

面向服务组合的复杂上下文本体模型^①

张龙昌, 万 君

(辽宁工程技术大学 营销管理学院, 葫芦岛 125105)

摘 要: 为了有效解决泛在网络环境下 Web 服务质量上下文描述问题, 在 OWL_S 基础上提出了面向服务组合的复杂上下文本体模型-OWL_S_QoS+, 包括用户上下文本体和服务上下文本体。用户上下文本体描述用户的环境状态、拥有的资源信息、用户需求及用户群组等对服务或组合服务运行产生影响的一系列属性; 服务上下文本体在原有 OWL_S 定义的基础上, 增加了服务对客户端与服务端的运行环境需求、不确定性 QoS 描述方法。通过模型比较验证了 OWL_S_QoS+ 具有较强的描述能力。

关键词: 复杂上下文本体; Web 服务; 服务质量; 服务组合; OWL_S

Web Service Composition-Oriented Complex Context Ontology Model

ZHANG Long-Chang, WAN Jun

(School of Marketing Management, Liaoning Technical University, Huludao 125105, China)

Abstract: To solve the context describe of quality of Web service in ubiquitous network environment, service composition-oriented complex context ontology model-OWL_S_QoS+ is proposed, which includes user context ontology and service context ontology. User context ontology describe user's environment, user's resource, user's requirements and user's group and so on, these information affects the running of Web service. Web service context ontology is defined Based on OWL_S, which added the running environment requirements of Web service in client and server and the QoS description method with uncertainty.

Key words: complex context ontology; Web service; quality of service; service composition; OWL_S

语义 Web 服务已经有较多研究, 本文在已有工作上基于下述问题提出新的 QoS 本体模型: 1) 互联网的开放性、动态性, 导致服务 QoS 具有很强不确定性, 服务本身可能暂时或者永久地关闭^[1]。QoS 数据来源于用户反馈、专家评估、自动监控, 某些 QoS 属性以不确定的形式给出能增强模型描述能力^[2]和反应 QoS 综合情况。对于可以精确描述并且变化频率较小的属性使用确定数据类型描述, 如价格。有些属性在一段时间内随机波动且有突发最大值和最小值用区间数表示, 如响应时间。有些 QoS 属性需要从用户的反馈中获得, 用户不是 QoS 专家不能给出明确的估计值, 用语言型数据(三角模糊数^[2]和直觉模糊集^[3])表示较合理。上述描述方法各有优缺点且分别适用不同的

QoS 属性, 而目前的 QoS 模型都不能支持上述四种数据类型混合。2) 泛在网络环境和交互信息动态变化, 为了提供适时适地的服务, 服务组合除了满足用户的需求之外, 还要适应环境的变化。用户环境主要指用户所拥有的设备资源、用户当前所处网络环境(也包括自然环境和社会环境)。因此需要建立一个能够描述用户上下文信息的用户上下文本体模型。

另外, 组合服务能够完成复杂功能, 应用广泛, 其 QoS 需要根据原子服务 QoS 计算。因此, 基于新的 QoS 本体模型, 需要给出组合服务 QoS 聚合算法。

为解决上述问题, 在 OWL-S^[4]基础上提出复杂上下文本体-OWL_S_QoS+, 并给出支持实数、区间数、三角模糊数和直觉模糊集的组合服务 QoS 聚合算法。

① 基金项目: 国家基础理论研究计划(973)(2009CB320406, 2009CB320504)

收稿时间: 2011-07-30; 收到修改稿时间: 2011-09-01

1 相关工作

1.1 OWL_S

OWL_S^[4]是基于 OWL 语言的描述 Web 服务的本体, W3C 推荐的标准, 从语义和逻辑角度来描述 Web 服务属性和功能。Resource 提供给使用者调用 Web 服务的基本信息、ServiceProfile 提供 Web 服务的相关信息供查找使用、ServiceModel 描述服务执行的具体过程、ServiceGrounding 描述服务是如何被访问。presents、describedby 和 supports 分别解决 Web 服务提供什么样的服务, Web 服务具体是怎样工作的以及如何访问这个 web 服务。目前语义 Web 服务的研究主要围绕 OWL-S 展开, OWL-S 提供了 Web 资源的服务化和语义化支持, 可以成为泛在、普适网络环境下资源描述的一种有效方法, 图 1 是 OWL-S 顶层本体表示。

