Web Service Provider Selection under QoS and Security Constraints*

Li Zhou and Yuqing Sun†
School of Computer Science & Technology, Shandong University, P.R.China
yxzl_855@163.com; sun_yuqing@sdu.edu.cn

Abstract

In pervasive computing environment, web services are widely adopted in dynamic business processes. When selecting service providers, quality of service (QoS) and security constraints should be considered simultaneously. However, most current research works about provider selection either focus on quality of service, or just take into account security constraints. In this paper, we propose an evaluation model for business process, which considers the above two factors at the same time. Two key criterias are introduced to measure the overall cost and benefit of QoS parameters respectively. Typical security constraints, such as separation of duty (SoD) constraints and binding of duty (BoD) constraints, are also considered. We define the service provider selection problem and present a linear programming based solution. Some experiments are performed to evaluate our model and the results show that our approach is effective in solving the problem.

Keywords: Web service provider, Security constraints, Quality of service.

1. Introduction

The rapid development of pervasive computing technologies facilitate the dynamic collaboration among organizations [1], in which web based business process has become one important form. Generally, a business process consists of multiple activities and each of them can be implemented with a web service. The web service technologies allow a business process to be established at running time according to a dynamic objective by selecting and binding web services from different providers under certain constraints [2] [3].

We take the tourism services as an example. Tourism services are value-added composite services consisting of several sub-services, which provide multiple kinds of services to people, such as designing traveling routes, reserving tickets, booking hotel, purchasing insurance, etc. In web based application environment, there often exist more than one providers for the same service, but they may exhibit different QoS parameters, such as response time, price, reliability. On the other side, the organizations who use these services may propose different QoS requirements related to business processes. At the same time, the widely adopted security constraints, such as separation of duty (SoD) and binding of duty (BoD), are also required. SoD constraints indicate that some sensitive activities should be assigned to or executed by a certain number of performers. For example, a cheque requires to be signed by two different people. BoD constraints dictate that two activities must be executed by the same performer. So, when selecting web service providers for the activities of a business process, both the QoS parameters of providers and organizations' requirements should be considered.

Current research works on web service selection mainly focus on QoS to select the services. [4] presented quality-driven approach to select component services during the execution of a composite service. [5] proposed a lightweight approach for QoS–aware service composition using genetic algorithms for the optimal QoS estimation and presented an algorithm for early triggering service re-planning. Some literatures focus on the computing of QoS [6] [7] [8]. However, most of them did not consider the security constraints. In the area of security, security constraints are mostly specified and enforced in one organization without consideration of web services. The only work corresponding to this aspect is resiliency policy which requires an access control system to be resilient to the absence of users but can not select users under QoS [12] [13].

In this paper, we investigate the problem of selecting web service providers under both QoS and security constraints for a dynamic business process. We propose a business process evaluation model and

*Supported by the Science Foundation of Shandong Province under Grant No. Y2008G28
†Corresponding author: School of Computer Science & Technology, Shandong University. Jinan, Shandong, China. 250101

give the formal schema of dynamic business process. We define two new evaluation criterias to measure the QoS of business processes, i.e. $Cost_N$ and $Benefit_P$. On these bases, we define the service provider selection problem for dynamic business process and present a method based on linear programming (LP) to solve the problem. We perform some experiments to verify the effectiveness of our method and analyze the experiments results.

The remainder of this paper is organized as follows. In section 2, we present the business process evaluation model. In section 3, we give the method based on LP to solve the service provider selection problem. In section 4, the system architecture is presented. In section 5, we review related work. Finally, section 6 concludes this paper and identifies directions for future research.

2. The Business Process Evaluation Model

In this section, we formally define the dynamic business process, discuss the considered factors in the selection of service providers and then present our problem definition of service provider selection.

2.1. The Formal Schema of Business Process

Definition 1 (Business Process). A business process is represented as a tuple $<A,\leq>$, where A is a finite set of activities, and $\leq \subseteq A \times A$ defines a partial order on activities in A.

If there are m activities in A, each activity is indicated as a_i, $1 \leq i \leq m$, then $A=\{a_1,a_2,...,a_m\}$. In a business process, $a_1 \leq a_2$ indicates that activity a_1 must be performed before a_2. Activity a_1 and a_2 may be performed concurrently.

