An Approach to Composite QoS Parameter based Web Service Selection

Rajeev Pratap Singh, K.K.Pattanaik

Information and Communication Technology
Atal Bihari Vajpayee -Indian Institute of Information Technology and Management, Gwalior, India

Abstract

In any enterprise application Quality of Service (QoS) is an important attribute for selecting a service during the service composition process. Although availability and reliability have been considered as the predominant factors for estimating reputation, two aspects are lacking in the literature. First, their use is limited to composite service level and does not count at the atomic level. Second, their combined effect is not evaluated. Hence, the methodology of estimating reputation and its use for atomic service will have compounding effect on the overall quality of the composite service. We feel, better estimation of QoS can be done with both factors considered together on the simple premise that availability tells about only the probability of that service being up/running, but does not tell about its failure trend. In our work we present mathematical modelling of these predominant QoS factors using Markov model and Weibull analysis. A scenario has been simulated using Colored Petri Nets (CPN) to study the behavioral aspects. The outcome of our research is two fold. First, counting on probability of a service being up/running and its failure trend, together, results in a better estimation of its behavior and helps selecting the most appropriate one. Second, this resulted in selection of a service with higher reputation but lower usage cost, as opposed to using a single factor that resulted in higher reputation with higher cost.

© 2013 The Authors. Published by Elsevier B.V. Selection and peer-review under responsibility of Elhadi M. Shakshuki

Keywords: Web Service, Composition, QoS Modelling, CPN, Reputation.

1. Introduction

Web service (WS) architecture make it possible to achieve interoperability among atomic business processes. This has enabled creation of complex composite WSs [1, 2] that are used by many B2B applications today.

Creating a composite WS may require study of the need for the two major types of requirements with respect to an atomic service: Functional and Non-Functional. The former specify the behavioral characteristics like: number of input parameters, their data type, operations performed etc. while the latter specify the qualitative characteristics [3]. In a WS marketplace where various WSs are performing same functional
task, choosing atomic service components to create a composite services is a trivial task. Parameter centric selection is the only choice for selecting the one from many.

Service ‘reputation’ [4] is an important parameter to be considered while selecting atomic services for composition. Availability and reliability when taken separately for computing reputation will either give the probability of a service being up/running or will analyze the change in failure trend of that service respectively. However, we are interested in studying the combined effect of both these parameters in computing the reputation of an atomic service and thereby on composite service. The above scenario can be exemplified as follows: Let there be two services $S_A, S_B$ with $A_A = 0.8612, R_A = 0.6664^1$ and $A_B = 0.8607, R_B = 0.7146^1$ as availability and reliability of $S_A$ and $S_B$ respectively. Now if reputation is computed only on availability then service $S_A$ would be the choice for selection. However, by comparing their reliability one will conclude that reliability of $S_B$ is better and is the choice for selection. Study of the behavior of a service in terms of its availability and reliability over a sample space would provide better knowledge about its suitability for selection. Current work is mainly being done on computing these QoS parameters for composite services and not at the granular atomic level. The major contribution of this paper is to propose a model to quantify availability and reliability of atomic services using Markov Chain model and Weibull analysis respectively. Also importance of modelling reputation as an aggregation of availability and reliability has been explained.

The rest of the paper is structured as follows: Section 2 discusses the related work. The proposed approach is presented in Section 3. Section 4 explains the experiment design and simulation. Results and analysis is presented in section 5. Finally, Section 6 concludes the paper.

2. Related Work

Research in the field of WS composition based on QoS can be broadly put into two directions. One focus on how the various QoS attributes can be quantified for their effective use and the other focus on their use at various stages of service selection or composition. In the following we will present efforts and outcomes of both the research directions.

In the direction of quantification of QoS attributes [5] proposed ‘WSrep’ a framework to model reputation based on user feedback. This approach uses maximization of subjective and objective view of past behavior of service providers and uses this knowledge to calculate reputation of a service. One thing that is worth considering here is that there could be a malicious user who could use a biased feedback to alter the reputation of the service, thus there is a need to include automated server side calculated attributes availability and reliability to give a precise estimate of reputation of service. In [6] the authors have used user feedback as the basis of reputation calculation. In order to quantify reputation, history of a service is considered as basis of reputation quantification. There is a need to include history of implicit attributes availability and reliability in order to achieve a better evaluation of reputation. In [7] Feedback Forecasting Model is proposed that considers two major aspects during the rating process: first to provide an automated feedback for customers who are fearful in giving feedback or don’t bother to provide feedback. Second to check that what is the credibility of feedback of a particular customer. Here user rating is considered for computing reputation and implicit QoS attributes are not considered. In [8] authors have used both availability and reliability to compute the reputation of the service. However both these attributes are used separately and their combined effect is not studied. Since availability gives the probability of success and reliability gives the change in failure trend, both must be used collectively to compute the reputation of a service. In [12] authors have proposed probabilistic methods to quantify QoS attributes: Cost, Throughput and Time, and studied their aggregated effect on composite service. However the quantification of QoS attributes Availability, Reliability and their collective effect to compute reputation is not studied.