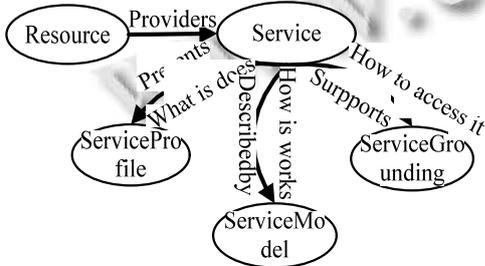


图 1 OWL_S 结构图

1.2 OWL_S+

OWL_S 在表达一般规则方面有局限性, 不能很好地表示用户的偏好和约束^[5]。文献[5]利用语义 Web 规则标记语言 OWLRule+对 OWL_S 的模型和本体进行一定的扩展和完善, 使其包含用户偏好和约束的语义描述 (OWL_S+), 图 2 是顶层本体表示。

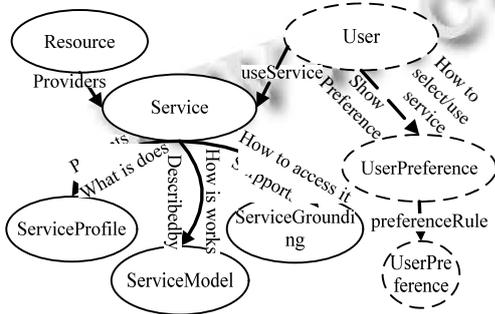


图 2 OWL_S+ 结构图

1.3 OWL_S_QoS

OWL_S+能够较好的描述用户偏好和偏好规则, 但是不能描述不确定 QoS。文献[2]提出了能够支持实

数、区间数、三角模糊数的混合 QoS 模型, 图 3 是 OWL_S_QoS 模型的顶层本体表示。

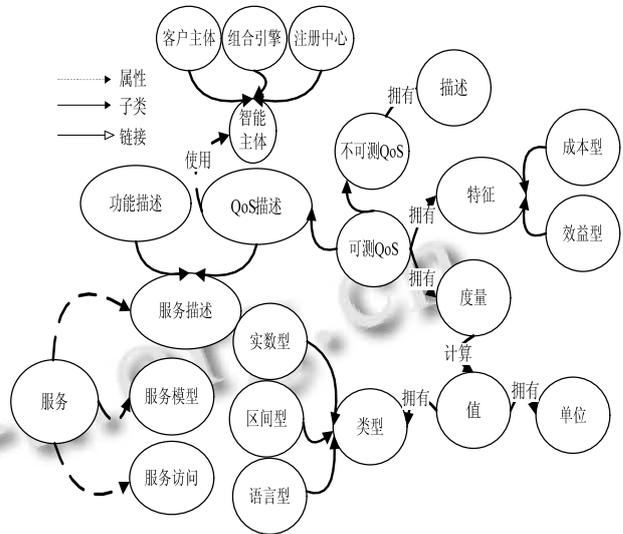


图 3 OWL_S_QoS 结构图

1.4 本文工作与相关工作比较

从 OWL_S、OWL_S+、OWL_S_QoS+中可以看出, 其面向的角色主要有服务提供者、服务调用者、终端用户。服务提供者提供 Web 服务并且使用 OWL_S 或者 OWL_S 的扩展版本发布其 Web 服务相关信息; 服务调用者通常依据 OWL_S 或者 OWL_S 的扩展版本的语义描述调用 Web 服务, 也可以生成新的服务并且使用 OWL_S 或者 OWL_S 的扩展版本描述; 终端用户直接使用 Web 服务, 描述其需求。泛在网络环境的自动化服务选择和组合需要 Web 服务本体能够描述 Web 服务和用户上下文信息。本文侧重点是复杂环境下的 Web 服务非功能属性上下文描述和用户非功能需求的上下文描述。