Definition 2 (Candidate Set for Activity). For a given activity a in a business process $<A,\leq>$, there may be a number of web service providers, we refer to these service providers as the candidate set for activity a, denoted by $S(a)$.

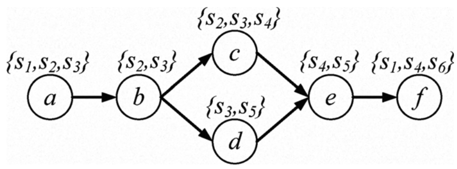


Fig.1. An example of business process

Fig.1 shows an example of business process. In this example, the business process is composed of six activities, namely $A=\{a, b, c, d, e, f\}$. The services labeled in the bracket above the activity are corresponding candidate set for activities. For example, the candidate set for activity a is $S(a)=\{s_1,s_2,s_3\}$.

2.2. Considered Factors

A service provider may provide services with different functions. For an activity of a business process, different candidate service providers may exhibit different QoS. For the evaluation of QoS, different perspectives can be measured, and the often used QoS parameters are as follows:

Response time(T): Time elapsed from the submission of a service request to the time the response is received.

Price(P): The amount of money paid by the service requestor for consuming the service.

Reliability(R): The probability that the service can be completed successfully.

Reputation(Rep): Used to describe the evaluation that the service requestors provided for the service they have invoked.

A composite service (Cs) is composed by a group of automic services (As). We express the QoS evaluation model for an atomic service As_i as a quadruple $Q_{As_i}(T_i,P_i,R_i,Rep_i)$. And the QoS evaluation model for a composite service Cs is also a quadruple $Q_{Cs}(T(Cs),P(Cs),R(Cs),Rep(Cs))$. The letters T, P, R and Rep denote the response time, price, reliability and reputation respectively. The QoS parameters' values for Cs can be computed through those of As_i.

Suppose a business process composing of n activities, which are executed in sequential order. As each activity is performed by an atomic service, the composite service for the business process consists n corresponding atomic services. Thus we get the QoS parameters' values of corresponding Cs as follows.

$$T(Cs)=\sum_{i=1}^{n}T_i \text{ , } P(Cs)=\sum_{i=1}^{n}P_i \text{ ,}$$

$$R(Cs)=\prod_{i=1}^{n}R_i \text{ , } Rep(Cs)=\sum_{i=1}^{n}Rep_i \Big/ n$$

For a business process which consists of n parallel activities, the QoS parameters' values of corresponding Cs are as follows.

$$T(Cs)=\mathrm{MAX}(T_1,T_2,\ldots,T_n) \text{ , } P(Cs)=\sum_{i=1}^{n}P_i \text{ ,}$$

$$R(Cs)=\prod_{i=1}^{n}R_i \text{ , } Rep(Cs)=\sum_{i=1}^{n}Rep_i \Big/ n$$

Besides the above QoS requirements, security constraints applied to the activities of the business process are the prerequisite for ensuring the execution of the business process.

Definition 3 (Security Constraints). Security constraints are represented as a tuple $<V,(a_1,a_2),\rho>$, where a_1 and a_2 are two activities in a business process, V represents the set of service providers that can perform activity a_1 and $V \subseteq S(a_1)$. ρ is the relationship on activity a_1 and a_2,

$\rho \in \{=, \neq\}$. $< V, (a_1, a_2), \rho >$ is satisfied if $x \in V$ performs a_1 and $y \in S(a_2)$ performs a_2, $(x, y) \in \rho$.

In fact, security constraints place some restrictions on the service providers that can perform activity a_1 and a_2. For instance, in the business process of Fig.1, $<V,(a,e),=>$ indicates that activity a and e must be performed by the same service provider. In other words, the service providers performing the two activities should satisfy BoD constraints. While $<V,(b,c),\neq>$ indicates that activity b and c must be performed by different service providers. That is to say, they must satisfy the SoD constraints.

2.3. Problem Definition

One fundamental problem in web service based business process is how to select the service providers that satisfy the constraints and requirements thus guarantee the execution of the business process.

