

一种基于 QoS 综合匹配的 Web 服务选择方法

张翠肖, 徐 丽, 冲 莉

(石家庄铁道大学 信息科学与技术学院 河北 石家庄 050043)

摘要: 当前基于 QoS 的 Web 服务选择方法, 缺乏同时考虑 QoS 语义和数值处理及用户的个性化需求。基于多属性理论, 提出一种 QoS 综合匹配方法用于筛选功能相同但 QoS 属性不同的候选 Web 服务。本方法包含 QoS 语义和数值综合匹配、构建多属性决策矩阵、个性化服务选择这几个阶段。实验验证本方法推荐结果符合用户需求。

关键词: 语义 Web 服务; 语义相似度; QoS 数值相似度

中图分类号: TP311 **文献标识码:** A **文章编号:** 2095-0373(2013)04-0098-05

1 QoS 驱动的 Web 服务选择方法

随着功能相同而非功能属性(QoS)不同的 Web 服务的增多, 如何从中准确地选出满足用户需求的最优服务是研究领域的一个热点。很多学者开始研究以 QoS 属性作为选择服务指标的各种方法^[1-4], 并取得了一定的成果。

最初的服务选择方法是基于 QoS 属性值计算的, 通过构建数学模型计算出每个候选服务的属性值, 选择属性值最大的服务作为最佳服务^[1], 该方法缺少语义支持不利于实现服务的自动发现。为解决上述问题, 研究者开始将语义中的本体技术应用到服务选择研究中^[2], 致力于构建 QoS 本体, 设计基于本体的服务选择方法。L Taher et al 提出基于 QoS 相似度的 Web 服务选择方法^[3-4], 但它只是考虑了服务 QoS 数值方面的匹配, 缺乏考虑 QoS 参数概念的语义相似度。

上述方法的缺点导致用户花费时间和精力却筛选不到更符合自身需求的服务。对语义和数值进行综合考量并考虑服务请求者的个性化需求, 提出了一种基于语义和数值综合匹配的 Web 服务选择方法, 包含 QoS 语义和数值综合匹配、构建多属性决策矩阵、个性化服务选择这几个阶段。

2 QoS 语义和数值综合匹配

考虑到 QoS 参数包括 QoS 概念和 QoS 数值两部分, 因此 QoS 的综合相似度是由 QoS 属性概念的语义相似度和 QoS 属性的数值相似度共同决定的。首先进行了两个 QoS 参数的语义匹配, 当语义上存在可比性时, 才进行下一步的数值处理和数值匹配。

2.1 QoS 语义匹配

在 QoS 本体树中语义相似度可以用来衡量两个概念的相关程度。语义相似度随着语义距离的增大下降速度很快, 采用下式为语义相似度的计算公式^[5]

$$S(c_i, c_j) = 1/e^{dis(c_i, c_j)} \quad (1)$$

式中, $dis(c_i, c_j)$ 为本体树中概念词 c_i 与概念词 c_j 之间的语义距离; $S(c_i, c_j)$ 为本体树中概念词 c_i 与概念词 c_j 之间的语义相似度。式(1)取值在 $[0, 1]$ 区间内, α 为可调节参数。如果 $c_i = c_j$ 则 $dis = 0$, $S(c_i, c_j) = 1$; 如果 $c_i \neq c_j$ 则 $dis(c_i, c_j)$ 采用通用余弦相似度度量距离算法^[6]计算, 如果 $dis(c_i, c_j) = \infty$, $S(c_i, c_j) = 0$ 。

在实际选择过程中, 很多不相关 QoS 属性参数作为本体树中的节点存在有限语义距离, 导致两者之间语义相似度不为 0。如“吞吐率”和“价格”在语义概念上来看是没有任何比较意义的, 但是在本体树

收稿日期: 2013-03-15

作者简介: 张翠肖 女 1969 年出生 教授

基金项目: 河北省自然科学基金(F2013210109); 河北省科学技术研究发展计划项目(10213516D)