OWL_S 在描述用户偏好和规则上具有很大缺陷, OWL_S+弥补了这缺陷, 但是仅仅考虑具有一定规则的用户偏好信息; 对于用户的不确定的 QoS 需求和当前用户所处的环境、拥有的资源丝毫没有考虑, 这种模型还很难适应泛在网络环境下的服务自动化选择和组合。OWL_S_QoS 提供了一套能够描述不确定信息表达的服务 QoS, 但是同样没有考虑用户的环境、拥有的资源信息以及不确定信息的 QoS 需求。因此, 本文对 OWL_S+和 OWL_S_QoS 本体进行一定的扩展和完善, 使其包含用户环境信息、拥有的资源信息和不确定语言描述的用户偏好以及用不确定语言描述的

Web 服务 QoS 信息和 Web 服务运行环境需求信息。

就我们的认知而言,以往 QoS 模型^[2,3]都不能支持实数、区间数、三角模糊数和直觉模糊集的混合,并且这些模型都没有考虑用户上下文对组合服务选择的影响。在文献[2]工作基础上,本文引入了用户上下文和 Web 服务上下文。在用户上下文中详细描述了用户偏好、资源上下文和用户群组。在 Web 服务上下文中引入了更先进、处理模糊能力更强的直觉模糊集描述那些用户犹豫度比较强的反馈数据(如信誉);三角模糊数描述用户反馈较确定的数据(如可靠性);实数描述能够精确表达的 QoS 属性(如价格);区间数描述一段时间内突发最大值、最小值的 QoS 属性(如可用性、响应时间)。另外,文献[2]给出的组合服务 QoS 聚合算法仅仅能支持实数、区间数、三角模糊数,本文在组合服务 QoS 聚合模型中给出了能够支持实数、区间数、三角模糊数和直觉模糊集的聚合算法。

2 OWL_S_QoS+

OWL_S_QoS+增加了用户类、群组类、用户需求类、用户描述类(不是本文重点不做详细说明)以及重新定义的服务非功能属性类。本文将描述用户信息的类称用户上下文本体,描述服务信息的类称服务上下文本体,具体定义如下。

定义 1: 用户上下文本体描述用户所处的环境状态、拥有的资源信息、用户需求及用户的社会背景等对运行的服务或组合服务产生影响的一系列属性。

定义 2: 服务上下文本体指除原有 OWL_S 定义的概念外,服务非功能属性应当包括服务客户端与服务端的运行环境、需求的资源等与服务相关的一系列属性以及用各种方式表达的服务 QoS 信息。

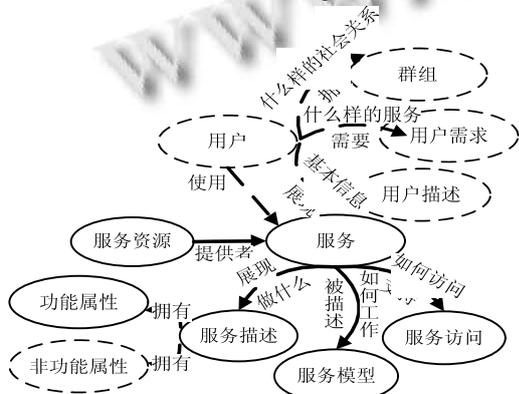


图 4 OWL_S_QoS+结构图

2.1 用户上下文本体

1. 顶层本体

用户上下文本体基于 OWL_S (见图 5), 图中虚线框为用户顶层本体。



图 5 用户上下文本体

用户需求包括功能需求和非功能需求;其中非功能需求是本文的研究重点,功能需求不做讨论。用户偏好和资源上下文为非功能需求的子概念。用户偏好描述了用户的 QoS 权重,是一个三元组(关键字、权重、阈值),其中关键字描述 QoS 属性名字,权重描述用户对 QoS 属性的期望程度(值或表达式),值是 QoS 偏好的具体数值,表达式是用户不能精确描述 QoS 的偏好表达(如价格大于响应时间等),阈值描述用户对某 QoS 的最低需求(如价格不能大于 100),阈值由单位和类型组成,阈值的单位链接到现有度量单位,阈值类型定义为实数型、区间型、语言型。资源上下文描述用户当前环境信息(网络环境、自然环境、社会环境等)和设备资源,其信息由传感器获得。泛在网络环境下,用户的资源上下文必须能够满足所运行服务的需求。用户资源上下文是一个四元组(类别、关键字、值、状态),类别表明资源所属的分类,关键

