

基于动态 QoS 的 Web 服务选取方法

方 晨 王晋东 于智勇

(解放军信息工程大学 郑州 450001)

摘 要 针对动态环境中 Web 服务 QoS 属性值波动的问题,提出一种基于动态 QoS 的服务选取方法。该方法首先建立区间 QoS 模型来表示 QoS 属性值的动态变化,然后用区间相似度衡量候选服务提供的 QoS 属性与用户需求值的接近程度。基于相似度的概念,采用逼近理想点的多属性决策方法计算出每个基本服务的 QoS 指标客观权重,并结合用户主观偏好得出 QoS 指标综合权重,最终基于推荐度对候选服务进行排序。仿真实验表明,此服务选取方法既充分考虑了用户的主观偏好,又克服了 Web 服务 QoS 属性值的波动影响,提高了服务选取的准确性。

关键词 服务选取,动态变化,相似度,区间数,用户主观偏好

中图分类号 TP181 文献标识码 A DOI 10.11896/j.issn.1002-137X.2017.05.044

Web Service Selection Method Based on Dynamic QoS

FANG Chen WANG Jin-dong YU Zhi-yong

(PLA Information Engineering University, Zhengzhou 450001, China)

Abstract In order to deal with the Quality of Services(QoS) value fluctuations of Web services in dynamic environment, a service selection method based on dynamic QoS was proposed. Firstly, the method establishes an interval QoS model to represent the dynamic changes of QoS values. Then it uses the interval similarities to measure the proximity of the QoS values provided by candidate services to the QoS requirement values provided by consumers. Based on the concept of similarity, the objective weight of QoS criteria for each basic service is calculated using TOPSIS(Technique for Order Preference by Similarity to Ideal Solution) method for MADM(Multiple Attribute Decision Making) problems. After calculating the comprehensive weight by combining with the consumers' subjective preference, recommendations are used to sort the candidate services. The experimental results show that this service selection method not only considers the consumers' subjective preference, but also overcomes the impact of fluctuations in QoS values of Web services, and enhances the correctness of service selection.

Keywords Service selection, Dynamic changes, Similarity, Interval number, Consumers' subjective preference

1 引言

随着云计算技术的迅速发展和商业应用的普及,越来越多的数据资源、计算资源和应用资源依托 Internet 以 Web 服务的形式提供应用。近年来,部署在 Internet 上的 Web 服务呈爆炸式增长,具有相同功能属性、不同非功能属性的 Web 服务也越来越多。因此,传统的依据功能属性进行服务选取的方法已经不能够满足用户的需求,如何依据服务的非功能属性如服务质量(Quality of Service, QoS)来进行服务选取,成为学术界关注的热点。

文献[5]建立了一个 QoS 属性云模型,利用期望、熵、超熵这 3 个数字特征来预测 Web 服务 QoS 数据的波动情况,根据用户提出的 QoS 约束条件,筛选出性能稳定的 Web 服务,

在此基础上以 QoS 属性相似度作为权重计算候选服务的 QoS 聚合值并排序,提高了服务选取的准确性。文献[6]提出了 Web 服务各随机 QoS 指标的度量方法,同时扩展了 Web 服务 QoS 管理体系结构,使 Web 服务的 QoS 指标值随着各因素的变化能够自适应地更新,从而使查询所得的 QoS 值与实际执行时的结果偏差较小,提高了 Web 服务选取的准确率。但是该文献在对 QoS 进行规范化处理时没有考虑用户的个性化需求。文献[7]提出了一种基于模糊层次分析法(FAHP)的多维 QoS 局部最优服务选择模型,该模型将 QoS 的真实度属性作为多维 QoS 向量的分量之一,然后建立了包含双重质量属性(客观属性和主观评价)的模糊层次结构,全面考查了主客观 QoS 属性对服务最优选择的影响,然而它并未考虑到动态环境下 QoS 属性值的波动。文献[8-9]提出了

收稿日期:2016-04-17 返修日期:2016-06-03 本文受国家自然科学基金资助项目(61309013, 61303074),河南省科技攻关计划项目(12210231003)资助。

方 晨(1993—),男,硕士生,主要研究方向为云服务选取和推荐;王晋东(1966—),男,教授,博士生导师,主要研究方向为云计算、网络安全;于智勇(1992—),男,硕士生,主要研究方向为云服务信任评估。

一个上下文感知的服务选取方法,它采用 Web 本体描述语言 (Web Ontology Language, OWL)/资源描述框架 (Resource Description Framework, RDF) 模型,将请求者的上下文转换为基于服务类别和用户领域相关的约束值,但是偏好权重是由用户根据一定语义主观设定的。以上研究方法在很大程度上提高了服务选择的准确性和可靠性,但是仍存在以下不足。

