QoS-Aware Recovery Strategies for Service Composition in Dynamic Network

Lili Sun, Yang Yang, Zenggang Xiong, Zhenqiang Mi, Shigang Li, Xiaoyong Zhao

School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing City, China, E-mail: lili@cs.toronto.edu

Corresponding Author

Abstract

Compared to traditional static wired network, dynamic network has a higher probability of failure due to the node mobility, constantly changing topology, as well as limited bandwidth and energy. Although the recovery strategies for service composition in dynamic network have aroused widespread concerns, most of the existing recovery strategies utilize rollback mechanism of post service composition interruption, which not only delay the process of service composition recovery, but also increase the overhead due to constantly searching for new service path and duplication of new service components, and also do not provide any Quality of Service (QoS) assurance. In this paper, we introduce the service composition model for dynamic network and propose two novel recovery strategies based on Backup Service Replacement Strategy (BSRS): Cold Backup Service Replacement Strategy (CBSRS) and Hot Backup Service Replacement Strategy (HBSRS). Furthermore, on the premise of guaranteeing the functionality of recovered service composition, we take the impact of different factors on the recovery performance of service composition into consideration and provide QoS guarantee for the recovery strategies. Experiments are conducted to show that the proposed strategies significantly improve the performance of service composition and effectively guarantee the availability and reliability of service composition in dynamic network.

Keywords: Dynamic Network, Service Composition, Service interruption, Service Recovery

1. Introduction

Service composition [1], as a basic problem of Service-oriented computing (SOC), enables standard, loosely-coupled and transparent implementation of shared service in the network. With the rapid development of wireless technology and mobile devices, the importance of service composition in dynamic network environment (DNE) [2] has been gradually realized. Compared to traditional static wired network, dynamic network has a higher probability of failure, both in mobile nodes and communication links. These failures are mainly caused by the following factors: the instability of the communication links (e.g., mobile wireless networks) [3], the higher probability of node damage (e.g., harsh environment, the battlefield environment) [4], and node mobility (e.g., mobile network) [5]. On one hand, the failures of service composition have become an inevitable phenomenon in dynamic network. On the other hand, a lot of research for service composition in traditional network is no longer applicable in dynamic network. Therefore, it is of great importance to develop recovery strategies for service composition in dynamic network.

The outline for rest of the paper is as follows. In section 2, we discuss related work and the motivation of our work. Next, the service composition model for dynamic network is introduced in section 3, and then the nature and characteristics of service composition in dynamic network is further specified. In section 4, we propose two novel recovery strategies for service composition in dynamic network: CBSRS and HBSRS, and describe the algorithms and QoS analysis of the two strategies respectively. After that, the simulation studies are given in section 5. Finally we conclude and summarize the future research work in section 6.
2. Related Work

Compared to network-level recovery, service-level recovery strategy has a better validity and reliability and mainly includes the following two aspects: Re-sending the Service Request Strategy (RSRS) [6, 7] and Backup Service Replacement Strategy (BSRS) [8-10].

A resource-oriented service composition recovery strategy is proposed for service composition in dynamic network [6]. In order to reduce response time of services and the load of the network, a dynamic monitor based service recovery strategy (DMBSRS) is proposed based on the distributed architecture of the service composition and dynamic surveillance [7]. However, the strategies do not take the declaration of the network stability into consideration.

It is efficient to set up the backup services for the unstable/failed services and replace them. According to the frequency and duration of service interruptions, a fresh concept of service interruption index and a framework for service composition recovery based on this concept is proposed [8]. However, because more paths of service composition have been selected, the load of the network has been increased. Zhou et al [9] has proposed a recovery strategy for service composition use backup services, but the study lacks of QoS guarantee. Based on the “transaction support”, an active and opportunistic replacement algorithm is put forward according to the feature of the transaction of web service [10]. However, the research only focused on simple environment of service application (e.g., the failure of single task), and could not well satisfy the actual demand of service composition recovery in dynamic network. Although the recovery strategies for service composition in the dynamic network have aroused widespread concerned, research on this issue still faces many challenges:

(1) The lack of in-depth analysis for the root-causes of service composition interruption and effective detection mechanism.
(2) Rollback mechanism after service composition interruptions is not enough to fundamentally optimize quality of service for dynamic network.
(3) The lack of QoS insurance for most proposed recovery strategies for service composition.

