Flexible structured QoS-aware service composition based on a general component service model

Abstract: The current paper address the drawbacks of traditional QoS-aware service composition on structural and performance limitations. The contribution of this paper is to: 1) proposed a general structural and QoS computing model for component service which allows flexible construction. 2) formulated a more general QoS-aware service composition problem based on dynamic constructed component services by relaxing the one-to-one mapping between subtasks and atomic services. 3) provided a solution based on SMT solver.
Keywords: service composition, flexible service construction, quality of service (QoS), SMT

Introduction

Service composition is one of the key technologies to build service oriented architecture (SOA) systems or microservice based systems. In the domain of traditional service oriented computing (SOC), a service refers to an Web service. As the concepts of cloud computing, cloud manufacturing, big data were proposed and drew wide-spread attention, increasing kinds of capabilities or resources are regarded as services, for example, software as a service (SaaS), manufacturing resource as a service, knowledge as a service, etc.

Since a service or microservice is usually for a relatively simple and specific purpose (to meet the requirement of "high cohesion and loose coupling" in system architecture), accordingly it is designed and developed as a relatively independent and single-functionality component, which we call it a **atomic service** (it is also a kind of **component service** of the smallest granularity). In order to fulfill a muti-functionality task, which is more common in real usage circumstances, several atomic services should be integrated together to meet the muti-functionality requirements. And this is the so-called **service composition**.

- Why and what is QoS-aware service composition

As a muti-functionality task arrives, the typical process of service composition to response to the task is includes: (1)

Traditional problem of QoS-aware service composition
- Drawbacks of the traditional problem
- New idea of flexible structured service composition

Related works

TODO

Formulation of the problem

We improve the model of QoS-aware service composition by using synergistic elementary service groups (*SESGs*) as the elementary services, instead of using atomic services directly(as shown in Fig.1).

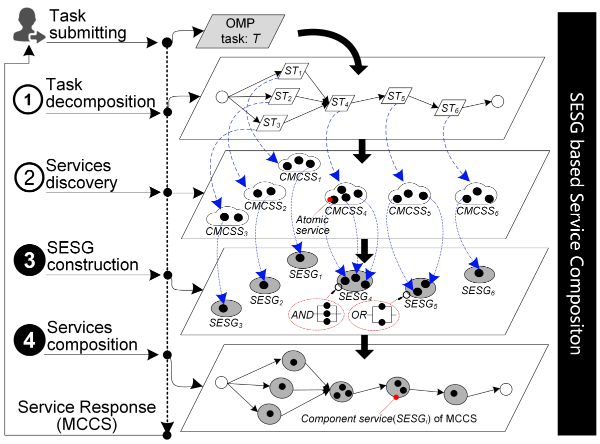

Figure 1: SESGs based service composition

A *SESG* is a general elementary service for flexible structured QoS-aware service composition. As shown in Fig.2, each *SESG* contains two parts: 1) the parallel structure, and 2) the selective structure. These two parts are both dynamically constructed with the runtime selected atomic services, and the two parts are bound together in a parallel mode.

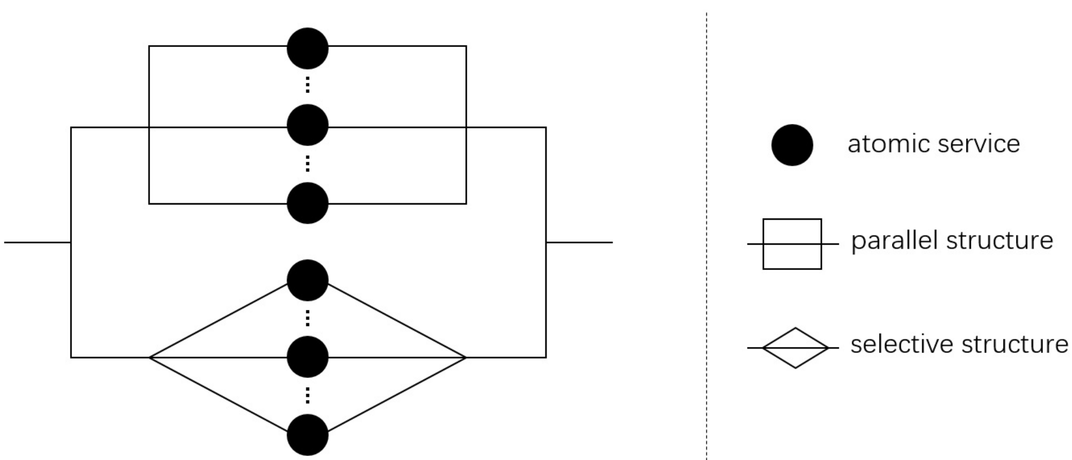

Figure 2: a general elementary service with runtime binding atomic services

Let s_{ij} denotes the j-th candidate atomic service that could be selected into the $SESG_i$ for a subtask ST_i , where $i \in [1, I]$, $j \in [1, J]$; the constant I and J denote the numbers of subtasks and atomic services, respectively.

