

一种基于 QoS 的扩展语义 Web 服务发现方法*

李小林¹, 张力娜², 李卫斌¹

(1. 咸阳师范学院 信息工程学院; 2. 咸阳师范学院 数学与计算科学学院, 陕西 咸阳 712000)

摘要 如何从大量的 Web 服务集合中找到满足用户要求的服务是 Web 服务技术的关键,然而现有 Web 服务缺乏对 QoS 信息的描述,不能满足用户对服务 QoS 的需求。本文基于 OWL-S 本体扩展了 profile 描述,将 QoS 本体引入 OWL-S,在基于 SOA 工作框架以及 P2P 网络的基础上,设计了 Web 服务发现框架,并在该框架中引入了 QoS 管理模块。服务匹配过程分为两个阶段:第一阶段为基于服务功能的匹配,服务的功能主要从 inputs、outputs、preconditions 和 effect 描述;第二阶段为基于服务 QoS 的匹配,QoS 属性从 time、cost 和 reliability 描述,以满足用户对服务质量提出的要求。实验结果表明该模型可以有效地增加 Web 服务发现的有效性。

关键词 Web 服务; QoS 本体; OWL-S; 服务发现模型

中图分类号 TP393.04

文献标识码 A

文章编号 1672-6693(2010)05-0055-04

Web 服务是架构在 XML 和 Internet 技术之上的分布式计算技术。借助于 SOAP、WSDL、UDDI 等基于 XML 的技术以及 Web 服务本身所具有的完好松散耦合性特点,各种基于网络的 Web 服务如雨后春笋般出现了。主流的 Web 服务发现方法采用基于 WSDL 描述和 UDDI 框架的服务匹配方法,缺乏对语义信息的支持,服务的查全率和查准率不高。现有的语义 Web 服务发现方法因缺乏对服务质量(QoS)的描述,而难以从众多功能相似的服务中选择满足用户需求的最佳服务。具有 QoS(Quality of services)保证的 Web 服务质量已成为服务提供商成功的关键和客户选择的依据以及 Web 服务研究的焦点之一。

QoS 指与服务相关的非功能属性,描述关于服务某些方面的性能。文献[1]将 QoS 属性分为性能、可用性、可访问性、完整性、完全性、可靠性、规范性和安全性。文献[2]提出一种支持 QoS 约束的 Web 服务发现模型,定义了一组描述 Web 服务 QoS 和信誉度的分类 tModel,采用了 QoS 协商和反馈机制,支持携带 QoS 描述信息的服务发布以及基于 QoS 约束的服务发现。文献[3]提出了一种新的服务发现模型,在服务查找过程中考虑了服务的 QoS 属性。该模型对 UDDI 进行了扩展,增加了一种新的数据结构,用于描述 Web 服务的 QoS 属性。

本文引入本体论,扩展了 OWL-S profile 本体,使得 Web 服务基于语义描述的基础上支持 QoS 属性的描述,在服务发现过程中对功能属性满足请求者需求的基础上进一步进行 QoS 性能比较,从而选取用户更满意的服务。

1 相关知识

1.1 本体

Ontology(本体)最初是一个哲学范畴,是指客观存在的一个系统的解释和说明。由 Gruber 提出的^[4]:Ontology is a formal, explicit specification of a shared conceptualization。即本体是共享概念模型的明确的、形式化规范说明。概念模型(Conceptualization)指通过识别世界上的一种现象的相关概念来对这种现象所建立的抽象模型。本体可以通过带标记的有向树来表示,节点描述概念,节点之间的连线对应于概念之间的某种语义关系。图 1 给出一个本体示例图。

1.2 QoS 本体建模

OWL-S 是构建于 OWL(Web ontology language)之上的用于描述 Web 服务的标记语言。OWL-S 主要由 ServiceProfile、ServiceModel 和 ServiceGrounding 构成。其中 ServiceProfile 提供了服务的基本描述,包括服务提供者的基本信息,如 ServiceName、text-

