

一种在用户偏好不确定情况下的 Web 服务选择方法*

文俊浩^a, 秦佳^a, 柳玲^b

(重庆大学 a. 计算机学院; b. 软件学院, 重庆 400044)

摘要: 首先由用户对各个 QoS 属性的偏好给出语言描述及其不确定度; 然后通过查找对照表将其换算成各个 QoS 属性的权重系数; 最后使用 QoS 属性值和权重系数进行候选服务的综合评价, 得到最接近满足用户不确定偏好的候选服务。模拟实验结果证明了该方法的有效性。

关键词: 服务选择; 服务质量; 用户偏好; 直觉模糊集

中图分类号: TP301.6 文献标志码: A 文章编号: 1001-3695(2010)06-2147-03

doi:10.3969/j.issn.1001-3695.2010.06.043

Approach to Web services selection under uncertain user preferences

WEN Jun-hao^a, QIN Jia^a, LIU Ling^b

(a. College of Computer Science, b. College of Software Engineering, Chongqing University, Chongqing 400044, China)

Abstract: Therefore, the linguistic descriptions and their uncertainty of individual QoS attributes were given by the user at first, and then was to lookup the reference table to convert those into weight factors of the various QoS attributes, the last step was to use the values and weight factor of QoS attribute to evaluate the candidate services comprehensively to get the one which was closest to meets the uncertain user preferences. The results of simulation experiment show the effectiveness of the approach.

Key words: Web services selection; QoS(quality of service); user preferences; intuitionistic fuzzy sets

0 引言

服务提供者将其服务发布到 UDDI 服务器上; 当服务消费者需要调用该服务时, 首先向 UDDI 服务器提交功能性需求, 利用 UDDI 服务器提供的目录去搜索该服务, 得到如何调用该服务的信息, 然后根据这些信息去调用服务提供者发布的服务^[1]。然而, 随着 Web 服务应用的逐渐普及, 服务的种类和数量不断增加, 可能使得 UDDI 服务器根据条件返回多个功能相似的服务, 需要服务消费者从中挑选最能够满足需求的候选服务。因此, 就需要利用服务的非功能性属性, 一般即指服务质量 QoS, 对得到的多个功能相似的候选服务进行选择。Web 服务的 QoS 属性主要包括: 时延、响应时间、持续时间、补偿率、惩罚率、信誉度、异常处理、可靠性、有效性等。文献[2~4]对 Web 服务的 QoS 属性进行详细地总结分类和讨论。在此研究基础上, 文献[5~7]指出 Web 服务还具有一类与特定应用领域相关的 QoS 属性。

目前, 已有研究提出了一些基于 QoS 的服务选择方法。文献[5]借助 QoS 本体和 XML 策略语言, 使用服务选择代理在服务消费者和服务提供者之间交换 QoS 信息, 在此基础上建立实现动态服务选择的 Web 服务代理框架(WSAF)。文献[7]利用矩阵表示与具体业务领域相关的 QoS 模型, 并在该模型的基础上提出形式化的动态服务选择方法。借鉴文献[7]的研究, 文献[8]进一步扩展已有描述语义 Web 服务的 WSMO

模型, 加入了 QoS 相关属性, 提出基于 QoS 的语义 Web 服务选择模型。文献[9]从服务组合的需求出发, 根据多属性决策(MCDM)思想, 提出整数规划方法, 使用 Web 服务的多个非功能性属性作为参考和约束条件来聚合 QoS 属性, 结合考虑用户偏好, 在执行路径的每个任务节点对应的一组候选服务中进行选择, 实现组合服务的 QoS 全局最优。

但是, 在上述研究中提出的 QoS 建模方法都基于一个假设: 即其中使用的 QoS 属性值和权重系数都必须以实数的形式明确给出。然而, 这个假设却不能满足实际应用需求, 还会丧失大量有用的数据信息。对于某些 QoS 属性值或权重系数而言, 应该以不确定的描述形式给出, 以增强模型的表达能力。文献[10,11]采用区间型数据描述某些 QoS 属性在某一时间段内的突发最大值和最小值, 同时采用语言型数据描述那些不能以具体数值表示的 QoS 属性; 然后建立适用于服务组合的混合 QoS 模型, 并给出基于该模型的语义 Web 服务组合全局优化决策算法。文献[12]利用直觉模糊集理论对多个候选服务进行评级分类, 根据人对相近事物作出的近似模糊评价进行分析, 将评价分为肯定、否定和不确定三种可能情况, 计算发生概率, 最后得到各个候选服务的综合评价。文献[13,14]分别在信息不完整或无法确定的情况下, 提出基于不确定知识的多层次的综合评价方法。