Let $x_{ij} = 1$ denotes the j-th atomic service is chosen into the parallel part of $SESG_i$, otherwise, $x_{ij} = 0$; Let $y_{ij} = 1$ denotes the j-th atomic service is chosen into the selective part of $SESG_i$, otherwise, $y_{ij} = 0$; such that:

$$\neg(x_{ij} = 1 \wedge y_{ij} = 1) \tag{1}$$

$$\sum_{j=1}^{J} x_{ij} + \sum_{j=1}^{J} y_{ij} \geq 1 \tag{2}$$

Eq.1 indicates one atomic service can be chosen into either the parallel or selective part of $SESG_i$, or not to be chosen; the case that one atomic service is selected into both two parts is not allowed. Eq.2 indicates that at least one atomic service will be chosen into either the parallel or selective part of $SESG_i$.

Let $T(\cdot)$, $C(\cdot)$, $R(\cdot)$ denotes the time, cost, and reliability property of QoS, respectively. Then, the formulas to calculate the QoS of $SESG_i$ are as follows:

$$C(SESG_i) = \underbrace{\sum_{j=1}^{J} C(s_{ij}) \times x_{ij}}_{cost\ of\ the\ parallel\ part} + \underbrace{\sum_{j=1}^{J} C(s_{ij}) \times p_{ij}}_{cost\ of\ the\ selective\ part} \tag{3}$$

$$T(SESG_i) = \frac{1}{\underbrace{\sum_{j=1}^{J} \frac{1}{T(s_{ij})} \times x_{ij}}_{work\ efficiency\ of\ the\ parallel\ part} + \underbrace{\sum_{j=1}^{J} \frac{1}{T(s_{ij})} \times p_{ij}}_{work\ efficiency\ of\ the\ selective\ part}} \tag{4}$$

$$R(SESG_i) = \underbrace{(\prod_{j=1}^{J} R(s_{ij})^{x_{ij}})^{sgn(\sum_{j=1}^{J} x_{ij})}}_{reliability\ of\ the\ parallel\ part} \times \underbrace{R(sel(SESG_i))}_{reliability\ of\ the\ selective\ part} \tag{5}$$

where the reliability of the selective part of $SESG_i$ can be defined as follows:

$$\begin{aligned} &((\sum_{j=1}^{J} y_{ij} > 0) \Rightarrow (R(sel(SESG_i) = \sum_{j=1}^{J} R(s_{ij})^{y_{ij}} \times p_{ij})) \quad \wedge \\ &((\sum_{j=1}^{J} y_{ij} = 0) \Rightarrow (R(sel(SESG_i)) = 1)) \end{aligned} \tag{6}$$

The variable p_{ij} in Eq.4 - 6 denotes the probability of the selected atomic service s_{ij} within the selective part of $SESG_i$ to undertake the subtask ST_i. We assume that the probability of s_{ij} is related to the reliability of s_{ij}; because in many real cases, the service of higher reliability is more likely to get the chance to undertake a task, when a group of services compete with each other in a selective mode. On account of this assumption, we express p_{ij} as follows:

$$((\sum_{j=1}^{J} y_{ij} \geq 1) \Rightarrow (p_{ij} = \frac{R(s_{ij}) \times y_{ij}}{\sum_{j=1}^{J} R(s_{ij}) \times y_{ij}})) \wedge ((\sum_{j=1}^{J} y_{ij} < 1) \Rightarrow (p_{ij} = 0)) \tag{7}$$

Eq.7 indicates that if at least one atomic service is chosen into the selective part of $SESG_i$, the probability of s_{ij} can be calculated according to the proportion of its reliability in the accumulated reliability of all chosen atomic services; otherwise, the probability equals zero.

Till now, the formulas of the time, cost and reliability property of QoS have been defined for a general elementary component service $SESG_i$ in *SESG*s based service composition. By using the aggregation formulas in traditional QoS-aware service composition (as shown in Fig.3), the QoS evaluation of the overall workflow of OMP can be performed.