Definition 4 (System Configuration). We define the system configuration as the union of candidate set for activities of the business process BP, expressed as $SC=S(a_1) \cup S(a_2) \cup ... \cup S(a_m)$, where BP refers to $<A, \leq>$, $A=\{a_1,a_2, \dots ,a_m\}$.

Definition 5 (Services Configuration for Business Process). For a given business process $<A,\leq>$, assuming that there are m activities in the business process, and each activity is identified by a_i, $1 \leq i \leq m$. We define the services configuration for the business process as $\{<a_1,s_{a1}>,<a_2,s_{a2}>,\dots,<a_m,s_{am}>\}$, where s_{ai} is the service provider for activity a_i.

For each activity in a business process, there may be a number of web service providers. So there may be several services configuration for a given business process. And selecting better service configuration to perform the activities can guarantee the execution of the whole business process.

In the often used QoS parameters, some parameters such as response time and price are negative, i.e., the higher the value is, the lower the quality of service is. While some parameters such as reliability and reputation are positive, i.e., the higher the value is, the higher the quality of service is.

In order to effectively distinguish these two kinds of QoS parameters, we define two new evaluation criterias to measure the QoS of business process.

We define $Cost_N$ to represent the overall cost of negative QoS parameters:

$$Cost_N = \sum w_i \cdot n_i \qquad (1)$$

the subscript N refers to Negative, w_i is the weight of QoS parameters which are negative, and $\sum w_i = 1$. n_i is the value of QoS parameters. w_i can be adjusted according to the organizations' preference on QoS.

We define $Benefit_P$ to represent the overall benefit of positive QoS parameters:

$$Benefit_P = \sum w_j \cdot p_j \qquad (2)$$

the subscript P refers to Positive, w_j is the weight of QoS parameters which are positive, and $\sum w_j = 1$. p_j is the value of QoS parameters. w_j can also be adjusted according to the organizations' preference on QoS.

Before computing the $Cost_N$ and $Benefit_P$, we use the method in [4] to normalize the values of the QoS parameters. The normalized values all satisfy the characteristic of positive QoS parameters, i.e., the higher the value is, the higher the quality of service is.

Then we can use the sum of of $Cost_N$ plus $Benefit_P$ to measure the overall QoS of the business process. To guarantee the QoS of the whole business process, we request the services configuration which has the maximum sum of $Cost_N$ plus $Benefit_P$ to perform the business process.

Definition 6 (Service Provider Selection Problem). The service provider selection problem for dynamic business processes takes the form

$SPSP<A, \leq, SC, Th\_Cost_N, Th\_Benefit_P, C>$

where A is the set of activities of a business process, $\leq$ indicates the relationship among the activities. SC is current system configuration. Th_Cost_N and $Th\_Benefit_P$ are two given thresholds and C is the set of security constraints applied to activities of the business process.

Given a business process $<A,\leq>$, two thresholds Th_Cost_N and $Th\_Benefit_P$ and C, in current system configuration SC, what we want is to find a group of service providers and the corresponding services configuration has the maximum sum of $Cost_N$ plus $Benefit_P$, and also $Cost_N$ and $Benefit_P$ are respectively greater than the threshold Th_Cost_N and $Th\_Benefit_P$ and the security constraints C are satisfied.

3. Linear Programming Based Solution

For the problem of selecting service providers for the business process, the most direct approach is the enumeration method. Assuming that there are n activities in a sequential-execution business process and there are m candidate service providers for each activity. Then the total number of web services configuration is m^n, and the time complexity of selecting an optimal services configuration is $O(m^n)$. If the activities number of a business process is large, and in web service application environment, there are a number of service providers which can provide the same service for an activity. So it is very costly to find all these web services configurations and select the best services configuration satisfying the security constraints and QoS requirements.

In this section, we present a method based on linear programming [9] which can be used to find an optimal services configuration without generating all the possible configurations.

Linear programming is a technique used to solve the problem of maximizing or minimizing a linear objective function subject to linear equality and inequality constraints. In a few words, linear programming provides the way to achieve the best outcome (such as maximum profit or lowest cost) in a given mathematical model and given some list of requirements represented as linear equations. Many practical problems in operations research can be expressed as linear programming problems.