当前基于模糊表达和模糊评价的服务选择方法大多是集中在关注 QoS 属性的本质特征, 很少涉及从服务消费者角度

收稿日期: 2009-12-30; 修回日期: 2010-01-19 基金项目: “十一五”国家科技支撑计划基金资助项目(2007BAF23B0302)

作者简介: 文俊浩(1969-), 男, 河南临颖人, 教授, 博导, 博士, 主要研究方向为服务计算与面向服务的软件工程(jwhen@cqu.edu.cn); 秦佳(1983-), 男, 硕士研究生, 主要研究方向为在面向服务架构中的 Web 服务组合及其优化、工作流的建模及其优化; 柳玲(1970-), 女, 副教授, 博士, 主要研究方向为服务计算与面向服务的软件工程。

出发研究当用户偏好不确定时的服务选择问题。由于应用上下文环境的复杂性、人的思维方法的模糊性和 QoS 属性的多样性等原因,服务消费者很难精确地表达对各个 QoS 属性的偏好程度。鉴于上述讨论,本文在已有研究的基础上,提出一种新的 Web 服务选择方法,以解决在用户偏好不确定情况下的 Web 服务选择问题。该研究将为进一步满足 Web 服务选择的实际应用需求提供新的思路和方法。

1 直觉模糊集与权重系数

模糊集理论为人类处理模糊信息提供了一套有效的方法,Atanassov^[15]将模糊集理论加以推广,提出了直觉模糊集(intuitionistic fuzzy sets,IFSs)的概念。它从隶属度和非隶属度两个方面描述不确定信息,为信息属性的描述提供了更多的选择方式,在处理不确定信息时具有更强的表现能力,目前已经在模式识别、神经网络、模糊优化等领域得到了很好的应用。本文将利用直觉模糊集理论,研究当用户偏好不确定时的服务选择问题。

定义 1 设 X 是一个非空有限集合,称 $A = \{ (x, \mu(x), \nu(x)) \mid x \in X \}$ 为直觉模糊集。其中: $\mu(x)$ 和 $\nu(x)$ 分别表示 X 中的元素 x 属于 A 的隶属(membership)度 $\mu_A: X \rightarrow [0,1]$ 和非隶属(non-membership)度 $\nu_A: X \rightarrow [0,1]$,且满足条件 $0 \leq \mu_A(x) + \nu_A(x) \leq 1, x \in X$ 。

定义 2 对于任意一个直觉模糊集 A ,称 $\pi_A = 1 - \mu_A(x) - \nu_A(x), x \in X$ 表示 X 中元素 x 属于 A 的不确定度(或称犹豫间隙)。其中 $0 \leq \pi_A(x) \leq 1$, 任取 $x \in X$ 。

定义 3 X 中的元素 x 属于 A 的隶属度与非隶属度组成的有序数对 $[\mu(x), \nu(x)]$ 称为直觉模糊数。可以将 X 上的直觉模糊集 A 看做是直觉模糊数的集合,记为 $A = \{ [\mu(x), \nu(x)] \mid x \in X \}$ 。

例如,用户对 Web 服务的某一 QoS 属性的偏好记为 $[0.5, 0.4]$,即表示隶属度 $\mu = 0.5$,非隶属度 $\nu = 0.4$ 。根据定义 2,可以得到不确定度 $\pi = 0.1$ 。由定义 3 推知 $\mu_A(x) + \nu_A(x) + \pi_A(x) = 1$,它表示描述直觉模糊集特征的参数 x 的所有取值都落在斜平面 $\mu_A(x) + \nu_A(x) + \pi_A(x) = 1$ 上。

定义 4 权重系数是表示单个 QoS 属性在所有 QoS 属性中的重要程度。它表示在其他 QoS 属性不变的情况下,这个 QoS 属性的变化,对综合评价结果的影响。