(1) 没有考虑到 Web 服务 QoS 属性值的动态不确定性 (简称动态性)。该动态性体现在两个方面: 1) Web 服务运行在动态的网络环境中,任何因素如网络带宽、位置、时间等都会影响到 Web 服务的 QoS 属性值,因此根据精确的 QoS 属性值进行服务选取的准确性不高; 2) 由于人的思维模式的模糊性,使得用户更倾向于使用区间数而不是精确值来表达其对服务的需求。考虑 QoS 的动态性应作为服务选取和服务组合的前提,而已有的大多数模型均没有考虑到 QoS 数据的动态性。

(2) 服务 QoS 属性权重确定不合理。目前主要有两种确定权重的方法,即主观赋权模式和客观赋权模式。在主观赋权模式下,服务的 QoS 权重完全由用户确定,这虽然体现了用户的个性化需求,但具有很大的主观随意性,且大多数用户不具有服务领域的相关知识,给定精确的权重值给用户带来了额外负担。相对来说,用户更加倾向于使用偏好顺序来表达其需求,如服务价格 > 响应时间 > 可靠性等。客观赋权模式由客观数据确定,虽然其结果具有较强的数学理论依据,但它忽略了用户的主观偏好。因此,在确定 QoS 权重时同时考虑用户主观偏好和客观数据更符合服务选取的实际情况。

针对以上两个问题,本文首先对服务中 QoS 的动态性进行分析,并设计了区间 QoS 来表示候选服务和用户需求的动态性;然后通过区间数排序的相似度方法对候选服务的 QoS 区间与用户需求的 QoS 区间进行比较,并利用逼近理想点的多属性决策方法求解出 QoS 指标客观权重;在此基础上结合用户主观偏好算出 QoS 指标综合权重,最终以推荐度的形式对所有候选服务进行排序,以提高服务选取的准确性。

2 QoS 属性定义及分析

Web 服务的 QoS 属性涉及多个方面,如响应时间、服务费用、吞吐量、可用性、安全性等,它们分别从不同的角度对 Web 服务的质量进行评估。然而在网络环境中,由于各种不确定性因素的影响,Web 服务的 QoS 属性呈现出动态性。基于此,本文将 Web 服务的 QoS 属性定义为一个四维向量 $Qos = (T, E, A, R)$, 并采用区间来表示其数值的波动范围。下面介绍各个分量的具体含义和区间表示形式。

1) 响应时间 T : 用户发出服务请求到获取到服务所花费的时间。由定义可知,响应时间受到网络运行状况的影响,因此具有动态变化性。其区间形式为 $[\mu_T - \sigma_T^L, \mu_T + \sigma_T^R]$, 其中 μ_T 为均值, σ_T^L 和 σ_T^R 为方差。

2) 信誉度 E : 服务的可信程度。服务的信誉度主要是由用户使用后的评价所决定的。由于服务环境的差异性和动态性,不同用户使用同一服务可能会产生不同的评价。其区间形式

为 $[\mu_E - \sigma_E^L, \mu_E + \sigma_E^R]$ 。

3) 可用性 A : Web 服务被正常调用的概率。服务的可用性也会随着网络环境的好坏而动态变化。其区间形式为 $[\mu_A - \sigma_A^L, \mu_A + \sigma_A^R]$ 。

4) 可靠性 R : Web 服务能够正确响应用户请求的概率。除了服务本身的服务质量以外,网络环境也会影响服务的可靠性。例如网络阻塞导致服务信息在传输过程中丢失或者失效。其区间形式为 $[\mu_R - \sigma_R^L, \mu_R + \sigma_R^R]$ 。

上述 QoS 区间中的 $\mu_M, \sigma_M^L, \sigma_M^R (M = \{T, E, A, R\})$ 的计算过程如下:

$$\mu_M = \frac{1}{n} \sum_{i=1}^n Q_i \quad (1)$$

$$\sigma_M^L = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_i^L - \mu_M)^2}, Q_i^L < \mu_M \quad (2)$$

$$\sigma_M^R = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_i^R - \mu_M)^2}, Q_i^R \geq \mu_M \quad (3)$$

3 区间型 QoS 模型

上述区间可以描述在动态环境中 Web 服务的 QoS 属性的波动情况。用户的 QoS 需求也用区间进行表示。为了刻画 Web 服务 QoS 属性区间与用户 QoS 需求区间的接近程度,引入区间数和相似度的概念。

定义 1 记 $Q = [x^-, x^+] = [x | x^- \leq x \leq x^+]$, 其中 $x^-, x^+ \in \mathbb{R}$, $\mathbb{R}$ 为实数集, 称 Q 为一个区间数, 若 $x^- = x^+$, 则 Q 退化为一个实数。在本文中, 区间数均指正区间数, 即 $0 \leq x^- \leq x^+$ 。

定义 2 若有两个区间数 Q_x 和 Q_y , 其中 $Q_x = [x^-, x^+]$, $Q_y = [y^-, y^+]$, 记 $l_x = x^+ - x^-$, $l_y = y^+ - y^-$, 称

$$p(Q_x \geq Q_y) = \min\{\max(\frac{x^+ - y^-}{l_x + l_y}, 0), 1\}$$

为 $Q_x \geq Q_y$ 的相似度。

根据定义可以证明, $p(Q_x \geq Q_y)$ 有如下性质:

(1) $0 \leq p(Q_x \geq Q_y) \leq 1$, 当且仅当 $y^+ \leq x^-$ 时 $p(Q_x \geq Q_y) = 1$, 当且仅当 $x^+ \leq y^-$ 时 $p(Q_x \geq Q_y) = 0$ 。

(2) 互补性: $p(Q_x \geq Q_y) + p(Q_y \geq Q_x) = 1$ 。

(3) 当且仅当 $x^+ + x^- \geq y^+ + y^-$ 时, 有 $p(Q_x \geq Q_y) \geq \frac{1}{2}$;

特别地, 当且仅当 $x^+ + x^- = y^+ + y^-$ 时, 有 $p(Q_x \geq Q_y) = \frac{1}{2}$ 。

(4) 传递性: 如果 $p(Q_x \geq Q_y) \geq \frac{1}{2}$, 且 $p(Q_y \geq Q_z) \geq \frac{1}{2}$,

则有 $p(Q_x \geq Q_z) \geq \frac{1}{2}$ 。

令用户对某 QoS 属性提出的需求为区间数 $Q_u = [q_u^L, q_u^R]$, 某候选服务的该 QoS 属性区间数为 $Q_c = [q_c^L, q_c^R]$, p 表示相似度。对于效益型 QoS 指标来说, 其区间数值越大越好; 对于成本型 QoS 指标来说, 其区间数值越小越好。则

$$p = \begin{cases} P(Q_c \geq Q_u), & \text{该属性为效益型} \\ P(Q_u \geq Q_c), & \text{该属性为成本型} \end{cases}$$

基于以上性质可知, 可以利用相似度来衡量候选服务提

供的 QoS 属性与用户请求的接近程度。相似度越大,表示该候选服务在这一 QoS 属性上越符合用户提出的需求,被选中的概率越大。

4 考虑用户偏好的服务选取方法

当得到每个候选服务的各个 QoS 属性的相似度以后,结合对应的权重,便可以得到每个候选服务的推荐值,从而对候选服务进行排序和选择。因此,如何得到一组合理的权重便成为正确进行服务选取的关键所在。目前在一些关于服务选取的文献中^[10-12],主要有主观赋权和客观赋权两种方法来确定 QoS 的指标权重。主观赋权方法主要有层次分析法、经验判断法等。但在实际生活中,大多数用户对于服务领域的相关知识缺乏了解,这会增加用户的负担。客观赋权方法主要有最短距离法、熵权计算法等。客观赋权法主要是根据候选服务的客观 QoS 数据进行 QoS 指标偏好的计算,不需要人为经验的介入,但是不能反映用户的主观偏好。

现实生活中,用户更加倾向于对自己偏好的 QoS 指标进行排序来表达需求,例如服务费用>响应时间>可靠性>可用性等。

定义 3 设用户的偏好向量为 $pref$,该向量中的组成元素为 Web 服务的 QoS 指标,用户对 QoS 指标的偏好程度按照向量中元素的排列顺序由前向后递减。例如 $pref=\langle T, E, A, R \rangle$ 表示用户的 QoS 指标偏好的优先顺序依次为响应时间、信誉度、可用性、可靠性。

综上所述,本文提出一种考虑了用户主观偏好的综合权重确定方法,该方法的基本思想是将基本服务 QoS 指标客观权重与用户主观偏好结合起来,生成一个综合偏好权重。该方法分为 3 个步骤:首先利用逼近理想点的多属性决策方法确定每个基本服务的 QoS 指标客观权重;然后根据基本服务的客观偏好与用户主观偏好的相似性计算出每个基本服务的重要性;最后综合所有基本服务的客观权重得到 QoS 指标综合权重。

4.1 基本服务 QoS 指标客观权重的确定

假设一个组合服务 WSC 由 m 个基本服务 WS_i 构成,记为 $WSC=\{WS_1, WS_2, \dots, WS_m\}$,每个基本服务 WS_i 有 n 个候选服务,记为 $WS_i=\{s_1, s_2, \dots, s_n\}$, k 个 QoS 指标记为集合 QCS。由第 3 节中的计算方法,可以得到每个候选服务的各个 QoS 属性的相似度。将每个基本服务下的 n 个候选服务列为一个 $n \times k$ 的矩阵 V_i ,可以得到 m 个矩阵。其中

$$V_i = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1k} \\ p_{21} & p_{22} & \dots & p_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nk} \end{bmatrix}, 1 \leq i \leq m$$

其中, p_{ij} 表示该基本服务的第 i 个候选服务的 QoS 指标 j 的相似度,即与用户需求的接近程度。

目前在解决多维 QoS 属性的服务选取问题中,通常采用多属性决策方法为候选服务打分,从而选择最优的候选服务。传统的多属性决策方法包括简单加权和法、层次分析法、加权