3.1. Expression of QoS Parameters

In order to use LP to select the optimal services configuration, we model the selection of service providers as an LP problem. We should give the linear function. And the linear function should reflect the overall QoS of the business process.

Assume S_j is the set of candidate service providers for activity a_j. We use the variables x_{ij} to represent the candidate service provider s_{ij} in the selected services configuration. The value of each variable x_{ij} is 1 if service s_{ij} is in the selected configuration, 0 otherwise.

We define the objective function as

$$\text{Max } f=Cost_N+Benefit_P \quad (3)$$

Here we consider two negative QoS parameters: response time and price, and two positive QoS parameters: reliability and reputation. And we consider a sequential business process. Then $Cost_N$ in (3) can be represented as

$$Cost_N = w_t \cdot n_t + w_p \cdot n_p \quad (4)$$

where w_t is the weight assigned to the QoS parameter response time and w_p is the weight assigned to the parameter price. n_t is the total response time of composite service for the business process, and n_p is the total price of composite service for the business process.

Similarly, $Benefit_P$ can be represented as

$$Benefit_P = w_r \cdot p_r + w_{rep} \cdot p_{rep} \quad (5)$$

where w_r is the weight assigned to the QoS parameter reliability and w_{rep} is the weight assigned to the parameter reputation. p_r is the total reliability of composite service for the business process, p_{rep} is the total reputation of composite service for the business process.

For a sequential business process, n_t, n_p and p_r, p_{rep} can be easily expressed by the following equations.

Assuming that variable t_{ij} indicates the response time of service provided by the web service provider s_{ij}, A is the set of all activities of the business process. Then we have the total response time of composite service for the business process:

$$n_t = \sum_{j\in A}\sum_{i\in S_j} x_{ij} t_{ij} \quad (6)$$

Again, the total price of composite service for the business process is as follows:

$$n_p = \sum_{j\in A}\sum_{i\in S_j} x_{ij} p_{ij} \quad (7)$$

where the variable p_{ij} represents the price of service provided by the web service provider s_{ij}.

The total reputation of composite service for the business process is as follows:

$$p_{rep} = (\sum_{j\in A}\sum_{i\in S_j} x_{ij} rep_{ij}) / n \quad (8)$$

where the variable rep_{ij} indicates the reputation of service provided by the web service provider s_{ij}, and n is the number of the activities of the business process.

The representation of reliability for composite service is a little different and complex. Assuming the variable r_{ij} is the reliability of service provided by the web service provider s_{ij}. Based on the calculation method for the total reliability, we have

$$p_r = \prod_{j\in A}\sum_{i\in S_j} x_{ij} r_{ij} \quad (9)$$

From above we can see that the reliability's function is non-linear. Then we can linearize it using a logarithm function as shown below:

$$ln(p_r)=ln(\prod_{j\in A} R_j) = \sum_{j\in A} lnR_j \quad (10)$$

Let $p_r'=ln(p_r)$, $R_j'=ln(R_j)$, $r_{ij}'=ln(r_{ij})$, we get

$$p_r' = \sum_{j\in A} R_j' = \sum_{j\in A}\sum_{i\in S_j} x_{ij} r_{ij}' \quad (11)$$

Then we transform all QoS aggregative functions to linear ones.

3.2. Expression of Constraints

For the problem of selecting the service providers for business processes, there exist two kinds of threshold which are Th_Cost_N and $Th\_Benefit_P$, then the constraints can be expressed as the following inequations:

$$Cost_N \geq Th\_Cost_N \quad (12)$$

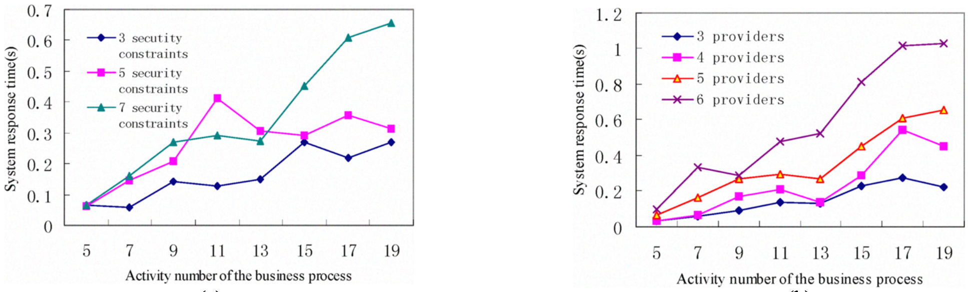
$$Benefit_P \geq Th\_Benefit_P \quad (13)$$