定义 5 设 x_1, x_2, \dots, x_n 为候选服务的一组 QoS 属性, w_1, w_2, \dots, w_n 为这组 QoS 属性对应的权重系数, $w_j \in [0, 1], \sum w_j = 1, j = 1, 2, \dots, n$,则加权平均和 $S = x_1 w_1 + x_2 w_2 + \dots + x_n w_n = \sum x_j w_j$ 。本文使用加权平均和 S 表示候选 Web 服务综合评价。

例如,一组功能相似的候选 Web 服务 ws_1, ws_2 , 具有一组 QoS 属性,依次为时延、可靠性、信誉度,与其对应的权重系数 $(w_1, w_2, w_3) = (0.2, 0.3, 0.5)$ 。第一个服务 ws_1 的 QoS 属性组 $(x_1, x_2, x_3) = (10, 20, 30)$,第二个服务 ws_2 的 QoS 属性组 $(x_4, x_5, x_6) = (20, 30, 10)$ 。根据定义 5,计算两个服务的 QoS 属性组的加权平均和,分别为 $S_{ws_1} = 10 \times 0.2 + 20 \times 0.3 + 30 \times 0.5 = 23; S_{ws_2} = 20 \times 0.2 + 30 \times 0.3 + 10 \times 0.5 = 18$ 。因为 $S_{ws_1} > S_{ws_2}$,所以推知, ws_1 的性能优于 ws_2 。

2 在用户偏好不确定情况下的 Web 服务选择

设有一组功能相似的候选服务 $WS = (ws_1, ws_2, \dots, ws_m)$, 且具有一组 QoS 属性 $C = (c_1, c_2, \dots, c_n)$ 。其中 $m, n \in N$ 。不失一般性,假设 QoS 属性值均大于 0,有关模糊 QoS 属性的相关研究不在本文讨论范围。设矩阵 $Q_{mn} = (q_{ij})_{m \times n}, i = 1, 2, \dots, m, j = 1, 2, \dots, n$ 。其中 q_{ij} 表示在服务组 WS 中的第 i 个候选服务的第 j 个 QoS 属性值。矩阵 Q_{mn} 的表达式为

$$Q_{mn} = \begin{matrix} ws_1 \\ ws_2 \\ \vdots \\ ws_m \end{matrix} \begin{pmatrix} c_1 & c_2 & \dots & c_n \\ q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{m1} & q_{m2} & \dots & q_{mn} \end{pmatrix}$$

由于用户不能确定对各个 QoS 属性的偏好程度,根据用户不确定的偏好得到近似合理的 QoS 属性的权重系数,是本文研究的重点内容。

2.1 归一化处理

由于各个属性的物理意义不完全相同,计量单位也不一定相同,从而使得属性值的量纲和数量级可能不同。为了便于数据的评估和比较,需要将矩阵 Q_{mn} 中的 QoS 属性值进行无量纲化(或称归一化)处理。使用的处理函数如下所示:

$$q'_{ij} = \begin{cases} (q_{ij} - \min_i q_{ij}) / (\max_i q_{ij} - \min_i q_{ij}) (q_{ij}), & \text{当 } q_{ij} \text{ 是效益型属性值} \\ (\max_i q_{ij} - q_{ij}) / (\max_i q_{ij} - \min_i q_{ij}) (q_{ij}), & \text{当 } q_{ij} \text{ 是成本型属性值} \end{cases}$$

其中: $\min_i q_{ij}$ 和 $\max_i q_{ij}$ 分别表示在服务组 WS 中第 j 个 QoS 属性 c_j 的最小值和最大值。效益型的属性值越大,表明属性质量越优;反之,成本型的属性值越小,表明属性质量越优。

2.2 确定权重系数

由于用户偏好不确定,无法直接得到 QoS 属性的权重系数。根据直觉模糊集的知识,首先给出 QoS 属性偏好程度的语言描述对照表,然后由用户根据实际应用需求对 QoS 属性的偏好进行语言描述并给出语言描述的不确定度;接着通过语言描述查找对照表确定与 QoS 属性对应的直觉模糊数表达式,最后将不确定度代入表达式把用户的语言描述转换为 QoS 属性的权重系数,其转换函数为: $w_j = \mu(q_j) - \nu(q_j) \times \pi(q_j)$, 其中 $w_j \in [-1, 1], j = 1, 2, \dots, n$ 。QoS 属性偏好程度的语言描述对照表的形式如表 1 所示^[12]。本文后面的实验部分将以此表为例。