Sequence pattern

if in_{seq} of the sequence pattern is constructed with the set of $\{\text{SESG}_{i_1}, \ldots, \text{SESG}_{i_n}\}$

$$Q(in_{\text{seq}}) = \begin{cases} q_1(\text{in}_{\text{seq}}) = \sum_{i=i_1}^{i_n} q_1(\text{SESG}_i) \\ q_2(\text{in}_{\text{seq}}) = \sum_{i=i_1}^{i_n} q_2(\text{SESG}_i) \\ q_3(\text{in}_{\text{seq}}) = \prod_{i=i_1}^{i_n} q_3(\text{SESG}_i) \end{cases}$$

Parallel pattern

if in_{par} of the parallel pattern is constructed with the set of $\{\text{SESG}_{i_1}, \ldots, \text{SESG}_{i_n}\}$

$$Q(\text{in}_{\text{par}}) = \begin{cases} q_1(\text{in}_{\text{par}}) = \max_{i=i_1 \ldots i_n} (q_1(\text{SESG}_i)) \\ q_2(\text{in}_{\text{par}}) = \sum_{i=i_1}^{i_n} q_2(\text{SESG}_i) \\ q_3(\text{in}_{\text{par}}) = \prod_{i=i_1}^{i_n} q_3(\text{SESG}_i) \end{cases}$$

Selective pattern

if in_{sel} of the selective pattern is constructed with the set of $\{\text{SESG}_{i_1}, \ldots, \text{SESG}_{i_n}\}$

Let λ_i denote the probability of SESG_i being selected, s.t. $\sum_{i=i_1}^{i_n} \lambda_i = 1$

$$Q(\text{in}_{\text{sel}}) = \begin{cases} q_1(\text{in}_{\text{sel}}) = \sum_{i=i_1}^{i_n} q_1(\text{SESG}_i) \times \lambda_i \\ q_2(\text{in}_{\text{sel}}) = \sum_{i=i_1}^{i_n} q_2(\text{SESG}_i) \times \lambda_i \\ q_3(\text{in}_{\text{sel}}) = \sum_{i=i_1}^{i_n} q_3(\text{SESG}_i) \times \lambda_i \end{cases}$$

Circular pattern

if in_{cir} of circular pattern is constructed with the set of $\{\text{SESG}_{i_1}, \ldots, \text{SESG}_{i_n}\}$

Let ω denote the circle number, then

$$Q(\text{in}_{\text{cir}}) = \begin{cases} q_1(\text{in}_{\text{cir}}) = \omega \times \left(\sum_{i=i_1}^{i_n} q_1(\text{SESG}_i)\right) \\ q_2(\text{in}_{\text{cir}}) = \omega \times \left(\sum_{i=i_1}^{i_n} q_2(\text{SESG}_i)\right) \\ q_3(\text{in}_{\text{cir}}) = \prod_{i=i_1}^{i_n} q_3(\text{SESG}_i) \end{cases}$$

Figure 3: Aggregation formulas to evaluate workflow fragments of different patterns

As shown in Fig. **??**, the workflow of OMP involves six subtasks. The first three subtasks are parallel, and the last three are sequential. Since the QoS properties of time, cost and reliability for each atomic service are given, the QoS value of a general elementary service $SESG_i$ for a given subtask ST_i can be calculated based on Eq. 3 - 7. Then, the QoS value of a possible composite service (*CS*) for the OMP task can be calculated as follows:

$$T(CS) = max(T(SESG_1), T(SESG_2), T(SESG_3)) + \sum_{i=4}^{6} T(SESG_i) \tag{8a}$$

$$C(CS) = \sum_{i=1}^{6} C(SESG_i) \tag{8b}$$

$$R(CS) = \prod_{i=1}^{6} R(SESG_i) \tag{8c}$$

A simple additive weighting (SAW) technique is used to deal with the different magnitudes of QoS properties and obtain a score from diverse QoS dimensions.Positive and negative properties are scaled in different ways, as defined in Eqs. and , respectively.

$$\begin{cases} (max(T(\cdot)) = min(T(\cdot)) \Rightarrow \overline{T}(\cdot) = 1) \wedge \\ (max(T(\cdot)) \neq min(T(\cdot)) \Rightarrow \overline{T}(\cdot) = \dfrac{max(T(\cdot)) - T(\cdot)}{max(T(\cdot)) - min(T(\cdot))})) \end{cases} \tag{9a}$$

$$\begin{cases} (max(C(\cdot)) = min(C(\cdot)) \Rightarrow \overline{C}(\cdot) = 1) \wedge \\ (max(C(\cdot)) \neq min(C(\cdot)) \Rightarrow \overline{C}(\cdot) = \dfrac{max(C(\cdot)) - C(\cdot)}{max(C(\cdot)) - min(C(\cdot))})) \end{cases} \tag{9b}$$

$$\begin{cases} (max(R(\cdot)) = min(R(\cdot)) \Rightarrow \overline{R}(\cdot) = 1) \wedge \\ (max(R(\cdot)) \neq min(R(\cdot)) \Rightarrow \overline{R}(\cdot) = \dfrac{C(\cdot) - min(C(\cdot))}{max(C(\cdot)) - min(C(\cdot))})) \end{cases} \tag{9c}$$