字描述资源名称, 值描述资源的度量值, 状态描述当前资源所处的状态 (状态根据实际情况确定可能包括正常、失效等语言类描述, 也可能是区间型、实数类型描述), 值由类型和单位组成, 类型包括实数型、区间型、语言型, 单位链接到现有度量单位。

群组是一个五元组 (关键字、所属类别、描述、成员、群组需求), 关键字描述群组的名称。所属类别描述群组的种类 (如运动、娱乐、休闲等等)。描述说明了群组的相关信息。群组需求描述群组的公共需求, 包括功能需求 (群组需要的功能, 文中假设备选服务已经满足功能需求) 和非功能需求 (群组内所有成员的公共的非功能需求); 非功能需求包括描述群组的共同资源的资源上下文和群组偏好, 群组偏好与用户偏好定义相同。成员描述组内其他用户, 链接到用户上下文本体 (尤其指的是用户的 QoS 偏好部分本体)。

2. 中层本体

本节定义用户偏好和资源上下文的中层本体, 群组定义简单这里不做说明。本文考虑具有代表性的 6 个用户环境、资源信息与资源上下文中的资源上下相对应, CPU、内存、网络带宽、信号强度、负载、网络状态。另外某些环境信息如温度、湿度可能会影响服务运行效果, 这些环境信息需要在一定的取值范围内, 因此用区间数表示。信号强度通常在手持设备上显示为 (弱, 较弱, 一般, 较强, 强), 因此用三角模糊数表示; 网络状态通常由用户描述, 用直觉模糊集表示; 其它四种属性能够精确获取, 用实数表示。

用户偏好中层本体定义与服务上下文中层本体中 QoS 属性相对应, 包括服务的价格、可用性、响应时间、可靠性和信誉 5 个属性的权重; 可能会用实数、区间数、三角模糊数和直觉模糊数中的一种描述。

2.2 服务上下文本体

1. 顶层本体

服务上下文本体扩展 OWL_S, 本文只考虑服务非功能属性。图 6 中虚线框内概念为 OWL-S 定义的顶层概念。资源上下文和 QoS 都是非功能属性的子概念。资源上下文是一个三元组 (关系、关键字、值), 用于描述服务正常运行所需的用户端资源信息, 如 CPU、内存、带宽、信号强度、负载、网络状态等; 其中关系指的是服务要求用户某个资源上下文必须符合数值关系 (等于、大于、大于等于、小于、小于等于)、集合关系 (包含于、真包含于、等价、包含、真包含)

中的一种, 数值中的等于关系和集合关系中的等价关系定义为等价关系, 其它关系定义为偏序关系; 关键字是资源上下文属性的唯一标识; 资源上下文属性的具体度量用值描述, 用度量单位和类型表示, 值的度量单位链接到现有度量单位, 类型定义为实数型、区间型和语言型。QoS 包括关键字、特征、值; 关键字指明 QoS 属性的名称; 特征表明属性的类型 (通常有成本型和效益型; 成本型属性描述消耗用户资源的指标, 因此成本型属性越小越好; 效益型属性描述给用户带来效益的指标, 因此效益型属性越大越好); 属性值描述某个属性的度量 (通常用实数表示能够精确度量的属性, 用区间数描述那些在某段时间有一定变换范围的属性, 用语言型数据描述不能精确度量但可以通过人类模糊评估的属性)。