积法等^[21]。由于本文在服务选取中需要同时考虑区间数表示的 QoS 属性值和需求值,不能直接利用传统的多属性决策方法。逼近理想点的多属性决策方法(TOPSIS)以计算复杂度小、可行性高的优点,经常被应用于解决区间数型多属性决策问题。因此,本文利用 TOPSIS 方法,通过建立多目标规划模型来求解基本服务 QoS 指标客观权重,具体方法如下。

设某基本服务的 QoS 指标权重向量为 $W=\langle w_1, w_2, \dots, w_k \rangle$,由 p_{ij} 的定义可知, p_{ij} 越大越符合用户需求。令正理想点为 1,则候选服务 i 与正理想点之间的加权距离为:

$$d_i^+ = \sum_{j=1}^k (1 - p_{ij}) w_k = 1 - \sum_{j=1}^k p_{ij} w_k$$

显然,候选服务越接近正理想点越优,即 d_i^+ 越小,候选服务越优。为此建立多目标决策模型:求解一个权重向量 W ,使得

$$\min d_i^+, 1 \leq i \leq n$$

$$\text{s. t. } \sum_{j=1}^k w_j = 1, w_j \geq 0, 1 \leq j \leq k$$

由于 $d_i^+ \geq 0, 1 \leq i \leq n$,且每个候选服务之间没有优劣之分,因此可以将该多目标优化问题转化为单目标优化问题,即

$$\min d = \sum_{i=1}^n d_i^+$$

$$\text{s. t. } \sum_{j=1}^k w_j = 1, w_j \geq 0, 1 \leq j \leq k$$

通过求解该线性模型,可以计算出某一基本服务 l 关于 QoS 指标的客观权重向量,记为 W_l 。

下面将基本服务的客观权重向量和用户的主观偏好进行结合。由于用户的主观偏好是由偏好向量 $pref$ 表示的,因此需要将基本服务的客观权重向量也转化为偏好向量的形式。

定义 4(客观偏好向量) 指基本服务的候选服务集的 QoS 指标偏好。由式(1)计算得出某一基本服务 l 的 QoS 指标权重向量 $W_l=\langle w_{l1}, w_{l2}, \dots, w_{lk} \rangle$,则其对应的客观偏好向量为 $pref_l=\langle C_{i1}, C_{i2}, \dots, C_{ij}, \dots, C_{ik} \rangle, C_{ik} \in QCS$,其中各 QoS 指标的权重值满足 $w_{i1} \geq w_{i2} \geq \dots \geq w_{ik}$,即按照计算出来的权重从大到小将 QoS 指标进行重新排序得出。

4.2 基本服务的重要性权重的确定

已知组合服务中有 m 个基本服务,则根据 4.1 节可以分别计算得到这 m 个基本服务的 QoS 权重向量 $W_l(1 \leq l \leq m)$,以及对应的客观偏好向量 $pref_l(1 \leq l \leq m)$ 。用户的主观偏好向量为 $pref$ 。由经验可知,如果某一基本服务的客观偏好向量与用户的主观偏好向量接近甚至相等,则说明该基本服务更加符合用户的心理预期,因此其在综合 QoS 权重的确定上应该占有更大的比例。本文把这种比例称为基本服务的重要性权重,则 m 个基本服务的重要性权重可组成一个向量,记为向量 $U=\langle u_1, u_2, \dots, u_m \rangle$ 。

如何计算两种偏好向量的相似性,是得到每个基本服务重要性权重的关键。本文采用了文献[14]中提出的一种基于偏移量的相似程度度量方法,其基本思路是用客观偏好向量中 QoS 指标元素相对于主观偏好向量中相应元素的偏移程度来刻画相似性。由于偏好向量具有有序性,因此不同位置的偏移在相似性度量中所占比重递减。

如假设用户主观偏好向量为 $pref=\langle \text{响应时间 } T, \text{信誉度 } E, \text{可用性 } A, \text{可靠性 } R \rangle$, 客观偏好向量为 $pref_i=\langle \text{可用性 } A, \text{响应时间 } T, \text{可靠性 } R, \text{信誉度 } E \rangle$, 可见相对于客观偏好向量, 响应时间 T 向后偏移了 1 位, 信誉度 E 向后偏移了 2 位, 而由于用户对于响应时间 T 的重视程度大于信誉度 E , 因此在计算相似性度量时响应时间 T 的偏移程度所占比重大于信誉度 E 。

4.3 计算 QoS 指标综合权重

由 4.1 节可以计算出一系列基本服务关于 QoS 指标的客观权重 $W_i=\langle w_{i1}, w_{i2}, \dots, w_{ik} \rangle (1 \leq i \leq m)$, 将其列为一个客观权重矩阵 W , 如下:

$$W = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_m \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1k} \\ w_{21} & w_{22} & \dots & w_{2k} \\ \dots & \dots & \dots & \dots \\ w_{m1} & w_{m2} & \dots & w_{mk} \end{bmatrix}$$

