

利用服务质量参数值预测的 Web 服务选择方法

翟社平¹, 魏娟丽², 李增智^{1,3}

(1. 西安交通大学电子与信息工程学院, 710049, 西安; 2. 西安体育学院网络信息中心, 710068, 西安;
3. 北京邮电大学网络与交换技术国家重点实验室, 100876, 北京)

摘要: 为了使得服务消费者获得的 QoS 值能够作为服务分级和排序的重要依据, 提出了一种基于 QoS 参数值预测的 Web 服务选择算法. 依据服务质量一致性指标生成服务质量报告, 并对可信报告进行标记, 再利用预处理算法来剔除欺骗报告. 采用模糊回归方法对服务质量值进行预测: 将服务质量的属性值和服务质量评价值分别作为方法的输入和输出, 以此来确定模糊系数; 把服务质量的评价模糊化为对称的三角模糊数; 最后将获得的 QoS 预测结果作为服务选择的依据. 实验结果表明, 随着质量报告样本数的增加, 服务质量预测值逐渐接近实际的 QoS 值. 在服务发现对比实验中使用服务质量预测方法, 通过调整服务质量的一致性、服务消费者能够接受的服务广告和服务请求之间的质量差, 实验结果的平均准确率误差可以控制在 2.7% 以内.

关键词: Web 服务; 服务质量; 服务选择; 模糊预测

中图分类号: TP393 **文献标志码:** A **文章编号:** 0253-987X(2010)06-0057-05

A Web Service Selecting Strategy Based on Prediction with Quality of Service Parameter Values

ZHAI Sheping¹, WEI Juanli², LI Zengzhi^{1,3}

(1. School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China; 2. Network Information Center, Xi'an Physical Education University, Xi'an 710068, China; 3. State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: To enable Web service consumer to acquire the quality of service (QoS) value as a reference for service ranking and sorting, a Web service ranking algorithm based on the QoS parameters prediction is proposed. According to the measure of service quality conformance, the service quality reports are produced, the creditable reports are marked, and the preprocessing method is chosen to exclude those cheating reports. The QoS value is predicted with the fuzzy regression method, whose I/O parameters get the property and evaluation values of service quality respectively, so the fuzzy coefficients can be determined and the QoS evaluation is hazed with symmetric triangular fuzzy number, and the results are employed as the service selecting basis. The experimental results demonstrate that the predicted value is close to the real QoS value gradually as the sample number of the service quality report increasing. The service quality prediction, compared with the others in the service discovery, enables the average R-precision error to be controlled within 2.7% with the varying parameters of service quality conformance and the QoS difference between service advertisement and requirement accepted by service consumer.

Keywords: web services; quality of service; service selecting; fuzzy prediction

从发现的大量具有相似功能的 Web 服务中选择最合适的服务是一项挑战性的任务,往往需要考虑服务的 QoS 属性^[1]. 服务消费者要么使用服务广告的 QoS 值选择服务,要么依赖自己的经验对服务性能做出判断,在没有直接的经验可用时,可以参照其他服务消费者对于服务提供者或服务本身的评价^[2]. 但是,在开放环境中,要参照其他服务提供者的服务评价存在两个问题:首先,服务消费者不知道怎样使用特定服务评价信息,这些信息往往具有模糊性且难以使用;其次,服务消费者不清楚这个评价信息的可靠性如何,一方面服务提供者出于自身的利益考虑,在服务广告中发布高于实际服务水平的 QoS,吸引服务消费者使用而获取利润,另一方面服务消费者反馈的 QoS 数据具有很大的随意性,甚至是一些恶意的虚假数据^[3]. 本文提出使用服务质量报告对服务的质量进行预测的方法. 由于存在虚假报告,首先设计算法对服务质量报告进行预处理,剔除那些不可信报告,使用可信报告对服务质量进行预测,并在此基础上对候选的服务进行选择和排序.

1 服务质量评价

参与 Web 服务发现的双方 Agent 往往之前从没进行过交互,对服务发现 Agent 来说,对服务提供者以及他提供的具体服务只有一些不完全信息^[4],这样使得服务提供者提供给消费者的服务未必是最好和最可靠的,并且服务消费者没有办法在服务交付使用前对服务质量进行验证. 因此,本文提出如下的服务质量评价方法.