In the direction of the use of QoS at selection and composition following works are considered. In [9] the Ontology Web Language for Service (OWL-S) is used to perform a functional match among the

---

1 Reliability tells the failure trend of a service. The values of $R_A$ and $R_B$ signify the gradient of the tangent to the curve at the given time. Since the gradient of service $S_B$ is greater than the gradient of service $S_A$, it means that $S_B$ has an increasing success rate or decreasing failure rate as compared to $S_A$. Thus $S_B$ will be considered to be more reliable than $S_A$. 

available services and then these services are rated according to their QoS score, however the methodology of calculating the QoS score is not mentioned. In [10] the authors proposed branch and bound for execution plan selection (BB4EPS) algorithm that creates a plan for service composition using the aggregated affect of the QoS attributes. The aggregation is studied for services connected in different structures/patterns. Both availability, reliability are studied separately and their combined effect is not evaluated. A relaxable QoS-based service selection algorithm (RQSS) is proposed in [11] that created a composite service by relaxing QoS criterion of the overall composite service using heuristic techniques like MMKP (Multi-Dimensional Multi Choice Knapsack). The work studies the composition process for a composite service based on QoS attributes: Execution Time, Reliability, Availability, Reputation and Price. Here reputation is based on user feedback. Also the combined behavior of availability and reliability to calculate reputation is not considered.

The novelty of our work are: first a probabilistic model for quantifying two different QoS attributes and second modelling the reputation of a service.

3. Composite QoS Parameter based Web Service Selection

Selection of an atomic service during composition process is based on various QoS parameters. Services are ranked based on their QoS values and selection is done based on rank. We propose an approach to model service reputation as a combination of availability and reliability.

3.1. Availability modelling

Service Availability is the probability that a service will be up and running during the course of study. The main events that reflect status of service availability are the transition of a service from up to down state or down to up state. As an up/running service could fail due to any of these reasons like: server going down, hardware failure, internal logical error etc., we have generalized these failure conditions to a single down state.

Markov Chain models [13] are used to study systems that could be represented as discrete states. Since service availability can have boolean discrete states we used Markov chain to model availability. Fig. 1 depicts the state transitions for a service along with their transition probabilities. $p_{ij}$ value represents the transition probability of moving from state $i$ to state $j$.

![Fig. 1. Markov model for availability](image)

All the transition probabilities of the given Markov chain model for availability could be represented in the form of Transition Probability Matrix($P$) as follows:

$$
P = \begin{bmatrix}
p_{00} & p_{01} \\
p_{10} & p_{11}
\end{bmatrix} = \begin{bmatrix}
(1 - q) & q \\
p & (1 - p)
\end{bmatrix}
$$

Entries of Matrix $P$ ($p_{ij}$) correspond to the Markov Chain’s single-path length transition probabilities. The row elements of matrix $P$ correspond to states that the system currently in and column elements denote the
next state. Current transition matrix gives the transition probability of single-path length. For better estimate of transition probabilities paths of all the possible lengths are considered to obtain $P^n$ matrix.

$$P^n = \begin{bmatrix} 1 - q & q \\ p & 1 - p \end{bmatrix}^n = \frac{1}{p + q} \begin{bmatrix} p & q \\ p & q \end{bmatrix} + \frac{(1 - p - q)^n}{p + q} \begin{bmatrix} q & -q \\ -p & p \end{bmatrix}$$

(2)

For sufficiently large values of $n$ the results are interpreted as long run averages or limiting probabilities $P_i$ of system being in state $i$. In such cases $P^n$ reduces to:

$$\lim_{n \to \infty} P^n = \begin{bmatrix} p & q \\ p(q - p) & q \end{bmatrix}$$

(3)

From above it is concluded that availability is the probability of moving from state 1 to again state 1 in a path-length $n$. Thus the availability of the system on the long run can be said as:

$$Availability = P^n_{11} = \frac{q}{p + q}$$

(4)

3.2. Reliability modelling

Reliability [14] of a service defines the rate of change of failure of a service under test. Various probabilistic distribution mechanisms are used to study reliability. In a realistic scenario services’ failure rate can be increasing, decreasing, or constant. To accommodate all these cases Weibull analysis has been considered. We make the following assumptions to model reliability: $x = threshold for successful samples$, $n = sample space in a given time frame$, $F(x) = failure rate$ and $R(x) = reliability$.

Failure of a service is a rare event as it does not occur as frequently as the success event, thus it is Poisson distributed. In the following we model failure of an atomic service.