由 4.2 节计算出一系列基本服务的重要性权重向量 $U=\langle u_1, u_2, \dots, u_m \rangle$ 。

将重要性权重向量乘以客观权重矩阵, 便可以得到考虑用户偏好的综合 QoS 指标权重, 记为 $\Omega=\langle \Omega_1, \Omega_2, \dots, \Omega_k \rangle$ 。

$$\Omega = U * W = \langle u_1, u_2, \dots, u_m \rangle * \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_m \end{bmatrix}$$

$$= u_1 W_1 + u_2 W_2 + \dots + u_m W_m$$

$$= \langle \Omega_1, \Omega_2, \dots, \Omega_k \rangle$$

4.4 服务排序推荐

定义 5(推荐度) 综合考虑客观数据和用户主观偏好后, 得到候选服务的综合评价值。在 4.1 节中得到了一系列基本服务的相似度矩阵 $V_i (1 \leq i \leq m)$, 将其乘以综合 QoS 指标权重, 便得到每一个基本服务所对应的推荐度矩阵, 记为 X_i 。

$$X_i = V_i * \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \vdots \\ \Omega_k \end{bmatrix} = \begin{bmatrix} p_{11} & \dots & p_{1k} \\ \vdots & \vdots & \vdots \\ p_{n1} & \dots & p_{nk} \end{bmatrix} * \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \vdots \\ \Omega_k \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

对于组合服务 $WSC=\{WS_1, WS_2, \dots, WS_m\}$ 来说, 其每个基本服务 $WS_i (1 \leq i \leq m)$ 都有 n 个候选服务 $s_i (1 \leq i \leq n)$, 每个候选服务的推荐度为 x_i , 将其按照由大到小的顺序进行排列, 并用链表形式进行存储, 如图 1 所示。

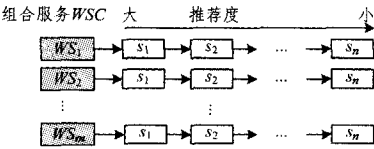


图 1 服务推荐链表

利用服务推荐链表进行服务选取时, 可以设定一个推荐度阈值, 根据这个阈值可以剔除掉推荐度低的候选服务, 从而缩小候选服务空间, 进一步提高服务选取的准确性。

5 实验分析

本文通过仿真实验来模拟动态环境下本文算法的有效性。实验采用了公共有效数据集 QWS, 它所有的数据均来自于互联网上的公共 Web 服务。该数据集包括了 2500 个 Web 服务及其对应的 9 个 QoS 属性值, 如响应时间、可用性、可靠性等, 对该数据集更详细的描述请见文献[18-20]。本文从 QWS 数据集中选取了 3 类基本服务, 每类基本服务的候选服务数量为 5 个, 并从 9 个 QoS 属性中选取出响应时间(秒)、信誉度(1-9 标度)、可用性(百分比)和可靠性(百分比)这 4 个 QoS 属性。实验最终结果为运行 50 次后的平均值。实验软硬件环境为: Pentium Dual 2. 6GHz, 4. 0GB RAM, Windows XP SP3, MATLAB 2009a, Java 1. 8。

本文设定用户的主观偏好排序为: 响应时间 > 可用性 > 可靠性 > 信誉度。用户主观需求通过人为设定。由于本实验主要验证基于动态 QoS 的服务选取问题, 因此表 1 仅仅列举出 QWS 数据集中的 一个基本服务类及它对应的候选服务集进行说明。

表 1 Web 服务 QoS 需求及其监测值

指标	时间 T/s [T ^L , T ^R]	可用性 A/% [A ^L , A ^R]	可靠性 R/% [R ^L , R ^R]	信誉度 E(1-9) [E ^L , E ^R]	用户 满意度/%
用户 需求	[0, 3]	[70, 100]	[80, 100]	[7, 9]	
服务 1	[0. 53, 1. 42]	[85. 3, 90. 6]	[85. 5, 88. 6]	[5, 7]	97. 9
服务 2	[1. 34, 2. 23]	[80. 1, 85. 5]	[80. 1, 83. 9]	[7, 8]	89. 7
服务 3	[1. 14, 1. 61]	[83. 5, 86. 9]	[82. 3, 85. 3]	[5, 7]	94. 2
服务 4	[2. 78, 3. 84]	[76. 2, 78. 5]	[78. 3, 81. 6]	[6, 7]	86. 3
服务 5	[2. 81, 3. 62]	[72. 3, 76. 5]	[73. 2, 82. 3]	[8, 9]	83. 2

根据表 1 中的数据, 利用本文的服务选取方法, 可计算出 QoS 指标综合权重为 $\Omega=\langle 0. 32, 0. 24, 0. 21, 0. 23 \rangle$, 各候选服务的推荐度为 $X_i=\langle 0. 583, 0. 467, 0. 512, 0. 414, 0. 289 \rangle$, 即服务排序为 $s_1 \phi s_3 \phi s_2 \phi s_4 \phi s_5$, 符号“ ϕ ”表示优于, 即 s_1 为最优服务, s_5 为最差服务。这与 QWS 数据集中用户满意度调查的排序结果一致。