For each activity a_j, only one web service provider in S_j should be selected. Then we get the following constraint:

$$\sum_{i\in S_j} x_{ij} = 1, \forall j \in A \quad (14)$$

Besides above constraints, we have to also consider the security constraints defined in our research. For two activities a_l and a_k, if there exist SoD constraints, then we have

$$x_{il} + x_{jk} = 1, \forall i = j, i \in S_l, j \in S_k \quad (15)$$



(a) (b)

Fig.2. Experiments results

For two activities a_l and a_k, if there exist BoD constraints, then we have

$$x_{il} + x_{jk} = 2, \forall i = j, i \in S_l, j \in S_k \quad (16)$$

3.3. Experiments Results

We have performed some experiments to verify the effectiveness of our method on a PC with OS Windows XP SP3, RAM 2G. We program the experiments with the tool of Matlab 7.1. The parameters that we should consider are: the activity number of business process, the number of security constraints and candidate service providers for each activity. The number of activity varies from 5 to 20, the number of security constraints we used is 3, 5 and 7. The number of candidate service providers increases from 3 to 6.

The experiments results are reported in Fig.2(a) and (b). Fig.2(a) shows the system response time of different activity numbers with the number of security constraints 3, 5, 7 and the number of candidate service providers for each activity 5. Because different security constraints determine different solution space, the system response time fluctuates as the security constraints change. Fig.2(b) shows the system response time of different activity numbers with the number of candidate service providers 3, 4, 5, 6 and the security constraints 7. We can see that for a given activity number, the system response time increases with the increase of the number of candidate service providers.

It is observed that the computation time is quite acceptable.

4. System Architecture

In this section, we present the architecture of web based composition for business process. As shown in Fig.3, the architecture consists of six modules, namely Process Definition module, Service Providers Searching module, Services Composition module, Services Execution module, Services Execution Monitor module, and Services Information Management module.

The Process Definition module is responsible for transforming user requests to abstract business process definition, and security constraints are also designed according to the security requirements of the business process of organizations.

The Service Providers Searching module finds the candidate web service providers for each activity of the business process. And the candidate web service providers for each activity have the capability to perform the corresponding activity.

In order to guarantee the security and overall QoS requirements of the business process, the Service Composition module selects the best services configuration for the business process. In the composition process, security constraints and QoS are considered as two important factors for selecting better services configuration.

Once the Services Composition module sends the result of services configuration to the Services Execution module, the execution module will invoke corresponding web services to execute the business process. And the execution results will return to the users. Meanwhile, the Services Execution Monitor module will monitor the execution process at every step and report the QoS information such as response time to the Services Information Management module which can update corresponding QoS parameters' values. Besides, the Services Information Management module can also keep the QoS information for services and add information for new registered web services. The Service Repository stores the services along with their QoS information and classifies the services according to their functions.

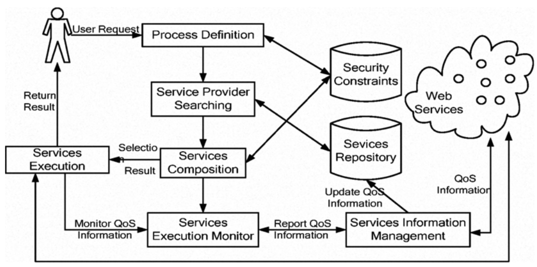


Fig.3. System architecture

5. Related Work

Many studies focus on the dynamic web service selection to guarantee the reliability of composite service for the business process. S.Y. Hwang et al.[10] study the dynamic web service selection problem in a failure-prone environment. They use the FSM(finite state machine) to model the permitted web services operations invocation sequences and propose two strategies to select services. The strategies select web services utilizing the aggregated reliability. The proposed approaches can maximize the chance to successful execute the composite service. But they only take the reliability into account and also do not consider the security constraints.