Failure rate $F(x)$ and reliability $R(x)$ can be related by:

$$F(x) = 1 - R(x)$$

(5)

where: $R(x) = e^{-(\frac{x}{\alpha})^\beta}$

On simplification equation 5 can be written as:

$$\ln \left[ \ln \left( \frac{1}{1 - F(x)} \right) \right] = \beta (\ln \alpha) - \beta (\ln x)$$

(6)

Comparing parameters in Eq. (6) with that of straight line provides information about the coordinates X and Y which in-turn are used to perform linear regression. This will provide the estimate for $\alpha$ and $\beta$ for computing the reliability of an atomic service. $\alpha$ is the Weibull Characteristic Life and is a measure of spread in the distribution and $\beta$ is the Shape Parameter that determine the nature of failure rate. In general failure rate of an atomic service $S_i$ is represented as:

$$FailureRate(S_i) = \begin{cases} \text{increasing, if } \beta > 1.0 \\ \text{constant, if } \beta = 1.0 \\ \text{decreasing, if } \beta < 1.0 \end{cases}$$

(7)

3.3. Reputation modelling

Reputation is modelled as a composite QoS attribute comprising both availability and reliability, so mathematically reputation is modelled as:

$$Reputation(S_i) = Availability(S_i) \times Reliability(S_i)$$

(8)

4. Simulation

The simulation was carried using CPN [15] and is mainly divided into following two sub-sections.
4.1. Input data

In order to compute the QoS parameters availability, reliability and reputation we need failure count data. Failure count signifies the number of samples for which the service is found to be down and not responding. As the event of service going down is a rare event, thus a Poisson distribution is taken to model failure count. QoS parameters are calculated over a sample space of 1440 samples taken over 24 hours with one sample taken at every minute. The failure count data of a service is recorded on daily basis and for a number of days for QoS estimation. In our simulation we have considered the acceptable failure count for an atomic service to be 10% of sample space. Such pattern of failure data is Poisson distributed with mean $\mu = 10\%$ of sample space, such that the distribution generates 75% values in interval of $\mu \pm \sigma$. Usage cost, another input data, signifies the cost that a user pays for a service. The usage cost of functionally same services generally lie in a range and thus uniformly distributed over that range.

4.2. Simulation details

![Diagram](image)

The nets used in our simulation are depicted in Fig. 2 through Fig. 4. Fig. 2 represents the overall architecture of our methodology in form of a hierarchical CPN. All the samples of all the services are taken as input for calculating availability and reliability. On the basis of Eq. (8) the net computes Reputation of the selected service which lies in range [0, 1].

Fig. 3 illustrates the process of computing availability. It takes all the samples of all the services as input and computes availability of one atomic service at a time. Availability for each atomic service is computed by dividing sum of failure count with total number of samples. One result token is generated for every atomic service reflecting its availability in the range [0, 1]. A value in this range represents the probability of a service being up/running.

Fig. 4 illustrates reliability computation process. Samples of a selected service form the input of the subnet. Linear regression is performed on these samples to evaluate regression parameters $S_{x}$, $S_{xx}$, $S_{yy}$, $S_{y}$ and $S_{xy}$. Applying Weibull analysis on regression parameters reliability is computed in the range [0, 1]. Here the numeric value of reliability signifies the success trend, more the reliability value more is success rate. Result also evaluates Weibull shape parameter $\beta$ that represents failure trend.
Fig. 5 illustrates various colorsets used in creating the CPN nets for availability, reliability and reputation.

5. Results and Analysis

We performed a series of experiments to evaluate effectiveness, performance and feasibility of the proposed system. To study the QoS parameters, 20 functionally equivalent services $S_1, S_2, S_3, S_4, S_5, \ldots S_{20}$ are chosen for our experiment. The usage cost of these services is uniformly distributed in range $[10, 50]$. The failure count of each service is poisson distributed with mean $\mu = 144.0$. QoS parameters of services are studied for samples collected over 2 days. Fig. 6 explains the failure trend $\beta$ for sample size in range $[1, 10]$. Fig. 6 illustrates that $\beta$ has a significant dispersion for sample size in range $[1, 4]$. A sample size of 2 days gives maximum dispersion. For samples of size more than 4 days, the value of $\beta$ becomes nearly constant which signifies failure rate does not increase further. This supports our hypothesis that analysis of past behavior for a short duration (2 days in our case) gives better estimate of failure trend of a service.

Fig. 7 illustrates the rationale for how a composite QoS based reputation modelling would give better outcome. If reputation is taken only on availability then service $S_{10}$ has highest reputation and if based only on reliability then service $S_{15}$ has highest reputation. Hence, when these parameters taken individually gives different estimates of reputation. As depicted in Fig. 7 service $S_{13}$ has highest reputation, calculated
6. Conclusion

Web service reputation is an important parameter in QoS based WS composition. Selecting atomic services with high reputation helps in creating robust, high performance, and cost effective composite services. Availability focuses on probability of a service being up/running and reliability computes the failure trend of service. This paper demonstrated a methodology that models reputation as a combination of both availability and reliability, giving a precise estimate of reputation. Finally CPN based simulation experiments are reported, demonstrating effectiveness and benefits of the proposed methodology. This research outcome
would help further to consider other related QoS parameters for modelling reputation and its relation to WS based business process solutions.

References