中,若二者存在语义距离,这导致语义相似度不为0,因此设计一个 QoS 属性参数之间的相容性参数 u 来描述两个语义概念是否有比较的必要性。 u 的取值取决于 $dis(c_i, c_j)$ 是否大于由 QoS 本体的设计者根据实际经验给出的语义距离的临界值 L ,当两个 QoS 属性参数之间的语义距离大于 L 时 $\mu = 0$; 否则 $\mu = 1$ 。因此,语义综合相似度函数 $SSem(c_i, c_j)$ 为相容性参数 u 和语义相似度 $S(c_i, c_j)$ 相乘的积,如公式(2)所示。

$$SSem(c_i, c_j) = uS(c_i, c_j) \quad (2)$$

式中, $S(c_i, c_j)$ 为两个 QoS 属性参数之间的相似度, u 为两个 QoS 属性参数之间的相容性参数。引入此公式后,避免了那些没有比较意义的 QoS 属性比较过程,提高了效率。

2.2 QoS 数值匹配

本文支持单值、模糊单值、区间值三种 QoS 数值表达方式。如,响应时间为单值属性;安全性描述为一个集合{高,中,低}对应的数值描述为{3,2,1},为模糊单值属性;价格区间100元以内,为区间型属性。本文将数值进行统一规划归成精确型单值数据来表示。

2.2.1 QoS 数据去模糊化

不同 QoS 属性参数特性不同,成本型 QoS 属性值越小代表 Web 服务质量水平越好,例如响应时间、费用等;效益型 QoS 属性则正好相反,例如吞吐量、可用性等 QoS 属性参数。去模糊化的过程中要对这两种类型的属性分别处理。

假设有 m 个具有相同功能的 Web 服务集合为 $S = \{S_1, S_2, \dots, S_m\}$ 。QoS 属性参数集合为 $Q = \{Q_1, Q_2, \dots, Q_n\}$ 。其中 $q_{ij} = [q_{ij}^L, q_{ij}^R]$ 为 Web 集合中第 i 个 Web 服务的第 j 个 QoS 属性的区间值, $i = 1, 2, \dots, m; j = 1, 2, \dots, n$ 。转化后的相应实数值设为 q_{ij}' 。

对于效益型 QoS 区间属性数据转化为实数。转化公式如下

$$q_{ij}' = q_{ij}^L / \max_{1 \leq i \leq m} \{q_{ij}^L\} + q_{ij}^R / \max_{1 \leq i \leq m} \{q_{ij}^R\} + (q_{ij}^R - q_{ij}^L) / \max_{1 \leq i \leq m} \{q_{ij}^R - q_{ij}^L\} \quad (3)$$

对于成本型 QoS 区间属性数据转化为实数。转化公式

$$q_{ij}' = \min_{1 \leq i \leq m} \{q_{ij}^L\} / q_{ij}^L + \min_{1 \leq i \leq m} \{q_{ij}^R\} / q_{ij}^R + \min_{1 \leq i \leq m} \{q_{ij}^R - q_{ij}^L\} / (q_{ij}^R - q_{ij}^L) \quad (4)$$

式中, q_{ij}' 表示经过一致性转化后的 QoS 数值; $\max_{1 \leq i \leq m} \{q_{ij}^L\}$ 表示一组候选 Web 服务集合中某一区间型 QoS 属性中区间下限最大值; $\max_{1 \leq i \leq m} \{q_{ij}^R\}$ 表示一组候选 Web 服务集合中某一区间型 QoS 属性中区间上限最大值; $\max_{1 \leq i \leq m} \{q_{ij}^R - q_{ij}^L\}$ 表示一组候选 Web 服务集合中某一区间型 QoS 属性中区间上下限的距离最大值; $\min_{1 \leq i \leq m} \{q_{ij}^L\}$ 表示一组候选 Web 服务集合中某一区间型 QoS 属性中区间下限最小值; $\min_{1 \leq i \leq m} \{q_{ij}^R\}$ 表示一组候选 Web 服务集合中某一区间型 QoS 属性中区间上限最小值; $\min_{1 \leq i \leq m} \{q_{ij}^R - q_{ij}^L\}$ 表示一组候选 Web 服务集合中某一区间型 QoS 属性中区间上下限的距离最小值。

2.2.2 QoS 数据预处理

最终是通过多个属性综合值来评判服务的质量,采用多属性决策理论中的比重变换法对 QoS 数据进行归一化处理

$$q_{ij}' = q_{ij} / \sqrt{\sum_{i=1}^m (q_{ij})^2} \quad j \in Q, i = 1, 2, \dots, m \quad (5)$$