* 收稿日期 2010-04-19

资助项目 陕西省自然科学基金研究项目(No. 2007D22); 咸阳师范学院专项科研基金(No. 07XSYK214; No. 07XSYK107)

作者简介 李小林,男,讲师,硕士,研究方向为网络计算。

Description 等 ;服务的功能属性 ,如 Inputs ,Outputs , Preconditions ,Effects ,即所谓的 IOPE ;服务其他信息 ,如 ServiceCategory 等。

定义 1 服务质量模型 $QoS = \langle time, cost, reliability \rangle$ 。 $time$ 代表服务响应时间 ,指客户提交服务请求到获得服务响应所花时间 ,包括传输时间、处理时间。 $cost$ 代表服务的费用 ,指服务需求者如果要调用某一服务必须支付的费用 ,一般由服务提供者指定。 $reliability$ 代表服务的可靠性 ,指服务能够正

确响应的比率 $reliability = s/k$,这里 s 表示成功执行的次数 k 代表一共执行的次数。本文在参考文献 [5]的基础上 ,利用 OWL-S Profile 定义中 ServiceParameter 扩充机制引入 QoS 本体的定义 ,用以描述服务的质量。

根据服务质量模型的定义 ,引入的 QoS 本体应该具有 $time, cost, reliability$ 3 个属性 ,分别从响应时间、费用和可靠性方面度量服务的质量。图 2 给出了 QoS 本体的描述模式^[5]。

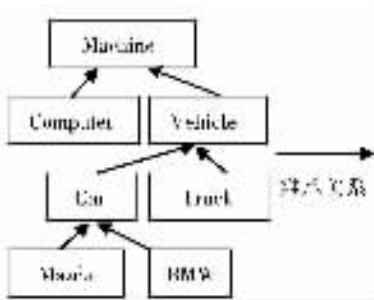


图 1 本体树

该模式分为 3 层。QoS Profile 层定义了 QoSProfile 类 ,具有两个数据属性 :Name 代表 QoS 本体描述的 Web 服务的名称 ,textDescription 表示对 QoS 本体简单的文本描述。一个对象属性 :hasParameter ,其取值范围为类 QoSParameter。QoSParameter 层主要是关于服务质量各个参数的描述层 ,具有两个数据属性 :paramName(参数的名称) ,paramDescription(参数的文本描述)。一个对象属性 hasMetric ,其取值范围为类 QoSMetric。QoSMetric 层主要给出各个参数的度量定义 ,包括度量的单位(unit) ,度量的值(value) ,度量的名称(MetricName)以及度量描述(MetricDescription)。

2 系统架构

本文设计的服务架构是在 SOA 的模型的基础上扩展的 ,基于 P2P 覆盖网络。如图 3 所示。

1) 服务提供者使用本文提出的带有 QoS 的扩展 OWL-S Profile 描述所提供的服务。所发布的 Web 服务信息可以发布在 P2P 网络中的任意节点。

2) 服务请求者可根据需求发布基于功能和 QoS 的服务请求 ,这个请求可通过客户端界面输入 ,形成 OWL-S 描述的服务请求。P2P 覆盖网络中的节点通过服务代理负责将服务请求路由给能提供这个服务的节点上。

3) 服务代理收到服务请求者的服务请求 ,负责

查找满足服务请求者功能及 QoS 需求的服务 ,并将结果返回给服务请求者。

4) 服务请求者和发布者绑定。

5) 提供服务注册的节点负责接收服务请求者反馈的关于它所获得服务的 QoS 性能评级 ,该节点把 QoS 报告传给 QoS 管理中心。QoS 管理中心负责为 QoS 评级进行管理 ,定时更新服务的 QoS 信息。

服务代理在服务查找过程中扮演了重要的角色 ,其内部结构如图 4 所示。

匹配模块由功能匹配 ,QoS 匹配组成 ,匹配模块是服务代理的核心部件。匹配模块在本地库 ,OWL-S 和 Advertisement 库和 QoS 数据库的支持下完成服务的匹配过程。