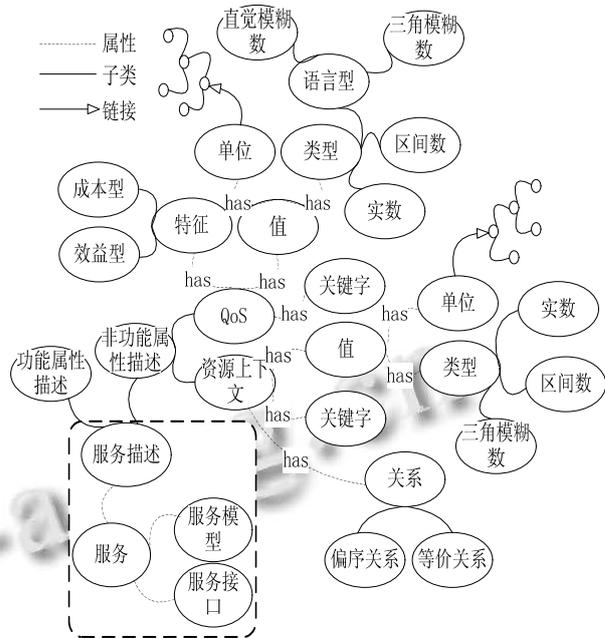


图 6 服务上下文顶层本体

2. 中层本体

在服务上下文中层本体中主要定义了一系列的 QoS 属性, 用来对服务质量的不同方面进行描述; 服务的资源上下文属性定义了服务运行所需要的资源设备、环境信息以及体现服务当前状态的其它相关信息, 包含于用户资源上下文的属性定义 (这里不做介绍, 具体定义见用户上下文中层本体)。本文考虑具有代表性的 5 个 QoS 属性, 服务的价格(price)、可用性(availability)、响应时间(response time)、可靠性

(reliability)、信誉(reputation)等指标。QoS 模型可以根据实际需要进行扩展, 当有新指标加进来时, 不需要对建立在该模型上的服务选择方法进行修改。下面给出组件 Web 服务的各指标的定义及其度量方法。

价格: 表示调用服务所需要支付的费用, 计算公式为: $P(ws) = p_{communication} + p_{provider} + p_{goods}$, 其中 $p_{communication}$ 调用服务的通信费, $p_{provider}$ 服务费用, 用户购买的虚拟商品或实际商品的费用, p_{goods} 该属性费用固定用实数(R)表示。

可用性: 表示 Web 服务可调用的概率, 从用户的反馈中获得, 是一段时间内用户成功访问次数与总访问次数的比值。各种复杂性和不确定性导致可用性会在一定范围内波动, 用区间数(IR)表示服务可用性:

$A(ws) = [\min(a^l, a), \max(a^u, a)]$, 其中 $a = Tr/n$, Tr 为某次监控成功访问次数, n 为总次数, a^l 和 a^u 为可用性当前最低概率和最高概率。

响应时间: 表示从客户请求消息发出到最终服务执行结果展现给用户所需时间开销, 单次模型为:

$t(ws) = t_{request} + t_{transmission} + t_{excute} + t_{display}$, $t_{request}$ 表示用户提交请求信息后软件经过一系列处理转换成标准的服务请求格式所需要时间, $t_{transmission}$ 指请求信息和服务执行结果返回信息在网络中的传输时间, t_{excute} 表示服务运行时间, $t_{display}$ 表示执行结果到达客户端到结果显示给用户需要的时间。响应时间在一定范围波动, 用区间数表示如下:

$$T(ws) = \left[\begin{matrix} \min(t(ws)_1, t(ws)_2, \dots, t(ws)_n) \\ \max(t(ws)_1, t(ws)_2, \dots, t(ws)_n) \end{matrix} \right]$$

可靠性: 表示 Web 服务正确执行其功能的能力。它是对服务质量的一个综合评判, 其计算数据来自于用户反馈。在以往研究中可靠性大多被定义为精确实数, 也有模糊集^[2], 本文定义为三角模糊数。对应的语言集合(高(very high), 较高(high), 一般(normal), 较低(low), 低(very low))及其对应的三角模糊数集合为([8,10,10], [5,7,9], [3,5,6], [1,3,4], [0,0,2])。

然后, 根据用户的反馈定义可靠性模型为:

$R(ws) = \left[\frac{1}{n} \sum_{i=1}^n a_i^l, \frac{1}{n} \sum_{i=1}^n a_i^m, \frac{1}{n} \sum_{i=1}^n a_i^u \right]$, 用户一次反馈的三角模糊数为 $[a_i^l, a_i^m, a_i^u]$ (a_i^l, a_i^m, a_i^u 分别表示下界、核、上界), n 表示一段时间用户反馈次数, 服务