1.1 服务质量评价场景

在 Web 服务发现系统中引入质量评价系统有助于服务的提供者和消费者 Agent 进行决策,它的基本思想是在一次服务发现并被交付后,让参与服务的双方相互评价,据此产生服务消费者和服务提供者的信任或名誉的评分等级,同时产生服务质量报告. 服务质量报告经过质量预测方法计算后能够作为其他服务 Agent 将来与之交互的决策依据. 如图 1 所示,假定服务 Agent A 同 Agent E、G,服务 Agent H 同 B、D 曾经有过服务交易的历史,并根据交易情况产生服务质量报告,当从未交易过的 Agent A 和 H 之间进行请求/提供服务时,已产生的服务质量报告可以作为 A 和 H 相互服务质量预测的依据(这里假定一个服务提供者同时也是服务的消费者).

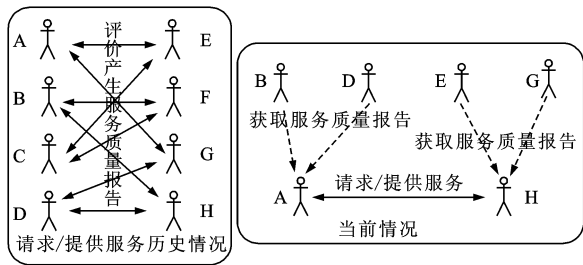


图1 使用服务质量报告场景

1.2 服务质量报告

为了对服务质量进行预测,从服务 Agent 的信任列表中提取一段时间内某个特定服务的服务质量值以及服务消费者反馈的实际交付质量值,根据这些数据对于服务的性能进行预测,并结合服务的分级和选择技术,发现最适合服务消费者的服务.

定义1 服务质量一致性反映了服务提供者广告的服务质量值和实际交付质量值之间的差异程度,其定义为 $c_{ij}^{ws} = \frac{|a_{ij}^{ws} - d_{ij}^{ws}|}{d_{ij}^{ws}}$,其中 i 表示服务质量数据

获取的时间, j 表示 Web 服务(ws)的第 j 个非功能属性, a_{ij}^{ws} 表示服务的广告质量值, d_{ij}^{ws} 表示服务的实际交付质量值. 服务消费者据此对信任 Agent 进行标记,服务质量一致性较好的服务提供者 Agent 被标记为诚实.

定义2 服务消费者的服务质量信任报告反映了服务消费者对于服务交付质量的最近评定结果,表示为一个向量 $r_q = \langle a, s, t, q_{ij}, c_{ij} \rangle$,其中 a 是提供报告的服务消费者 Agent 的标识符, s 表示被报告的 Web 服务的名称, t 表示报告的时间戳, q_{ij} 表示服务质量值, c_{ij} 表示服务质量一致性值. 服务消费者对来自于信任 Agent 的服务质量报告标记为可信赖报告,其提供的服务质量数据能够直接参加服务质量的预测.

定义3 服务消费者的非功能属性需求表示为 $r = \langle q_{i1}, q_{i2}, \dots, q_{im} \rangle$,其中 q_{im} 反映了服务消费者对某个服务质量值的需求. 用向量 Δ 表示服务消费者能够接受的服务广告和服务请求之间的质量差, $\Delta = \langle \delta_{q_{i1}}, \delta_{q_{i2}}, \dots, \delta_{q_{im}} \rangle$,对于服务消费者来说, Δ 越接近于 0,表示其对服务质量越满意.

1.3 服务质量报告预处理算法

为了提高服务质量预测的准确性,算法首先对服务质量报告进行预处理,剔除掉那些具有欺骗性的质量报告,使得返回的服务提供者的服务质量数据尽可能接近服务消费者的需求. 预处理算法如图 2 所示.

```
Algorithm1 QoSReportsPreprocessing()
设置阈值(C和Δ值)并从能够交互的Agent处
获取所有的服务质量报告
将诚实Agent提供的报告标记为可信的
Repeat
    查找所有提供报告的未标记服务Agent
    For each unmarked report do
    IF 报告的服务质量一致性值大于等于 C
        报告标记为可信报告
    Else
        此报告标记为欺骗报告
    End if
End for
Until 所有提供报告的服务Agent被处理
For 所有标记为可信任服务质量报告
    计算Δ值
    IF Δ大于等于设定的阈值
        将该报告加入服务质量预测报告集合
    End if
End for
输出服务质量报告集合
```