式中, q_{ij}' 表示经过归一化处理后的 QoS 数据;对于效益型数据 q_{ij} 表示 Web 服务集合中第 i 个 Web 服务中第 j 个 QoS 属性参数的取值;对于成本型数据, q_{ij} 表示 Web 服务集合中第 i 个 Web 服务中第 j 个 QoS 属性参数取值的倒数。 Q 表示效益型和成本型数据集合。

2.2.3 QoS 数值匹配

在进行完上述两步处理后,再计算精确型 QoS 数值相似度比较简单,公式如下

$$DSem(m_r, m_p) = \begin{cases} 1 - (r_j - q_{ij}) / r_j & Q_r > Q_p \\ 1 & Q_r = Q_p \\ 1 - (q_{ij} - r_j) / r_j & Q_r < Q_p \end{cases} \quad (6)$$

式中, $i \in [1, m]$, $j \in [1, n]$; $DSem(m_r, m_p)$ 表示 QoS 数值相似度函数; m_r 表示服务请求者请求的 QoS 数值约束; m_p 表示服务提供者提供的 QoS 数值约束; r_j 表示服务请求者所请求的第 j 个 QoS 属性参数的取值; q_{ij} 表示第 i 个候选 Web 服务中的第 j 个 QoS 属性参数的取值。

3 多属性 QoS 匹配

综合评价 Web 服务时考虑多个 QoS 属性综合匹配度,如费用、响应时间、吞吐量和可用性等。考虑到

Web 服务 QoS 属性的多样化,构建一个 $m \times n$ 的多属性决策矩阵 $M = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & & \vdots \\ q_{m1} & q_{m2} & \cdots & q_{mn} \end{bmatrix}$ 。

矩阵 M 说明如下:

(1) 行表示候选 Web 服务,列表示每个候选 Web 服务的一种 QoS 属性参数。

(2) $M_{m \times n}$ 中的元素 q_{ij} 表示服务请求者请求的 QoS 属性参数与第 i 个候选 Web 服务中的第 j 个 QoS 属性参数相匹配的综合相似度。

(3) 元素 q_{ij} 作为 QoS 的综合相似度是由概念语义综合相似度和数值相似度共同决定的。QoS 综合相似度函数 $QMatch(m_r, m_p)$ 为 QoS 语义综合相似度 $SSem(m_r, m_p)$ 与 QoS 数值相似度 $DSem(m_r, m_p)$ 的聚合乘积。当 QoS 语义可比时,数值匹配才有意义。当 QoS 语义相似度 $SSem(m_r, m_p)$ 不变, QoS 数值相似度越大越匹配;当 QoS 数值相似度 $DSem(m_r, m_p)$ 不变,语义相似度越大越匹配。因此 QoS 综合相似度函数如下所示

$$QMatch(m_r, m_p) = SSem(m_r, m_p) \times DSem(m_r, m_p) \quad (7)$$

式中, m_r 表示服务请求者定义的 QoS 属性参数; m_p 表示具有相同功能的候选 Web 服务集合中提供的 QoS 属性参数。

考虑服务请求者的个性化需求,采用权重分配法来描述请求者的愿望。将所构建的多属性决策矩阵与服务请求者所设定的权重分配数值进行相乘然后再将计算出的单个 QoS 属性评价结果相加,得到 Web 服务综合评价价值。如评价公式(8)所示

$$Sorce(S_i) = \sum_{j=1}^n w_j q_{ij} \quad (8)$$

式中, $Sorce(S_i)$ 为综合评价函数, $i = 1, 2, \dots, m$; q_{ij} 表示服务请求者请求的 QoS 属性参数与第 i 个候选 Web 服务中第 j 个 QoS 属性参数相匹配的综合相似度; w_j 表示服务请求者对所请求的 QoS 属性参数中第 j 个 QoS 属性的关注程度,且满足 $w_1 + w_2 + \dots + w_n = 1$ 。第 j 个 QoS 属性的权重 w_j 其取值范围在 $[0, 1]$ 区间内,如果值越趋近于 1,则表示服务请求者的个性化需求对此项的要求越高。

