical B/S structure. All aggregating and reasoning work are carried out at server, including:

a) Server, the main component of the system, including MySQL 5 database storing actual Web services collected from Internet, service relations and QoS information, and several background processes such as, service annotation, service relation mining and QoS inspector, etc., as well as automatic service compositor based on service relation tracing. The server is a PC with Windows 7, equipped with Intel Q8200 CPU, 4GB RAM and 500GB SATA hard disk.

b) PC with Windows 7 as client, equipped with Intel Q8200 CPU, 4GB RMB and 500GB SATA hard disk, IE8, connected with server via 100 Mbps Ethernet. In addition to receiving users input and displaying composition result via IE, client does not need to undertake any composition or reasoning.

Experimental scene setting simulated information inquiry in Section 2, in order to validate that snowball composition based on service relation tracing is rational and efficient.

Test dataset: 300 actual Web services related with tourism are downloaded from Seekda, and added to SN system prototype by WSDL archives validation, service annotation and service relation mining, etc., including 1,026 abstract services and 49,321 service relationships. Response time is simulated with the response time of host that each Web service resides; success-ability and reputation are set as default value 1, to construct test dataset.

Test process: client submits initial input (including destination country and license key that drives the inquiry) and expected output information (like weather, airport, railway station and exchange rate of the destination) from clients machine; abstract business process and binding results satisfying clients demand are output and average time of five operation times is recorded.

4.2. Experimental result

In the above experimental scene, the abstract business process constructed with snowball composition based on service relation tracing and binding result, say, solution A, as shown in Figure. 2.1 in Section 2, takes 3,959 ms. Considering that the performance of the server we used in the experiment is rather low, the efficiency of our method is completely within users tolerance.

To further measure the effectiveness of the composition method, we get new result after adding one specific Web services, Weather, and corresponding three abstract services, GetWeatherInformation(1027), GetCityForecastByZIP(1028), and GetCityWeatherByZIP(1029) into test dataset on purpose (solution B) as shown in the Figure. 4.1. It takes 4,001 ms.

By comparison, solution A and solution B are consistent in the service quantity they involve, say, both are 10 services. And there is not much difference in time cost. However, solution A has more complex composite structure. Due to the abstract service, GetCityWeatherByZIP, to join, the service 628 in solution A was replaced by the service 1013, 544 by 1029, which makes two Synchronization pattern resulted by partial sequential relation disappear in solution B that can be considered as consisting of two standalone subprocesses. Comparatively speaking, each subprocess is simpler too.

5. Conclusion

The most important contribution of this paper is to propose a snowball method for parallel split and synchronization pattern in automated WSC issue to make the aggregation of relevant Web services efficient and on-demand. Efficient means it bases on service relation tracing mechanism and reduces search space by coordination relation mined off-line. On-demand means the composition is distinguished abstract business process based on functionality, which is stable relatively and can be reused in future, and dynamic specific service binding based on non-functionality which is variable in the reality.

Future work will focus on two aspects. One is to design proposal selection policy of more sophisticated and ingenious abstract composition, to select strong composite proposal with simple structure instead of only considering service quantity involved. The other is to analyze service web structure based on complex web theory and method, to guide the improvement of composite computing and further increase automation level, efficiency and users experience of services composite.

Acknowledgement

The authors wish to thank Guangquan Xu for his valuable advice and the members of service computing group from Institute of Knowledge Science and Engineering, Tianjin University. This work was supported by the NSF of China under Grant No. 61173155, the 863 Program of China
under Grant No. 2007AA01Z130, the 985 Project of Tianjin University under Grant No. 06050110000, and the Innovation Foundation of Tianjin University under Grant No. 2010XG-0009. The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References


Qixuan Liang, born in 19- 88, is a master of Software Engineering at Tianjin University, China. His current research include service computing and SOA.

Shizhan Chen, born in 19- 75, is a lecturer at Tianjin University, China. He obtained his PhD at Tianjin University in 2010 and his current research interests include service computing and SOA.
Zhiyong Feng, born in 1965, PhD, is a full tenured professor and head of the School of Computer Science and Technology, Tianjin University, China. His current research interests include knowledge engineering, service computing, and security software engineering.