In order to solve the problems mentioned above, we present two novel recovery strategies. On the premise of guaranteeing the functionality of recovered service composition, we take the impact of different factors on the recovery performance of service composition into consideration and provide QoS guarantee for the recovery strategies.

3. Service Composition Model for Dynamic Network

Specifically, the nature of service composition in dynamic network is to find the service path to meet users’ service requests through service discovery and dynamic integration of atomic services. A number of researches [11, 12, 13] have been conducted to address the problem of cross-level integration of the service layer and the network layer, and the proposed service composition model in dynamic network has been described in detail in our previous work [14].

In this paper, we adopt the following three QoS parameters in the service layer:

(1) Service time $ST(s, n_j)$: the time of a successful implementation of the service $(s, n_j)$.
(2) Service cost $SC(s, n_j)$: the cost of a successful implementation of the service $(s, n_j)$.
(3) Service availability $P(s, n_j)$: the available time of the service $(s, n_j)$ during the service lifetime $T$. Assuming $m$ is the number of service disruptions and $\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_m$ is the sequence of interruption durations. The service availability can be calculated as:

\[
P(s, n_j) = \frac{T - \sum_{i=1}^{m} \tilde{t}_i}{T}
\]
The failures of service composition could mainly attribute to the following two aspects. First, the increase of the network load results in the decline of the service’s end to end QoS in the service layer. Second, the interruptions of the wireless communication links in the network layer lead to the interruptions of the service links in the service layer. If the interruption of service composition occurs, then recovery is required immediately.

Figure 1 shows a process of service composition recovery. When the atomic service \( s_n \) failed to perform in the service process, a new service component \( s_n \) will be discovered and activated, and new service links \( P'_2 \) and \( P'_3 \) will be established to restore the service composition. Assuming the execution time \( ET(s_n) \) starts from sending the request of service \( s_n \) to receiving the result which includes the recovery time after the service disruptions during the execution of the service \( s_n \). The generalized execution time \( \tilde{T}(s_n) \) and the generalized execution cost \( \tilde{C}(s_n) \) of the atomic service \( s_n \) can be calculated as:

\[
\tilde{T}(s_n) = ET(s_n) + CT(s_n) \\
\tilde{C}(s_n) = EC(s_n) + CC(s_n)
\]

The connection time \( CT(s_n) \) and connection cost \( CC(s_n) \) of new service links \( P'_2 \) and \( P'_3 \) can be calculated as:

\[
CT(s_n) = CT_{P'_2}(s_n) + CT_{P'_3}(s_n) \\
CC(s_n) = CC_{P'_2}(s_n) + CC_{P'_3}(s_n)
\]

Where \( CT_{P'_2}(s_n) \) and \( CT_{P'_3}(s_n) \) are the connection time of the service link \( P'_2 \) and \( P'_3 \); \( CC_{P'_2}(s_n) \) and \( CC_{P'_3}(s_n) \) are the connection cost of the service link \( P'_2 \) and \( P'_3 \).

4. Recovery Strategies for Service Composition in Dynamic Network

4.1. Backup Service Replacement Strategy

Backup Service Replacement Strategy (BSRS) is a recovery strategy designed for handling service with a greater probability of failures in dynamic network. These failures are mainly caused by the mobility of the nodes, link failures, etc. BSRS is based on the assumption that the mobility of the nodes which leads to the failures of the service composition also brings new services at the same time.
4.2. Improvement Strategies and the QoS Analysis

According to the different execution processes of the service restoration, we propose two recovery strategies for service composition in dynamic network: Cold Backup Service Replacement Strategy (CBSRS) and Hot Backup Service Replacement Strategy (HBSRS).

4.2.1. Cold Backup Service Replacement Strategy (CBSRS) and QoS Analysis

In the process of service recovery, CBSRS is triggered upon the detection of the service failures. In CBSRS, the backup services will not be triggered when the service is available. The algorithm has been described in our previous work [16]. The generalized execution time and the generalized execution cost of the atomic service \( (s_i, n_j) \) can be calculated as:

\[
\tilde{T}_{CBSRS}(s_i, n_j) = ET_{CBSRS}(s_i, n_j) + CT_{CBSRS}(s_i, n_j)
\]

\[
\tilde{C}_{CBSRS}(s_i, n_j) = EC_{CBSRS}(s_i, n_j) + CC_{CBSRS}(s_i, n_j)
\]