图 2 服务质量报告预处理算法

2 服务质量预测

2.1 服务质量的模糊回归分析

假定 Web 的服务质量使用变量 $q_1^{ws}, q_2^{ws}, \dots, q_m^{ws}$ 表示,在 n 个不同的时刻对服务的质量报告进行获取,并得到服务消费者对 Web 服务交付质量的整体评价. q_{nm}^{ws} 表示服务质量属性 m 在时刻 n 的获取值, e_n^{ws} 表示在时刻 n 对服务质量的整体评价. 不同时刻的质量评价可表示成如下一组等式

$$\begin{cases} e_1^{ws} = \beta_0 + \beta_1 q_{11}^{ws} + \beta_2 q_{12}^{ws} + \dots + \beta_m q_{1m}^{ws} + \epsilon_1 \\ e_2^{ws} = \beta_0 + \beta_1 q_{21}^{ws} + \beta_2 q_{22}^{ws} + \dots + \beta_m q_{2m}^{ws} + \epsilon_2 \\ \vdots \\ e_n^{ws} = \beta_0 + \beta_1 q_{n1}^{ws} + \beta_2 q_{n2}^{ws} + \dots + \beta_m q_{nm}^{ws} + \epsilon_n \end{cases}$$

式中: $\beta_0, \beta_1, \beta_2, \dots, \beta_m$ 是待定的 $m+1$ 个回归系数; $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ 为 n 个相互独立且服从正态分布 $N(0, \sigma^2)$ 的随机误差变量. 现令

$$e = \begin{pmatrix} e_1^{ws} \\ e_2^{ws} \\ \vdots \\ e_n^{ws} \end{pmatrix}; q = \begin{pmatrix} 1 & q_{11}^{ws} & q_{12}^{ws} & \dots & q_{1m}^{ws} \\ 1 & q_{21}^{ws} & q_{22}^{ws} & \dots & q_{2m}^{ws} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & q_{n1}^{ws} & q_{n2}^{ws} & \dots & q_{nm}^{ws} \end{pmatrix}$$
$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}; \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

这样,总体质量评价与服务的各个属性质量之间的关系可以使用 $e=q\beta+\epsilon$ 来表示. 由于服务系统结构的不确定性以及服务质量描述的不精确性,有必要

引入模糊数对服务质量进行评价^[5]. 本文将服务质量的各个属性值和服务质量评价值作为输入和输出,确定出模糊系数 $\beta_j (j=1, 2, \dots, m)$. 把服务质量 $e_i^{ws} (i=1, 2, \dots, n)$ 的评价模糊化为对称的三角模糊数 $E_i^{ws}(e_i^{ws}, c_i)$, 其中 c_i 根据实际问题确定, β_j 表示为对称模糊数 $\beta_j(a_j, e_j)$. 为了得到 Web 服务整体质量的预测值,问题变为需要确定参数 a_j 和 $e_j (j=1, 2, \dots, m)$. 根据服务质量模糊预测模型 $\tilde{e}_i^{ws*} = \tilde{\beta}_0 + q_{i1}\tilde{\beta}_1 + \dots + q_{im}\tilde{\beta}_m$, 服务质量总体评价模糊数可表示为 $\tilde{e}_i^{ws*} (a_0 + \sum_{j=1}^m q_{ij}^{ws} a_j, e_0 + \sum_{j=1}^m q_{ij}^{ws} e_j)$.

2.2 服务质量拟合度

当 $|e_i^{ws} - (a_0 + \sum_{j=1}^m q_{ij}^{ws} a_j)| \leq c_i + e_0 + \sum_{j=1}^m |q_{ij}^{ws}| e_j$ 时,服务质量总体评价的经验值和估计值之间的拟合

度 h_i 为 $1 - \frac{|e_i^{ws} - (a_0 + \sum_{j=1}^m q_{ij}^{ws} a_j)|}{c_i + e_0 + \sum_{j=1}^m |q_{ij}^{ws}| e_j}$, 其他时为 0.