在此基础上, 设计 4 个场景来验证本文基于动态 QoS 的服务选取模型的有效性。

(1) 在用户 QoS 需求、用户主观偏好不变的情况下, 对某候选服务 i 的各项 QoS 指标逐渐进行优化, 即对于效益型指标其值逐渐增大, 对于成本型指标其值逐渐缩小, 从而提高其与用户 QoS 需求的接近程度, 其他候选服务保持不变。此时, 候选服务 i 的推荐度应该增大, 其排名顺序应该提前。

为验证上述假设, 将当前最差服务 s_5 的各项 QoS 指标值逐渐进行优化, X_i 的变化情况如图 2 所示。

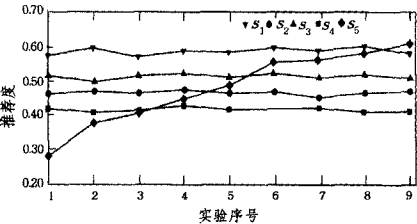


图 2 QoS 指标值优化时推荐度的变化

由图 2 可知,随着候选服务 s_5 各项指标的不断优化,其推荐度也不断增大,从而其排名也不断上升,并最终由最差服务变为最优服务。由于 QoS 指标值的变化可能会引起综合 QoS 指标权重 Ω 的变化,因此其他候选服务的推荐度也会出现小幅度的波动。

(2)设用户 QoS 需求和所有候选服务的 QoS 指标值保持不变,而用户对某一 QoS 指标的重视程度逐渐增大,即该指标在用户的主观偏好向量中的排名逐渐上升,而其他指标重视程度不变或减小。此时,对该指标具有较大优势的候选服务的推荐度会逐渐增大,相反,对于该指标不具有优势的候选服务的推荐度会逐渐减小。

为验证上述假设,将原用户主观偏好排序为响应时间>可用性>可靠性>信誉度,现增大用户对信誉度的重视程度。每次实验使信誉度在主观偏好排序中提升一位, X_i 的变化情况如图 3 所示。

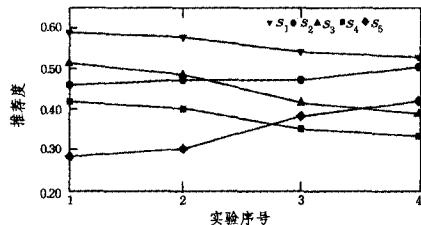


图 3 用户主观偏好变化时推荐度的变化

如图 3 所示,服务排序由之前的 $s_1 \phi s_3 \phi s_2 \phi s_4 \phi s_5$ 变为 $s_1 \phi s_2 \phi s_5 \phi s_3 \phi s_4$ 。随着用户对信誉度重视程度的增加,服务 s_2 和 s_5 的推荐度不断增长。由于服务 s_5 在信誉度指标上比服务 s_2 更具有优势,因此其推荐度的增长速度比服务 s_2 更快,最终服务 s_5 的推荐度排名上升了 2 位,服务 s_2 的推荐度排名上升了 1 位。相反,由于服务 s_1, s_3, s_4 的信誉度指标比较低,其推荐度逐渐减小。

(3)为了验证在基本服务类增多的情况下本文算法的准确性,将从 QWS 数据集选取的基本服务由 3 类逐个扩展到 7 类,并分别采用本文算法和文献[8]中的服务选取算法进行仿真实验,判断两种算法的服务排序结果是否与实际用户满意度的调查结果一致。每轮实验分别进行 100 次,最终计算出服务排序结果的出错率,如图 4 所示。

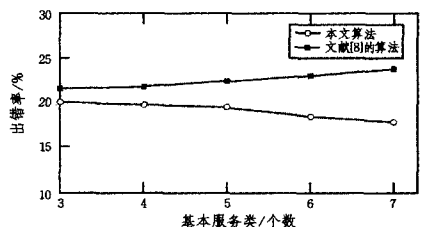


图 4 基本服务类增多时两种算法的出错率变化

从图 4 中可以看出,随着基本服务类的增多,本文算法的出错率在不断下降,而文献[8]中服务选取算法的出错率在不断上升。

由于本文算法综合考虑了基本服务 QoS 指标客观权重和用户主观偏好,因此当基本服务类增多时,算法能够融合更

多关于基本服务的 QoS 指标客观权重的信息,从而计算出的 QoS 指标综合权重也就更加合理,提高了服务排序结果的准确性。而文献[8]中的服务选取算法仅仅考虑用户的主观偏好权重,因此当基本服务类增多时,其得出的 QoS 指标权重具有更大的主观随意性,服务排序结果的出错率也随之增加。