通信模块相当于一个信息总线 ,连接系统内其他内部模块 ,并与系统外部实体交互 ,例如服务请求用户、QoS 管理中心、服务提供者 ,用以获得服务请求信息、QoS 数据以及用户 QoS 性能反馈 ,并将这些信息提供给内部模块。另外它还是节点与其他对等代理节点的接口。它使代理节点之间能互相传播、交换注册的服务信息和 QoS 数据。它接受用户提交的用户服务请求 ,通过匹配模块找到包含 QoS 需求的语义匹配的服务并返回执行结果。

3 匹配过程

服务发现是一个不断筛选、逐步求精的过程。

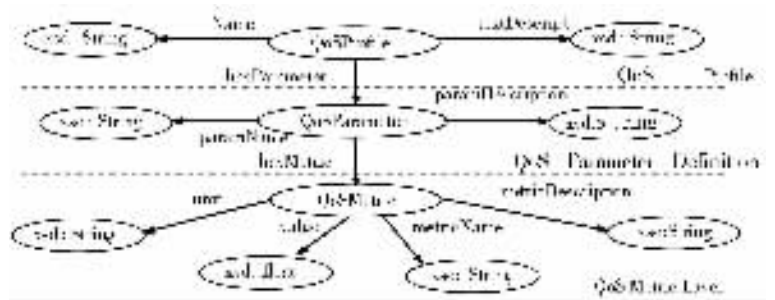


图 2 QoS 本体

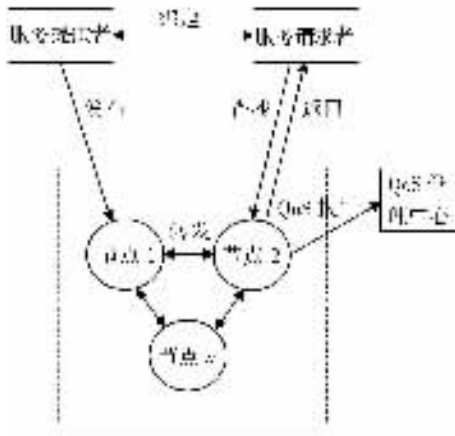


图 3 Web 服务架构

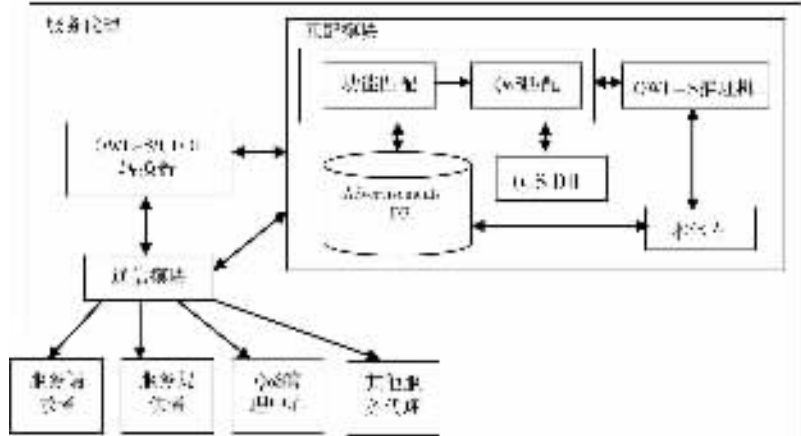


图 4 服务代理模块

整个匹配过程分为两个阶段。第一阶段为基于服务功能的匹配,第二阶段为基于服务 QoS 的匹配,以满足用户对服务质量提出的要求。

3.1 基于功能的匹配

服务的功能主要从输入(Inputs)、输出(Outputs)、前提(Preconditions)和结果(Effects)描述。通过本体描述,服务的功能属性可以映射到本体树中的概念节点。根据 GCSM^[6],利用本体树计算出两个概念节点的相似度