表 1 QoS 属性偏好程度的语言描述对照表

QoS 属性偏好程度 的语言描述	直觉模糊数	QoS 属性偏好程度 的语言描述	直觉模糊数
非常重要(VI)	$[0.9, 0.1 - \pi]$	不重要(U)	$[0.3, 0.7 - \pi]$
重要(I)	$[0.7, 0.3 - \pi]$	非常重要(VU)	$[0.1, 0.9 - \pi]$
一般(M)	$[0.5, 0.5 - \pi]$	未知(N)	$[0.0, 0.0]$

在语言描述对照表中,直觉模糊数表达式中的未知变量 π 表示用户在对 QoS 属性偏好程度给出语言描述时的不确定度,不同的用户给出的不确定度是不完全相同的。当用户忽略不确定度时,即有 $\pi(q_j) = 0$,可以得到 $w_j = \mu(q_j) - \nu(q_j) \times 0 = \mu(q_j)$ 。根据定义 5, $w_j \in [0, 1], \sum w_j = 1, j = 1, 2, \dots, n$ 。因此,还需要对得到的权重系数进行转换处理,使得其值落在区间 $[0, 1]$ 上,其转换函数为 $\tilde{w}_j = w_j / (w_1 + w_2 + \dots + w_n) (j = 1, 2, \dots, n)$ 。

2.3 综合评价

最后由定义 5 中的加权平均和公式,得到候选服务的综合评价,其综合评价函数为: $S_{ws_i} = \sum_{j=1}^n q_j \tilde{w}_j, i = 1, 2, \dots, m$ 。其中 $\max(S_{ws_i})$, 对应的候选服务 WS_i 即为最终所求,即在这组候选服务中,综合评价结果最高的服务 WS_i 的性能优于同组的其他候选服务,最能够符合用户对实际应用需求的不确定偏好。

3 实验分析

根据本文前面的研究内容,使用 MATLAB 进行数据模拟,通过计算得到最接近符合用户不确定偏好的候选服务。在实验中,生成一组功能相似的候选服务 $WS = (ws_1, ws_2, ws_3, ws_4, ws_5)$, 且具有一组 QoS 属性 $C = (c_1, c_2, c_3, c_4, c_5)$, 依次为价格、时延、信誉度、可靠性、有效性。QoS 属性偏好程度的语言描述对照表参见表 1。实验步骤如下:

a) 生成一组功能相似的候选服务 $WS = (ws_1, ws_2, ws_3, ws_4, ws_5)$, 这组候选服务具有一组 QoS 属性 $C = (c_1, c_2, c_3, c_4, c_5)$, 其 QoS 属性值如表 2 所示。

表 2 候选服务 QoS 属性值

候选 Web 服务	价格 c_1	时延 c_2	信誉度 c_3	可靠性 c_4	有效性 c_5
ws_1	61.2420	15.0059	49.4032	0.0675	3.5788
ws_2	60.8259	14.9635	49.6966	0.0985	4.2308
ws_3	61.5887	14.8336	49.1894	0.0966	3.9619
ws_4	61.2308	14.8866	49.3684	0.2934	4.3957
ws_5	61.3930	15.0091	49.6477	0.1323	3.9805

b) 为了便于数据的评估和比较,对在步骤 a) 中得到的 QoS 属性值进行归一化处理,处理结果如表 3 所示。

表 3 候选服务 QoS 属性值的归一化处理结果

候选 Web 服务	价格 c_1^*	时延 c_2^*	信誉度 c_3^*	可靠性 c_4^*	有效性 c_5^*
ws_1	0.4545	0.0180	0.4216	0	0
ws_2	1.0000	0.2598	1.0000	0.1374	0.7981
ws_3	0	1.0000	0	0.1287	0.4690
ws_4	0.4982	0.6978	0.3529	1.0000	1.0000
ws_5	0.2565	0	0.9037	0.2867	0.4918

c) 模拟产生用户对各个 QoS 属性偏好程度的语言描述和用户语言描述的不确定度,然后根据 2.2 节的方法计算出各个 QoS 属性的权重系数,如表 4 所示。

表 4 候选服务 QoS 属性的用户偏好转换

QoS 属性	偏好程度	不确定度	权重系数
价格 c_1^*	不重要	0.005 9	0.149 2
时延 c_2^*	非常重要	0.613 1	0.612 3
信誉度 c_3^*	一般	0.189 7	0.222 4
可靠性 c_4^*	非常不重要	0.083 2	0.016 2
有效性 c_5^*	未知	0	0