(4)为了验证在动态环境中本文算法的有效性,采用 M/M/1 排队模型来模拟动态环境的不稳定性。按照表 1 中 Web 服务参数的设置,产生 30 组这样的候选服务,每组有 500 个顾客,到达率为 $\lambda=0.2$,服务率为 $\mu=0.4$ 。分别采用本文算法和文献[1]中的 NPT_QoSDDMM 算法进行仿真实验,判断两种算法的服务排序结果是否与实际用户满意度调查结果 $s_1 \phi s_3 \phi s_2 \phi s_4 \phi s_5$ 一致。取前 20 组实验数据,其服务推荐度如图 5 所示,其中实线代表按照本文算法得到的排序结果,虚线代表按照 NPT_QoSDDMM 算法得到的排序结果。

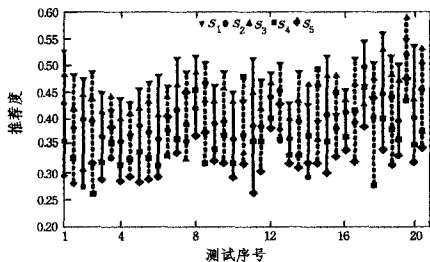


图 5 在动态环境下两种算法的服务排序结果

从图 5 中可以看出,由于动态环境的不稳定性,每个候选服务的推荐度也在不停地变化,导致每次实验的服务排序结果也不尽相同。具体情况如表 2 所列。

表 2 模拟动态环境下两种算法的仿真结果

算法	与实际结果一致的测试序号	准确率/%
本文算法	1,3,4,5,6,7,9,10,11,12,13,15,16,18,19,20	80
NPT_QoSDDMM 算法	1,4,5,9,11,13,18	35

从表 2 中可以看出,在 20 次仿真实验中,根据本文算法得到的服务排序结果中有 16 次与实际用户满意度排序结果一致,准确率为 80%;按照 NPT_QoSDDMM 算法,除了第 1,4,5,9,11,13,18 次测试以外,其余测试的服务排序结果均与实际用户满意度排序结果有不同程度的出入,其准确率为 35%。可见,在动态环境中本文算法的抗干扰能力要强于 NPT_QoSDDMM 算法。

本文算法相比于 NPT_QoSDDMM 算法的优势在于,NPT_QoSDDMM 算法在大部分情况下都只是使用均值来比较服务的优劣,而均值仅仅能够反映一个服务在某个 QoS 属性上的平均取值,不能反映出其 QoS 数据的波动性,故在动态环境中其排序结果不一致。而本文算法使用区间 QoS 模型,能够有效克服动态环境的不稳定性,所以其排序结果更加准确。

结束语 本文首先对动态环境下 Web 服务的 QoS 不确定性进行了分析,进而设计了区间 QoS 模型,用相似度的概念来表征候选服务 QoS 属性区间与用户需求值区间的接近程度,并以此为基础用逼近理想点的多属性决策方法来确定每个基本服务的客观 QoS 权重,然后计算基本服务的客观偏

好与用户主观偏好的相似程度来得出基本服务的重要性权重,通过该重要性权重将各基本服务的客观 QoS 权重结合起来计算出 QoS 指标综合权重;最终计算出推荐度对每个基本服务下的候选服务进行排序,这样用户就可以根据排序结果对服务进行选取。实验验证了该服务选取模型的有效性。未来将研究 Web 服务各个 QoS 指标不确定性之间的关联关系,进一步提高服务选取的准确率。