2.3 服务质量整体评价估计值的模糊度

对于给定的 Web 服务质量反馈经验值,当 $\tilde{e}_i^{ws*}$ 的中心位置 $a_0 + \sum_{j=1}^m q_{ij}^{ws} a_j$ 固定时,模糊度越大,则拟合度 i 也越大. 在减小模糊度的情况下,需尽可能提高拟合度,因此应该使质量预测的整体模糊系数 $(\beta_0, \beta_1, \dots, \beta_m)$ 的综合模糊度达到最小. 这样,服务质量预测方程的模糊系数求解就转化为如下的线性规划问题. 目标函数

$$\min S = \frac{1}{3} (e_0 + \sum_{j=1}^m q_{ij}^{ws} e_j), i = 1, 2, \dots, m$$

约束条件

$$\sum_{j=1}^m q_{ij} \alpha_j + (1-H) \sum_{j=1}^m |q_{ij}| e_j \geq e_i^{ws} - (1-H)\epsilon_i$$
$$\sum_{j=1}^m q_{ij} \alpha_j - (1-H) \sum_{j=1}^m |q_{ij}| e_j \leq e_i^{ws} + (1-H)\epsilon_i$$

式中: $e_j > 0$; H 为给定拟合度; $i=1, 2, \dots, n$; $j=1, 2, \dots, m$.

在对具体的服务质量预测中,参数 H 和 e_j 根据具体的应用给定,将目标函数简化为如下等式

$$\min S = w_0 e_0 + w_1 e_1 + \dots + w_j e_j$$

$$w_0 = \frac{1}{\sum_{k=1}^m (\max_{1 \leq j \leq n} |q_{jk}|)}; w_j = \frac{\max_{1 \leq j \leq n} |q_{ij}|}{\sum_{k=1}^m (\max_{1 \leq j \leq n} |q_{jk}|)}$$

2.4 基于质量预测的 Web 服务选择和排序算法

对于特定的服务请求,首先将功能匹配成功的

服务列表,对其中满足消费者 QoS 需求的服务进行质量预测,并给出最终依据预测结果排序的服务列表,算法如图 3 所示.

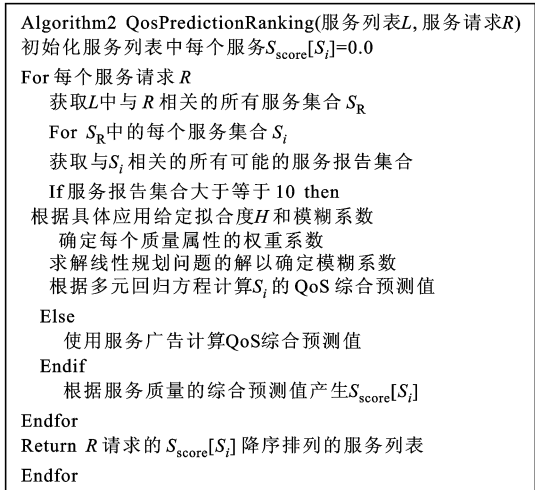


图 3 QoS 预测的服务排序算法

3 仿真实验

3.1 服务质量预测方法的有效性验证

为了验证模糊线性回归方法对于服务质量预测的有效性,设计一组服务 Agent 作为服务提供者,并设计不同的服务请求(改变服务质量属性值)去发现相应的服务,根据实际交付的服务,服务消费者 Agent 对服务的 QoS 属性质量进行评分,产生其对应的 d_j^{qs} . 使用模糊线性回归预测方法对其实际交付质量进行预测,将其同实际的 QoS 值进行比较,实验时固定服务提供者的 QoS 值保持在 8.5,其结果如图 4 所示. 将服务质量预测的消费者评价样本固定在 10、50、100、150 和 200,发现随着参加预测的消费者评价报告样本的增加,服务的预测值逐渐接近实际服务质量属性的总体评价,样本数为 10 个时,预测的误差大约为 34%. 当报告样本数增加到 200 个时进行预测,预测误差减小到约 1.2%. 随着样本数的增加,该方法的预测值更加接近服务的真实质量值.

3.2 服务选择和评价算法的验证

为了评价基于模糊预测的服务选择和评价算法,本文选择了房屋租赁场景^[6]来进行服务质量测试,服务请求者 Agent 根据实际提供的服务属性交付值对服务提供者 Agent 进行评价,产生服务质量报告. 在不同的背景下对服务发现的 R-Precision 指标进行了实验评价,即只针对满足服务请求者