可靠性定义为一段时间用户反馈的平均值。

信誉: Web 服务的信誉反映服务可信程度, 来源于用户反馈。在以往研究中信誉大多被定义为精确实数, 也有定义为直觉模糊集(IFS)^[3]。本文定义为直觉模糊数(IFSs), 对应的直觉模糊集(高(very high), 较高(high), 一般(normal), 较低(low), 低(very low))及其与直觉模糊数的对应关系如表 1, 其中 π 表示用户评分的犹豫度。根据用户的反馈定义信誉模型为:

$C(ws) = \frac{1}{n} \sum_{i=1}^n A_i$, n 表示一段时间服务访问的总次数 (表示某次用户的反馈值)。价格和响应时间是成本型数据, 其它 3 种类型是效益型数据。

表 1 信誉的语言短语与直觉模糊数对应关系

Linguistic terms	Very high(VH)	High(H)
IFNs	$\langle 0.9, 0.1-\pi \rangle$	$\langle 0.7, 0.3-\pi \rangle$
Normal (N)	Low(L)	Very low(VL)
	$\langle 0.5, 0.5-\pi \rangle$	$\langle 0.3, 0.7-\pi \rangle$
		$\langle 0.1, 0.9-\pi \rangle$

3 组合服务QoS聚合模型

组合服务由原子服务按照一定逻辑结构进行组织的能够完成复杂功能的逻辑实体, 经过封装在互联网上发布后便成为 Web 服务。Web 服务组合方法从组合方案生成方法可以分为基于工作流的方法和基于语义推理的方法。在工作流中用来建立组合的基本结构包括串行结构(sequence)、选择结构(switch)、并行结构(parallel)和循环结构(loop) (如图 7 所示)。

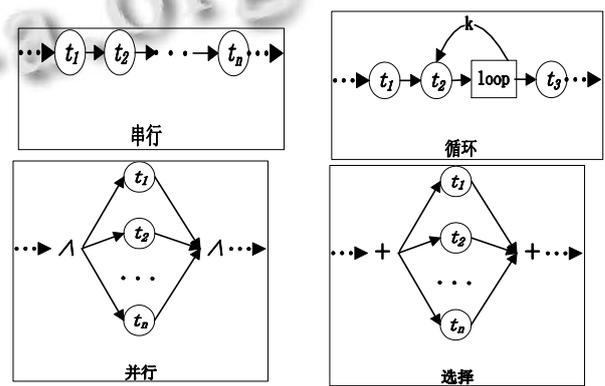


图 7 工作流的基本结构

为简化复杂组合服务的 QoS 计算, 文献[6]将上述基本结构简化为串行结构和并行结构。本文基于串行和并行结构对组合服务 QoS 计算方法进行研究。根据实数、区间数^[7]、三角模糊数^[8]、直觉模糊数^[9]四则运

算得组合服务 QoS 聚合模型, 见表 2。

表 2 组合服务 QoS 聚合模型

QoS	串行	并行
Price	$P(ws) = \sum_{i=1}^n P_i$	$P(ws) = \sum_{i=1}^m P_i$
Availability	$A(ws) = \left[\prod_{i=1}^n a_i^t, \prod_{i=1}^n a_i^u \right]$	$A(ws) = \left[\prod_{i=1}^m a_i^t, \prod_{i=1}^m a_i^u \right]$
Response	$T(ws) = \left[\sum_{i=1}^n t_i^t, \sum_{i=1}^n t_i^u \right]$	$T(ws) = \left[\max(t_i^t), \max(t_i^u) \right]$
Reliability	$R(ws) = \left[\frac{1}{n} \sum_{i=1}^n r_i^t, \frac{1}{n} \sum_{i=1}^n r_i^m, \frac{1}{n} \sum_{i=1}^n r_i^u \right]$	$R(ws) = \left[\frac{1}{m} \sum_{i=1}^m r_i^t, \frac{1}{m} \sum_{i=1}^m r_i^m, \frac{1}{m} \sum_{i=1}^m r_i^u \right]$
Reputation	$C(ws) = \frac{1}{n} \sum_{i=1}^n C_i$	$C(ws) = \frac{1}{m} \sum_{i=1}^m C_i$