Assuming \( k_i \) is the executive number of the backup services of the atomic service \( (s_i, n_j) \). The execution time \( ET_{CBSRS}(s_i, n_j) \) and execution cost \( EC_{CBSRS}(s_i, n_j) \) of the service \( (s_i, n_j) \) can be calculated as:

\[
ET_{CBSRS}(s_i, n_j) = ET'_{CBSRS}(s_i, n_j) + ET_{CBSRS}(s_i, n_j)
\]

\[
EC_{CBSRS}(s_i, n_j) = EC'_{CBSRS}(s_i, n_j) + EC_{CBSRS}(s_i, n_j)
\]

Where the time and the cost that has been executed by the atomic service \( (s_i, n_j) \) when the service disruption occurs can be calculated as:

\[
ET'_{CBSRS}(s_i, n_j) = P_i(s_i, n_j)ST(s_i, n_j)
\]

\[
EC'_{CBSRS}(s_i, n_j) = P_i(s_i, n_j)SC(s_i, n_j)
\]

The execution time and the execution cost of the backup service after the service disruptions can be calculated as:

\[
ET_{CBSRS}(s_i, n_j) = \sum_{i=1}^{k-1} P_i(s_i, n_j)ST(s_i, n_j) + P_i(s_i, n_j)ST(s_i, n_j)
\]

\[
EC_{CBSRS}(s_i, n_j) = \sum_{i=1}^{k-1} P_i(s_i, n_j)SC(s_i, n_j) + P_i(s_i, n_j)SC(s_i, n_j)
\]

4.2.2. Hot Backup Service Replacement Strategy (HBSRS) and QoS Analysis

In HBSRS, the backup services always remain in the activated states. If the service interruption occurs, the interrupted service will be replaced by its first backup service, and activate the second backup service at the same time. The algorithm has been described in our previous work [16]. The generalized execution time and the generalized execution cost of the atomic service \( (s_i, n_j) \) can be calculated as:

\[
\tilde{T}_{HBSRS}(s_i, n_j) = ET_{HBSRS}(s_i, n_j) + CT_{HBSRS}(s_i, n_j)
\]
\[
\tilde{C}_{\text{HBSRS}}(s_i, n_j) = EC_{\text{HBSRS}}(s_i, n_j) + CC_{\text{HBSRS}}(s_i, n_j)
\] (15)

The execution time \( ET_{\text{HBSRS}}(s_i, n_j) \) and execution cost \( EC_{\text{HBSRS}}(s_i, n_j) \) of the service \((s_i, n_j)\) can be calculated as:

\[
ET_{\text{HBSRS}}(s_i, n_j) = \min\{ST(s_i, n_j), ST(s_k, n_j)\}
\] (16)

\[
EC_{\text{HBSRS}}(s_i, n_j) = EC'_{\text{HBSRS}}(s_i, n_j) + EC_{\text{HBSRS}}(s_k, n_j)
\] (17)

Where the cost that has been executed by the atomic service \((s_i, n_j)\) when the service disruption occurs can be calculated as:

\[
EC'_{\text{HBSRS}}(s_i, n_j) = P(s_i, n_j)SC(s_i, n_j)
\] (18)

The execution cost of the backup service after the service disruptions can be calculated as:

\[
EC_{\text{CBSRS}}(s_k, n_j) = \sum_{k=1}^{K-1} P(s_k, n_j)SC(s_k, n_j) + P(s_k, n_j)SC(s_k, n_j) + P(s_{k+1}, n_j)SC(s_{k+1}, n_j)
\] (19)

4.3. Comparison of Various Recovery Strategies

According to the aforementioned analysis, we can conclude that the advantage of using backup service strategy to recovery the service composition lies in only when the service components and their all backup services are failed, the failure of the service composition will occur, thus greatly reducing the failure probability of service composition.