Recent research on resiliency policies in access control requires an access control system to be resilient to the absence of users [11]. Q.Wang et al.[12] discuss the authorization constraints and the satisfiability problem in workflows. They also define three levels of resiliency in workflow systems. In addition, F.Paci et al.[13] investigate the problem of WS-BPEL process resiliency in RBAC-WS-BPEL, and a new constraint—resiliency constraint is proposed to guarantee even some users become unavailable, the remaining users can still complete the execution of the process. In a word, the goal of research on the resiliency policies is to guarantee the execution of the business process. Although they consider security constraints in the execution of process, they can not select users according to the absence probability.

On the whole, almost all previous researches on web service composition or access control system only consider one factor that influences the execution of business process. The biggest difference between of our research and previous work is that we consider both the security constraints and QoS requirements.

6. Conclusions and Future Work

In this paper, we have studied the problem of web service provider selection in dynamic business process. In the proposed business process evaluation model, we present the formal schema of dynamic business process and two factors that influence the execution of business process, namely the security constraints and QoS requirements. We define two new evaluation criterias $Cost_N$ and $Benefit_P$ to measure the overall cost of negative QoS parameters and the overall benefit of positive QoS parameters. We take SoD and BoD constraints into account in the aspect of security constraints. We define the service provider selection problem for dynamic business process and give the solution based on LP. And the experiments show that our method is quite effective.

Our approach mainly investigated the service provider selection problem in sequential business process. In practice, the business process may be complex and constituted by paralleled, sequential and switched activities. As part of our future work, we plan to study the service providers selection problem in complex business process.

References

[1] D.Saha and A.Mukherjee, "Pervasive computing: a paradigm for the 21st Century," Computer, 2003, vol.36, no.3, pp.25–31.

[2] K.Yue, X.Wang, A.Zhou, "Underlying techniques for web services: A survey," Journal of Software.2004, vol.15, no.3, pp.428-442.

[3] G.Alonso, H.Kuno, F.Casati, and V. Machiraju, "Web services: concepts, architectures and applications," Springer, 2004.

[4] L. Zeng, B. Benatallah, M. Dumas, J. Kalagnanam, and Q.Z. Sheng, "Quality driven web services composition," Proc.12th Int'l Conf.World Wide Web (WWW '03), 2003, pp. 411-421.

[5] G.Canfora, M.Di.Penta, R.Esposito, M.L.Villani, "A lightweight approach for QoS-aware service composition," In Proc.2nd International Conference on Service Oriented Computing (ICSOC'04).

[6] T.Yu, Y.Zhang, and K.J.Lin, "Efficient algorithms for web services selection with end-to-end QoS Constraints," ACM Transactions on the Web(TWEB).2007,vol. l, no.1.

[7] V.Grassi and S.Patella, "Reliability prediction for service-oriented computing environments," IEEE Internet Computing, 2006,10(3), pp.43-49.

[8] S.-Y.Hwang, H.Wang, J.Srivastava, R.A.Paul, "A probabilistic qos model and computation framework for web services-based workflows," In Proceedings of 23rd International Conference on Conceptual Modeling,2004, pp.596-609.

[9] H.Karloff. Linear programming. Birkhauser, 1991.

[10] S.Y.Hwang; E.P.Lim, C.H.Lee, C.H.Chen, "Dynamic web service selection for reliable web service composition," Services Computing, 2008, vol.1, no.2, pp.104-116.

[11] N. Li, M.V.Tripunitara, and Q.Wang," "Resiliency policies in access control," In Proc. ACM Conference on Computer and Communications Security (CCS).New York: ACM TISSEC, Nov. 2006, pp.113-123.

[12] Q.Wang, N.Li, "Satisfiability and resiliency in workflow systems," ESORICS 2007, pp.90-105.

[13] F.Paci, R.Ferrini, Y.Sun, E.Bertino, "Authorization and user failure resiliency for WS-BPEL Business Processes," Proceedings of the 6th International Conference on Service-Oriented Computing, 2008, pp.116-131.