4 实验仿真与分析

采用 Protégé3.4.8 构建简单的 QoS 属性本体树;使用 Jena2.6.2 对本体树中的层次进行验证,如 WaitingTime 和 ExecutionTime 为 ResponseTime 的两个并列子节点。假设用户请求的 Web 服务 r 对于 QoS 属性参数要求为:费用(Price)为 10 元、兼容性(Compatibility)为 0.12 至 0.20 之间、吞吐率(Throughput)为 11.7 和响应时间(ResponseTime)为 0.8 s。具体候选 QoS 属性参数如表 1 所示。

基于 QoS 综合匹配的语义 Web 服务选择方法过程中,两个 QoS 属性参数之间相关性的临界距离 L , 本文将其设定为 1。QoS 语义匹配成功后,对相应的 QoS 数值进行匹配。

若服务请求者对于所请求的 Web 服务需求的权重分配为:费用(Price)占 0.4、兼容性(Compatibility)占 0.1、吞吐率(Throughput)占 0.2、和响应时间(ResponseTime)占 0.3。根据用户分配权重得出的 Web 服务评价结果: $S_1 = 0.759\ 93$, $S_2 = 0.705\ 45$, $S_3 = 0.987\ 90$, $S_4 = 0.577\ 75$, $S_5 = 0.449\ 44$, $S_6 = 0.646\ 85$, $S_7 = 0.757\ 47$, $S_8 = 0.454\ 05$, $S_9 = 0.618\ 33$, $S_{10} = 0.869\ 55$ 。可见 S_3 为最符合该用户需求的候

选服务 S_{10} 较符合用户要求。

表1 候选 Web 服务的 QoS 属性参数值

候选 Web 服务	Price/元	Cost/元	Compatibility	Throughput	ResponseTime/s	
					WaitingTime	ExecutionTime
S_1	11	15	[0.18 0.23]	10.52	0.75	0.68
S_2			[0.10 0.12]	18.18	0.89	
S_3		10	0.18	11.80	0.82	
S_4	18	19	[0.09 0.21]	21.88	0.91	0.78
S_5			[0.12 0.18]	21.12	[0.56 0.86]	
S_6		15	≤ 0.09	20.97	0.75	
S_7	20	13	[0.10 0.16]	6.82	0.68	0.58
S_8			[0.16 0.18]	7.87	[0.78 0.92]	
S_9		9	[0.13 0.20]	9.8		0.66
S_{10}		8	≤ 0.16	11.58	0.87	

若用户权重分配依次为:0.1、0.2、0.3、0.4,则得出的评价结果: $S_1 = 0.813\ 28$, $S_2 = 0.588\ 25$, $S_3 = 0.981\ 62$, $S_4 = 0.601\ 53$, $S_5 = 0.421\ 00$, $S_6 = 0.618\ 70$, $S_7 = 0.748\ 18$, $S_8 = 0.442\ 31$, $S_9 = 0.530\ 62$, $S_{10} = 0.933\ 18$ 。可见 S_3 为最符合该用户需求的候选服务, S_{10} 较符合用户要求。

从以上两次用户权重的分配来看, S_3 均很好的满足了用户的需求,这表明 S_3 在这四个属性中的匹配度均高于其他的服务。现在人为来考虑单个属性值在 QoS 数值上的相似程度。费用数值相似程度 $S_3 > S_2 > S_9 > S_{10}$, 由于 $S_3 = 10$ 恰好跟用户需求相同所以匹配度为 1; 兼容性在数值上没有扩大范围的候选服务有 S_3 、 S_5 、 S_8 、 S_9 , S_3 恰为单值数据可以看作恰好满足了区间型数据的核; 吞吐量数值相似程度 $S_3 > S_{10} >$ 其他服务; 响应时间需求为 0.8 s 可以看到 $S_3 = 0.82$ s, 完全可以匹配用户需求而且相较于其他服务而言为精确型单值数据。请求者期望所选择的服务中 QoS 属性的侧重度为费用大于其他三种属性所占比例时, 结果明显 S_3 从各个方面来说是最满足服务请求者需求的服务。