![Figure 2](image.png)

**Figure 2.** The comparison of various recovery strategies

Figure 2 shows the comparison of various recovery strategies. Assuming the service composition recovery succeed through a recovery after the service \((s_i, n_j)\) interrupt. The generalized execution time of the service \((s_i, n_j)\) in different strategies can be calculated as:

\[
\tilde{T}_{\text{RSRS}}(s_i, n_j) = P(s_i, n_j)ST(s_i, n_j) + ST(s_i, n_j)
\] (20)

\[
\tilde{T}_{\text{CBSRS}}(s_i, n_j) = P(s_i, n_j)ST(s_i, n_j) + ST(s_i, n_j) + T_e(s_i, n_j) + T_e(s_i, n_j)
\] (21)

\[
\tilde{T}_{\text{HBSRS}}(s_i, n_j) = \min\{ST(s_i, n_j), ST(s_i, n_j)\} + T_e(s_i, n_j) + T_e(s_i, n_j)
\] (22)

The generalized execution cost of service \((s_i, n_j)\) in different strategies can be calculated as:

\[
\tilde{C}_{\text{RSRS}}(s_i, n_j) = P(s_i, n_j)SC(s_i, n_j) + SC(s_i, n_j)
\] (23)
\[
\hat{C}_{CBSRS}(s, n_j) = P_i(s, n_j)SC(s, n_i) + SC(s, n_i) + C_i(s, n_i) + C_i(s, n_i)
\]  \tag{24}

\[
\hat{C}_{CBSRS}(s, n_j) = SC(s, n_i) + SC(s, n_i) + C_i(s, n_i) + C_i(s, n_i)
\]  \tag{25}

Table 1. Comparison of different recovery strategies for service composition

<table>
<thead>
<tr>
<th>Recovery Strategy</th>
<th>Execution time</th>
<th>Network load</th>
<th>Execution cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSRS</td>
<td>With the decrease of (P_i(s, n_j)) tend to endless</td>
<td>Minimum</td>
<td>Lowest</td>
</tr>
<tr>
<td>CBSRS</td>
<td>Long</td>
<td>Small</td>
<td>High</td>
</tr>
<tr>
<td>HBSRS</td>
<td>Short</td>
<td>Maximun</td>
<td>Highest</td>
</tr>
</tbody>
</table>

The comparison of different recovery strategies for service composition is shown in table 1. In RSRS, with the decrease of \(P_i(s, n_j)\), the service will be executed again and again. When the failure probability reaches a certain threshold, the execution time become infinite and the service composition will be invalid permanently. In CBSRS, because the backup services are not needed when the original service is valid, the execution cost is less, but the execution time is longer. Compared with the CBSRS, HBSRS reduced service recovery time. However, because backup services and original services are executed at the same time, the recovery cost increases accordingly. Obviously, HBSRS is a typical transcendental recovery strategy for service composition in dynamic network. It is a method that restores the service composition dynamically according to the service availability and the current state of service composition before the services interrupt.

5. Experiments and Results Analysis

5.1. Experimental Setup

All the experiments in this paper run on 50 mobile routing nodes in the scope of \(100m \times 100m\). According to the service stability analysis, the service availability \(P_i(s, n_j)\) of each atomic service in service composition \(SC_i\) is known. First, we control the service availability by switching on and off the services during the fixed experimental period \((\tau_1, \tau_2) \in [0, T_{exp}]\) to ensure the consistency of the stability of the same service node. The experimental parameters are shown in Table 2.

Table 2. The experimental parameters settings

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T_{exp})</td>
<td>The experimental time of each experiment</td>
<td>60</td>
</tr>
<tr>
<td>(D_{net})</td>
<td>Network range</td>
<td>(100m \times 100m)</td>
</tr>
<tr>
<td>(N_r)</td>
<td>The total number of routing nodes</td>
<td>50</td>
</tr>
<tr>
<td>(N_a)</td>
<td>The number of atomic services in a service composition</td>
<td>10</td>
</tr>
<tr>
<td>(N_i)</td>
<td>The number of interrupt services in a service composition</td>
<td>(\leq 10)</td>
</tr>
<tr>
<td>(B_i)</td>
<td>The number of backup services of the service ((s, n_j))</td>
<td>5</td>
</tr>
<tr>
<td>(k_i)</td>
<td>The interrupt times of the service ((s, n_j)) / The execution times of backup service</td>
<td>(\leq 5)</td>
</tr>
<tr>
<td>(P_i(s, n_j))</td>
<td>The availability of the atomic service ((s, n_j))</td>
<td>0.5</td>
</tr>
<tr>
<td>(P_b(s, n_j))</td>
<td>The availability of the backup service ((s, n_j))</td>
<td>0.5</td>
</tr>
<tr>
<td>(ST(s, n_j))</td>
<td>The service time of the atomic service ((s, n_j))</td>
<td>2s</td>
</tr>
<tr>
<td>(ST_b(s, n_j))</td>
<td>The service time of the backup service ((s, n_j))</td>
<td>2s</td>
</tr>
</tbody>
</table>
5.2. Experimental Results and Analysis