参 考 文 献

- [1] ZHANG Y H, FU F R, RU C. Semantic Web service selection based on QoS ontology[C]//Proceedings of the 2010 International Conference on Information, Network and Automation. Washington, DC, USA; IEEE Computer Society, 2010: 2492-2496.
- [2] WAN C, WANG H B. Uncertainty-aware QoS description and selection model for web services [C]//SCC'07. Salt Lake City: IEEE Press, 2007: 154-161.
- [3] LI L, LIU M, CHENG G Q. A local optimal model of service selection of Muti-QoS based on FAHP[J]. Chinese Journal of Computers, 2015, 38(10): 1997-2010. (in Chinese)
李玲, 刘敏, 成国庆. 一种基于 FAHP 的多维 QoS 局部最优服务选择模型[J]. 计算机学报, 2015, 38(10): 1997-2010.
- [4] ZHU R, WANG H M, FENG D W. Trustworthy services selection base on preference recommendation[J]. Journal of Software, 2011, 22(5): 852-864. (in Chinese)
朱锐, 王怀民, 冯大为. 基于偏好推荐的可信服务选择[J]. 软件学报, 2011, 22(5): 852-864.
- [5] SHEN L M, CHEN Z, LI F. Service selection approach considering the uncertainty of QoS data[J]. Computer Integrated Manufacturing Systems, 2013, 19(10): 2652-2663. (in Chinese)
申利民, 陈真, 李峰. 一种考虑 QoS 数据不确定性的服务选取方法[J]. 计算机集成制造系统, 2013, 19(10): 2652-2663.
- [6] FAN X Q, JIANG C J, WANG J L, et al. Random-QoS-aware reliable Web service composition[J]. Journal of Software, 2009, 20(3): 546-556. (in Chinese)
范小芹, 蒋昌俊, 王俊丽, 等. 随机 QoS 感知的可靠 Web 服务组合[J]. 软件学报, 2009, 20(3): 546-556.
- [7] HU J Q, LI J Z, LIAO G P. A Multi-QoS based local optimal model of service selection[J]. Chinese Journal of Computers, 2010, 33(3): 0526-0534. (in Chinese)
胡建强, 李涓子, 廖桂平. 一种基于多维服务质量的局部最优服务选择模型[J]. 计算机学报, 2010, 33(3): 0526-0534.
- [8] STEPHAN R M, YU H Q, TILLY M. Service selection based on non-funtional properties[C]//Proceedings of International Conference on Service-Oriented Computing (ICSOC 2007). Berlin, Germany: Springer-Verlag, 2009: 128-138.
- [9] YU H Q, STEPHAN R M. A method for automated Web service selection[C]//Proceedings of IEEE Congress on Services. Washington, DC, USA; IEEE Computer Society, 2008: 513-520.
- [10] WU J, CHEN L, DENG S G, et al. QoS-Skyline based dynamic service selection [J]. Chinese Journal of Computers, 2010, 33(11): 2136-2146. (in Chinese)
吴健, 陈亮, 邓水光, 等. 基于 Skyline 的 QoS 感知的动态服务选择[J]. 计算机学报, 2010, 33(11): 2136-2146.
- [11] WEN T, SHENG G J, GUO Q, et al. Web service composition based on modified particle swarm optimization[J]. Chinese Journal of Computers, 2013, 36(5): 1031-1046. (in Chinese)
温涛, 盛国军, 郭权, 等. 基于改进粒子群算法的 Web 服务组合[J]. 计算机学报, 2013, 36(5): 1031-1046.
- [12] PAN J, XU F, LU J. Reputation-based recommender discovery approach for service selection[J]. Journal of Software, 2010, 21(2): 388-400. (in Chinese)
潘静, 徐锋, 吕建. 面向可信服务选取的基于声誉的推荐者发现方法[J]. 软件学报, 2010, 21(2): 388-400.
- [13] LI F, SHEN L M, SI Y L, et al. A trust forecasting model based on entity context and time tamp[J]. Journal of Electronics & Information Technology, 2011, 33(5): 1217-1223. (in Chinese)
李峰, 申利民, 司亚利, 等. 一种基于实体上下文和时间戳的信任预测模型[J]. 电子与信息学报, 2011, 33(5): 1217-1223.
- [14] ZHANG X Y. Research on QoS evaluation and service selection in Web service composition[D]. Nanjing: Nanjing University, 2011. (in Chinese)
张旭云. Web 服务组合中 QoS 评估及服务选择的关键技术研究[D]. 南京: 南京大学, 2011.
- [15] ZHU X L, WANG B. Web service selection based on Interval QoS[J]. Journal of Beijing University of Posts and Telecommunications, 2011, 34(4): 0080-0084. (in Chinese)
祝希路, 王柏. 支持区间型 QoS 的 Web 服务选择[J]. 北京邮电大学学报, 2011, 34(4): 0080-0084.
- [16] WANG S G, SUN Q B, ZHANG G W, et al. Uncertain QoS-aware skyline service selection based on cloud model[J]. Journal of Software, 2012, 23(6): 1397-1412. (in Chinese)
王尚广, 孙其博, 张光卫, 等. 基于云模型的不确定性 QoS 感知的 Skyline 服务选择[J]. 软件学报, 2012, 23(6): 1397-1412.
- [17] ALRIFAI M, SKOUTAS D, RISSE T. Selecting Skyline services for QoS-based Web service compotion[C]//Proc. of the 19th Int'l Conf. on World Wide Web(WWW 2010). 2010: 11-20.
- [18] AL-MASRI E, MAHMOUD Q H. Discovering the best Web service[C]//Proc. of the 16th Int'l Conf. on World Wide Web (WWW 2007). 2007: 1257-1258.
- [19] AL-MASRI E, MAHMOUD Q H. QoS-Based discovery and ranking of Web services[C]//Proc. of the 16th Int'l Conf. on Computer Communications and Networks (ICCCN 2007). 2007: 529-534.
- [20] AL-MASRI E, MAHMOUD Q H. Investigating Web services on the World Wide Web[C]//Proc. of the 17th Int'l Conf. on World Wide Web(WWW 2008). 2008: 795-804.
- [21] XIA Y Q, WU Q Z. A technique of order preference by similarity to ideal solution for hybrid multiple attribute decision making problems[J]. Journal of Systems Engineering, 2010, 19(6): 630-634. (in Chinese)
夏勇其, 吴祈宗. 一种混合型多属性决策问题的 TOPSIS 方法[J]. 系统工程学报, 2010, 19(6): 630-634.