$$GCSM(uri1, uri2) = \frac{depth(uri1) + depth(uri2)}{2 \times depth(LCA(uri1, uri2))}$$

depth 函数代表本体树中一个节点的深度,并规定分类树根节点的深度为 0,每扩展一层,节点深度加 1。例如,depth(Machine) = 0,depth(Car) = 2。LCA(Lowest Common Ancestor, 最低共同祖先)函数是指本体树中两个概念节点共同祖先中深度最深的祖先节点。把 GCSM 所代表的概念的语义距离求倒数得到两个概念节点的语义相似度

$$sim(uri1, uri2) = \frac{1}{GCSM(uri1, uri2)}$$

IOPE 的相似度公式为

$$Sim_{fun}(req, adv) = \mu_1 Sim_{in}(req, adv) + \mu_2 Sim_{out}(req, adv) + \mu_3 Sim_{pre}(req, adv) + \mu_4 Sim_{effect}(req, adv)$$

其中 Sim_{in}、Sim_{out}、Sim_{pre}、Sim_{effect} 分别是输入、输出、前提和结果的相似度。μ₁、μ₂、μ₃ 和 μ₄ 代表输入、输出、前提和结果等属性在整个功能匹配中所占的比例,且 $\sum_{j=1}^4 \mu_j = 1$ 。

当广告服务与请求服务的功能相似度 Sim_{fun} 小于最小相似度的阈值 T_{min} 时,该广告服务不符合请求者的要求,需要从服务候选集中删除该服务,得到

阶段的服务匹配结果。如果服务需求者对服务质量有要求,则进入下一阶段,基于服务质量匹配。

3.2 服务 QoS 属性的匹配

得到 QoS 属性权值后,因为多个 QoS 属性之间不可公度性和矛盾性,很难认为 QoS 属性加权值越高服务越好。所以把 QoS 属性分为两种类型:一种是减量型的 QoS 属性,如时间、费用,它们值越大对于整体的服务质量贡献越小;另一种是增量型 QoS 属性,如可靠性,它的值越大对整体服务质量的贡献越大。

假设经过功能匹配后得到阶段性候选服务集 $S = \{S_1, S_2, \dots, S_n\}$,服务质量属性集 $A = \{time, cost, reliability\}$ 。利用集合 S 和 A 构造矩阵 $Q = (Q_{ij})_{n \times 3}$,Q_{ij} 表示第 i 个服务的第 j 个质量属性的取值。对于 time 和 cost 采用公式(1)计算, reliability 采用公式(2)计算^[6-7]

$$Y_{ij} = \begin{cases} \frac{Q_j^{max} - Q_{ij}}{Q_j^{max} - Q_j^{min}} & \text{当 } Q_j^{max} - Q_j^{min} \neq 0 \\ 1 & \text{当 } Q_j^{max} - Q_j^{min} = 0 \end{cases} \quad (1)$$

$$Y_{ij} = \begin{cases} \frac{Q_{ij} - Q_j^{min}}{Q_j^{max} - Q_j^{min}} & \text{当 } Q_j^{max} - Q_j^{min} \neq 0 \\ 1 & \text{当 } Q_j^{max} - Q_j^{min} = 0 \end{cases} \quad (2)$$

其中 $Q_j^{max} = \text{Max}(Q_{ij}), 1 \leq i \leq n; Q_j^{min} = \text{Min}(Q_{ij}), 1 \leq i \leq n$ 。

得到新的矩阵 $Y = (Y_{ij})_{n \times 3}$,利用公式 $Sim_{qos}(req, adv) = \sum_{j=1}^3 Y_{ij} \times \omega_j$ 算出每个服务 QoS 的综合评价价值。ω_j 为各质量属性的权重,且 $\sum_{j=1}^3 \omega_j = 1$ 。Sim_{qos}(req, adv) 值的范围限定在 [0, 1]。Sim_{qos}(req, adv) 的值越趋近 1,说明服务的 QoS 相对越好。服务的