泛在网络环境下的服务组合: 服务组合场景为一个简单的移动导航服务, 该服务包括移动定位(Task1)、线路显示(Task2)、路况显示(Task3), 这 3 项任务串行执行, 三个任务的 QoS 值和组合服务 QoS 如表 3。

表 3 服务 QoS

任务	价格	可用性	响应时间	可靠性	信誉
Task1	30	[0.93, 0.97]	[75, 130]	[3, 5, 6]	<0.1, 0.7>
Task2	35	[0.96, 0.98]	[85, 135]	[1, 3, 4]	<0.9, 0.05>
Task3	50	[0.92, 0.95]	[75, 105]	[8, 10, 10]	<0.3, 0.5>
组合服务 QoS	115	[0.82, 0.90]	[235, 370]	[4, 6, 7]	<0.937, 0.0175>

泛在网络环境下不仅仅需要考虑服务 QoS 信息还需要考虑用户的资源环境信息以及 Web 服务所能够适应的用户上下文条件, 如表 4。用户网络带宽 512K, Task3 无法在该用户上运行, 因此不能被选中。

表 4 服务需求上下文和用户上下文

	CPU	内存	网络带宽	信号强度	网络状态
Task1	512M 以上	256M 以上	512K 以上	“一般” 以上	<0.1, 0.7> 以上
Task2	256M 以上	512M 以上	256K 以上	“较弱” 以上	<0.2, 0.1> 以上
Task3	512M 以上	256M 以上	1M 以上	“较强” 以上	<0.3, 0.3> 以上
用户上下文	512M	512M	512K	“一般”	<0.5, 0.3>

3 结语

本文针对 Web 服务本体研究较少考虑用户资源上下文、群组、用户偏好等问题, 提出一种新的 Web 服务上下文本体—OWL_S_QoS+。该本体不仅能够描述用户的资源上下文、用户偏好、用户群组, 还能描述实数、区间数、三角模糊数、直觉模糊集表示的服务

QoS。同时本文还提出一个能够支持上述四种数据类型的组合服务 QoS 聚合算法。在该上下文本体基础上, 我们将提出一种基于多属性群决策理论的组合服务选择算法用于完成服务与用户上下文匹配、获得群体 Pareto 最优组合服务选择。

参考文献

- 1 Su S, Li F, Yang FC. Iterative selection algorithm for service composition in distributed environments. Science In China Series F—Information Sciences, 2008,51(11):1841–1856.
- 2 Yang FC, Su S, Li Z. Hybrid QoS-aware semantic web service composition strategies. Science In China Series F—Information Sciences, 2008,51(11):1822–1840.
- 3 Wang P. QoS-aware web services selection with intuitionistic fuzzy set under consumer’s vague perception. Expert Systems with Applications, 2009,36(3):4460–4466.
- 4 Martin D, Burstein M, Hobbs J, Lassila O, McDermott D, McIlraith S, Narayanan S, Paolucci M, Parsia B, Payne T, Sirin E, Srinivasan N, Sycara K. OWL-S: Semantic markup for Web services. <http://www.w3.org/Submission/2004/SUBM-OWL-S-20041122/>.
- 5 梁晟. 基于语义 Web 的服务自动组合技术的研究.[博士学位论文]. 北京: 中国科学院软件研究所, 2004.
- 6 Ardagna D, Pernici B. Adaptive service composition in flexible processes. IEEE Trans. on Software Engineering, 2007,33(6):369–384.
- 7 Kulisch UW. Numerical Validation in Current Hardware Architectures. International Dagstuhl Seminar, Dagstuhl Castle, Germany. 2009.
- 8 Giachetti RE, Young RE. A parametric representation of fuzzy numbers and their arithmetic operators. Fuzzy Sets and Systems, 1997,91(2):185–202.
- 9 Atanassov K. New operations defined over the intuitionistic fuzzy sets. Fuzzy Sets and Systems, 1994,61(2):137–142.