根据算法仿真实验, 本文提出的方法可以有效的给用户推荐满足且服务质量较高的 Web 服务。

5 结束语

本文提出了一种多属性 QoS 综合匹配方法用于解决 Web 服务选择问题。在语义匹配后给出 QoS 数值相似度计算方法, 采用去模糊化方法将异构 QoS 数值参数统一转化为单值型 QoS 数值, 解决异构 QoS 参数值之间的匹配问题。构建多属性决策矩阵解决 QoS 属性多样化, 并考虑了服务请求者个性化需求。本方法侧重于区间型 QoS 数值, 对于参数值得确定人为因素影响也较大, 还需要进一步的考虑。

参 考 文 献

- [1] 王安华, 国林, 晓娟, 等. 基于服务质量的 Web 服务发现研究与实现[J]. 计算机工程与设计, 2007(21): 5112-5114.
- [2] CHEN ZHOU, LIANG TIEN CHIA, BU SUNG LEE. DAML-QoS ontology for web services[C]//Proceedings of the International Conference on Web Services(ICWS04). San Diego, California: IEEE Computer Society, 2004: 472-479.
- [3] Taher L, Khatib H El. A framework and QoS matchmaking algorithm for dynamic web services selection[C]//Proceedings of the 2nd International Conference on Innovations in Information Technology (IT05). [S. l.]: [s. n.], 2005.
- [4] MICHAEL MAXIMILIEN E, MUNINDAR P S. Toward autonomic web services trust and selection[C]//Proceedings of the 2nd International Conference on Service Oriented Computing. New York: ACM Press, 2004: 212-221.
- [5] GANJISAFFAR Y, ABOLHASSANI H, NESHATI M. A similarity measure for OWL-S annotated web services[C]//Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence. Hong Kong: IEEE Computer Society, 2006: 621-624.
- [6] Ganesan P, Garcia-Molina H, Widom J. Exploiting hierarchical domain structure to compute similarity[J]. Transaction on Information Systems, 2003, 21(1): 64-93.

(下转第 108 页)

- [4]朱川曲. 采煤工作面瓦斯涌出量预测的神经网络模型[J]. 中国安全科学学报, 1999, 9(2): 42-45.
- [5]史峰, 王小川, 郝磊, 等. MATLAB 神经网络 30 个案例分析[M]. 北京: 中国航空航天大学出版社, 2010.
- [6]韩玉芳, 谢士宏, 杨龙. 马家坡瓦斯隧道设计等级的判定研究[J]. 石家庄铁道大学学报: 自然科学版, 2011, 24(4): 74-77.

Gas Emission Forecast Based on GRNN

Ge Jiang

(School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: This paper introduces the characteristics of the gas emission forecast and the GRNN, and points out the rationality of using GRNN to forecast the gas emission and its advantage over the BP. The forecast is carried out based on a case, proving that GRNN can solve the problem of gas emission.

Key words: neural; gas emission; forecast; GRNN

(责任编辑 车轩玉)

(上接第 101 页)

A Web Services Selection Method Based on QoS Comprehensive Matching

Zhang Cuixiao, Xu Li, Chong Li

(Major Computer Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang 050043, China)

Abstract: Currently, many selection methods of Web Services based on QoS don't have a comprehensive consideration of the QoS semantics, numerical processing and service requestors' personality needs on service quality. The method presented in this paper is to apply multi-attribute decision-making theory to the QoS value pretreatment and QoS comprehensive matching and to design comprehensive matching method of QoS parameter. The efficiency and feasibility of the method is verified.

Key words: semantic Web service; semantic similarity; QoS value similarity

(责任编辑 车轩玉)

(上接第 104 页)

Campus WLAN Deployment and Access Authentication

Yan Shiqiang, Xiao Zhong

(Educational Technology of Center Shijiazhuang Tiedao University, Shijiazhuang 050043, China)

Abstract: In view of campus WLAN deployment and access authentication method for the construction of the campus WLAN, the advantages and disadvantages of several common ways of WLAN deployment are analyzed, pointing out the campus WLAN should be constructed by a local forwarding mode. Campus WLAN access method using 802.1x is given, as well as how to achieve a unified user authentication with an existing wired network. WLAN construction of the campus network through this program will be able to safeguard the campus WLAN efficiency, convenience and security, and achieve unified management of wired and WLAN.

Key words: campus network; WLAN; 802.1x; authentication method

(责任编辑 车轩玉)