通信模块负责将具有高 QoS 值的服务返回给用户。

为了验证本文模型的合理性,进行了模拟实验。到目前为止,缺乏一个语义 Web 服务相关的标准平台和测试数据集,且网络上很少语义标注的 Web 服务。本文采用人工标注已有 Web 服务的方法产生测试集。从网络上下载 WSDL 描述的服务,用扩展的 OWL-S 进行语义标注,生成了 200 多个 Web 服务。在基于功能匹配阶段将最小相似度阈值 T_{min} 设置为 0.7。从实验结果来看基于关键字的服务发现查全率比较低,但查准率比较高。本文算法的查全率为 86%,基于关键字匹配算法查全率为 20%;本文算法的查准率为 60%,基于关键字匹配算法查准率为 56%。本文设计的模型的查全率比 UDDI 上基于关键字的匹配高 66%,在不失查准率的基础上提高了查全率。

4 结论

本文基于 QoS 对 Web 服务发现进行了探讨^[8-9],设计了一种基于 SOA 工作框架的 Web 服务发现框架。在 OWL-S 的基础上扩展了 profile 描述,引入了 QoS 本体。提出了基于服务功能和 QoS 的两阶段匹配模型。在此模型中不但实现了服务发现的语义匹配,而且能满足用户的 QoS 需求。通过实验分析,该模型可以有效地提高服务发现的有效性。

参考文献:

- [1] Liu Y T, Ngu A H H, Zeng L Z. QoS computation and policing in dynamic web service selection[C]. Feldman S I, Uretsky M, Najork M, et al. Proc of the 13th Int'l Conf on World Wide Web (WWW 2004). New York: ACM Press, 2004: 66-73.
- [2] 杨胜文, 史美林. 一种支持 QoS 约束的 Web 服务发现模型[J]. 计算机学报, 2005, 28(4): 589-594.
- [3] Ran S. A model for web services discovery with QoS[J]. ACM SIGCOM Exchanges, 2003, 4(1): 1-10.
- [4] Gruber T R. Towards principles for the design of ontologies used for knowledge sharing[J]. International Journal of Human-Computer Studies, 1995, 43: 907-928.
- [5] 任波. 基于语义的 Web 服务发现研究[D]. 浙江: 浙江工业大学, 2005.
- [6] Ganesan P, Molina H G, Widom J. Exploiting hierarchical domain structure to compute similarity[J]. ACM Transactions on Information Systems, 21(1): 64-93.
- [7] Ran S. A model for web services discovery with QoS[J]. ACM SIGecom Exchanges, 2003, 4(1): 1-10.
- [8] 曹进明, 李广泉. 基于 SOA 的中小型企业 IT 系统设计与实现[J]. 重庆师范大学学报: 自然科学版, 2008, 25(4): 55-59.
- [9] 陈蜀宇, 刘刚国. 面向 Web 服务的数字化营区系统架构[J]. 重庆工学院学报: 自然科学版, 2008, 22(9): 103-107.

Discovery Method of Extension Semantic Web Service Based on the QoS

LI Xiao-lin¹, ZHANG Li-na², LI Wei-bin¹

(1. Dept. of Information Engineering; 2. Dept. of Mathematics and Information Science, Xianyang Normal University, Xianyang Shanxi 712000, China)

Abstract: How to discover services for ours is the key problem for the application of the Web service technology. However, current descriptions of Web services lack QoS information, can not meet the needs of the user's QoS of services. In this paper, after extending the profile description based on OWL-S, the QoS was introduced into OWL-S. Based on the SOA framework and P2P networks, the Web service discovery framework is designed, and the QoS management module was introduced into. The matching process is divided into two phases. The first is based on function. The service function is described from inputs, outputs, preconditions and effect; the second is based on services QoS. QoS attribute is described from the time, cost, and reliability. The experimental result shows this model of Web service discovery greatly increases the recall rate, and improves the performance of Web service discovery.

Key words: Web services; QoS ontology; OWL-S; services discovery model