d) 根据步骤 b) 和步骤 c) 得到的结果,计算候选服务 $ws_1, ws_2, ws_3, ws_4, ws_5$ 综合评价,计算结果如表 5 所示。

表 5 候选服务的综合评价

候选 Web 服务	综合评价结果	候选 Web 服务	综合评价结果
ws_1	0.172 6	ws_4	0.591 9
ws_2	0.532 9	ws_5	0.243 9
ws_3	0.614 4		

e) 从步骤 d) 得到的计算结果可以看出,候选服务的性能由优至劣的顺序依次为 $ws_3 > ws_4 > ws_2 > ws_5 > ws_1$ 。由此可知,服务 ws_3 为最能够符合用户不确定偏好的候选服务。

4 结束语

用户偏好不确定是在服务选择过程中存在的客观问题,本文在已有研究的基础上,提出一种近似满足用户不确定偏好的服务选择方法。在本文讨论过程中,QoS 属性偏好程度的语言描述对照表是以表 1 为例的,但是对照表并不是固定的。如何根据实际情况确定恰当的语言描述对照表,需要进行专门的讨论。在下一步的研究工作中,将同时考虑 QoS 属性中的模糊属性,然后将用户偏好不确定问题扩展到 Web 服务组合中。此外,分析在选择过程中不确定的表达误差对真实结果的影响也是值得研究的问题。

参考文献:

- [1] 岳昆,王晓玲,周傲英. Web 服务核心支撑技术—研究综述[J]. 软件学报,2004,15(3):428-442.
- [2] RAN S. A model for Web services discovery with QoS[C]//Proc of ACM SIGecom Exchanges. New York:ACM Press,2003:1-10.
- [3] O'SULLIVAN J, EDMOND D, HOFSTEDE A. What's in a service? [J]. Distributed and Parallel Databases, 2002, 12(2-3):117-133.
- [4] MENASCE D A. QoS issues in Web services [J]. IEEE Internet Computing, 2002, 6(6):72-75.
- [5] MAXIMILIEN E M, SINGH M P. A framework and ontology for dynamic Web services selection [J]. IEEE Internet Computing, 2004, 8(5):84-93.
- [6] 杨文军,李涓子,王克宏. 领域自适应的 Web 服务评价模型 [J]. 计算机学报,2005,28(4):514-523.
- [7] LIU Y, NGU A H, ZENG L. QoS computation and policing in dynamic Web service [C]//Proc of the 13th International World Wide Web Conference on Alternate Track Papers & Posters. New York: ACM Press, 2004:66-73.
- [8] WANG X, VITVAR T, KERRIGAN M, et al. A QoS-aware selection model for semantic Web services [C]//Proc of Service-Oriented Computer-ICSOC 2006. Heidelberg: Springer, 2006:390-401.
- [9] ZENG L, BENATALLAH B, NGU A H, et al. QoS-Aware middleware for Web services composition [J]. IEEE Trans on Software Engineering, 2004, 30(5):311-327.
- [10] 李祯,杨放春,苏森. 基于模糊多属性决策理论的语义 Web 服务组合算法 [J]. 软件学报, 2009, 20(3):583-596.
- [11] 杨放春,苏森,李祯. 混合 QoS 模型感知的语义 Web 服务组合策略 [J]. 中国科学 E 辑:信息科学, 2008, 38(10):1697-1716.
- [12] WANG P. QoS-aware Web services selection with intuitionistic fuzzy set under consumer's Vague Perception [J]. Expert System with Applications, 2009, 36(3):4460-4466.
- [13] 何贵青,陈世浩,田,等. 多传感器图像融合效果综合评价研究 [J]. 计算机学报, 2008, 31(3):486-492.
- [14] 王飞,邹仕洪,陈山枝,等. 基于模糊数学的 Web 服务 QoS 建模 [J]. 计算机应用研究, 2007, 24(4):214-216.
- [15] ATANASSOV K T. Intuitionistic fuzzy sets [J]. Fuzzy Sets and Systems, 1986, 20(1):87-96.