According to the analysis of the algorithm mentioned above, we can draw the conclusion that the performances of the recovery strategies for service composition are affected by the service availability, the number of service interruptions, and the number of the backup services. In this paper, we mainly consider the average execution time (AET) of service composition and the successful implementation probability (SIP) of service composition to measure the performance of the different recovery strategies. AET contains the normal service time and the recovery time of service composition, while SIP is measured by the successful times of service composition in 100 times of experiments, where the observation time of each experiment is 60 seconds. We compare four strategies, i.e., Re-sending the Service Request Strategy (RSRS), Cold Backup Service Replacement Strategy (CBSRS), Hot Backup Service Replacement Strategy (HBCRS), and Global search strategy (GSS) (i.e., re-establishing a new path after service composition interruption) in the same network.

5.2.1. The Impact of Service Availability on Service Composition Recovery

Figure 3 shows the AET and SIP of different recovery strategies when $\lambda_i = 1, k_i = 1, N_x = 10$. In GSS, AET is comparatively high, because when the service composition interruption occurs, GSS strategy will begin to re-discovery and re-establish a new service path with the very first atomic service. In RSRS, AET will increase along with the increase of the node failure probability. When the failure probability reaches a certain threshold, AET become infinite and the service composition will be invalid permanently. In CBSRS, AET is close to a linear function of the failure probability. But in HBSRS, once the service composition interruption occurs, the performance of service composition recovery is only relevant to the backup services and has nothing to do with the availability of the atomic services.

5.2.2. The Impact of the Number of Interrupt Services on Service Composition Recovery

Figure 4 shows the AET and SIP of different recovery strategies when $P(s_i, n_j) = 0.5, k_i = 1, N_x = 10$. It can be obtained that along with the increase of the number of interrupt services $N_x$, AET of the four strategies also increase linearly approximately. However, because the backup service replacement strategies eliminate service re-discovery and re-combination of GSS strategy, AET significantly decreases after the backup services being introduced. On the other hand, with the increase of the number of interrupt services $N_x$ and the rise of AET, SIP of four service composition strategies decreases in a certain period of time. However, in HBSRS, the successful implementation probability of service composition is obviously increased and the average time of service composition is reduced effectively.
5.2.3. The Impact of Execution Times of Backup Service on Service Composition Recovery

Figure 5 shows the AET and SIP of different recovery strategies when \( P(s_n, n_j) = 0.5, N_i = 1, N_e = 10 \). It can be obtained that with the increase of execution times \( k_i \) of the backup service. AET of the backup service replacement strategies increases continuously, while SIP continues to reduce. It’s because that the recovery time of CBSRS is a linear function of \( k_i \) when \( P(s_n, n_j) = 0.5, N_i = 1, N_e = 10 \). AET will continuously accumulate with the increase of the number of the service composition interruptions according to the algorithm analysis. By contrast, HBSRS takes the minimum execution time to recovery the service composition, so AET decreases, while SIP increases significantly accordingly.

6. Conclusions and Future Work

In this paper, we develop a hierarchical model for service composition in dynamic network. To solve the service composition interruption in dynamic network, we propose two recovery strategies: Cold Backup Service Replacement Strategy (CBSRS) and Hot Backup Service Replacement Strategy (HBSRS). Based on the nature and characteristics of dynamic network and the different failure causes of service composition, we further analyze the impact of service availability, the number of interrupt services, and execution times of backup service upon the performance of the recovery strategies for service composition in dynamic network. The simulation studies are conducted to show that the proposed strategies significantly improve the performance of service composition and effectively guarantee the availability and reliability of service composition in the dynamic network. The experimental results show that the proposed recovery strategies can significantly improve the reliability and availability of service composition in dynamic network.

This paper mainly focuses on the service-level recovery, and our future work will take the other factors such as the impact of the routing, restore and optimization of the network-level) in the dynamic network into consideration, and conduct the cross-layer research on the service composition recovery in the dynamic network.
7. Acknowledgement

This work was supported by the National Nature Science Foundation of China (Grant No. 61070182 and Grant No. 61170209), Beijing Municipal Education Commission Science Technology Development Plan Key Projects (KZ201010009008), and Natural Science Foundation of Hubei Province of China (No. 2011CDC029).

8. References