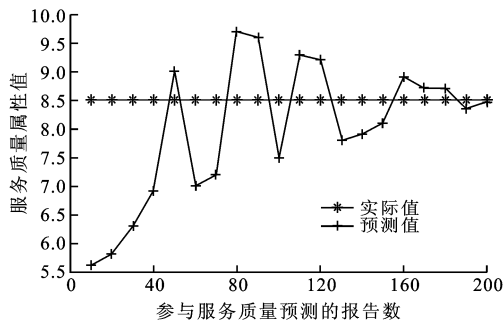


图 4 服务质量预测曲线

Agent 的 R 个服务准确度进行计算,分 4 种情况进行:①理想情况是参加服务发现的所有 Agent 都是诚实的,即它们提供的服务质量报告都是可信的,服务提供者的服务广告质量和交付质量完全相同;②实际情况是部分服务质量报告不可信赖,使用算法 QoSReportsPreprocessing 进行预处理;③所有的服务质量报告进入服务质量预测;④不使用报告对服务质量进行预测,根据服务提供者的广告作为服务选择的依据. 使用 R-Precision 参数反映在特定服务请求下,与之相关的 R 个服务的准确度,通过改变参加预测的服务质量报告的可信报告的百分比, R-Precision 参数的实验结果如图 5 所示. 在实际情况下,由于使用报告预处理算法对不可信报告进行剔除,使得预测结果非常接近理想情况,通过调整一致性参数 C 和质量差参数 Δ 值,服务发现的 R-Precision 准确度能够控制在 2.7%(平均值)以内. 在不对质量报告进行预处理的情况下,服务 QoS 预测的准确度较差,但随着可信报告数的增加,服务 QoS 预测的准确度曲线逐渐接近理想情况. 使用服务广告的 QoS 值进行预测,造成服务发现的 R-Precision 曲线变化较大,这也说明不能直接使用服务提供者发布的服务 QoS 值进行服务的选择.

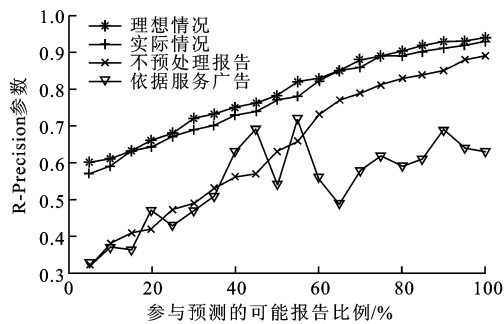


图 5 服务质量预测的 R-Precision 参数比较曲线

4 结 论

从多个满足功能需求的候选服务中选择消费者最满意的服务,使用服务质量预测机制可以明显改善消费者使用服务的整体效用. 本文提出使用模糊线性回归对服务质量属性进行预测的方法来进行服务的选择和排序,首先通过服务质量报告预处理算法产生可信的服务质量报告,然后对质量属性值进行预测,依据预测结果,产生满足服务消费者需求的服务排序结果. 下一步我们将对影响服务性能的上下文参数进行处理,获取影响服务性能和服务消费者需求的环境信息的变化情况,据此对服务的提供做出动态的响应.

参考文献:

[1] ABRAMOWICZ W, KACZMAREK M, ZYSKOWSKI D. Managing web service quality: measuring outcomes and effectiveness[M]. Hershey, USA: IGI Global, Information Science Reference, 2008:96-113.

[2] KLAN F, KONIG-RIES B. A personalized approach to experience-aware service ranking and selection[C]// Second International Conference on Scalable Uncer-

tainty Management(SUM2008). Napoli, Italy: Springer, 2008:270-283.

[3] 李研,周明辉,李瑞超. 一种考虑 QoS 数据可信性的服务选择方法[J]. 软件学报, 2008, 19(10): 2620-2627.

LI Yan, ZHOU Minghui, LI Ruichao. Service selection approach considering the trustworthiness of QoS data[J]. Journal of Software, 2008, 19(10): 2620-2627.

[4] JØSANG A, ISMAIL R, BOYD C. A survey of trust and reputation systems for online service provision[J]. Decision Support Systems, 2007, 43(2): 618-644.

[5] 吴冲,潘启树,李汉铃. 模糊线性回归预测[J]. 西安交通大学学报, 2000, 34(9): 100-102.

WU Chong, PAN Qishu, LI Hanling. Fuzzy linear regression prediction[J]. Journal of Xi'an Jiaotong University, 2000, 34(9): 100-102.

[6] ZHAI Sheping, WANG Hai, WEI Juanli. A pragmatics web service oriented approach to understanding the semantics of concepts [C]//IEEE Asia-Pacific Conference on Services Computing. Piscataway, NJ, USA: IEEE Computer Society, 2008:1069-1074.

(编辑 杜秀杰 苗凌)