

基于语义匹配和 QoS 的 Web 服务混合选择方法

张佩云¹ 黄 波¹ 孙亚民¹

(1 南京理工大学计算机科学与技术学院,南京市孝陵卫 200 号,210094)

摘 要:提出了一种 Web 服务语义匹配方法,在此基础上分析了基于语义匹配和基于 QoS(服务质量)的服务选择策略,以及面向独立服务和组合服务的服务选择方法,并给出了结合语义匹配和 QoS 指标值的 Web 服务混合选择策略及算法。实例分析结果表明,该混合选择策略能针对不同的服务需求作相应的处理,增加了服务选择的正确性,提高了服务组合效率。

关键词:Web 服务;语义匹配;语义匹配选择;QoS 选择

中图法分类号:P208

目前,电子商务活动往往依赖第三方提供的应用和服务^[1]。随着 Web 服务数量的日益增加,用户在请求相应的 Web 服务时可能不知道存在什么样的服务,而服务在制定时也不能预期都会存在什么样的请求,当服务消费者寻求特殊的 Web 服务时,仅仅基于句法匹配的服务选择已经不能满足用户的需求,因此有必要结合语义匹配和服务的 QoS 指标来选择恰当的服务。基于语义及服务的 QoS 进行服务选择是服务组合研究的一个热点^[2-7]。本文结合服务的 QoS 属性,提出了基于语义匹配度的服务混合选择机制。

1 基本概念

定义 1 语义距离。概念 c_1 与概念 c_2 间的语义距离指两个概念在本体中的最短几何距离,其值为两个概念结点间最短几何距离中边的数量,为 $[0, \infty)$ 之间的整数。

定义 2 语义匹配度。语义匹配度指概念的相似度,两个概念的相似度值越大,则语义匹配度越高。语义匹配度离不开本体中概念间的语义距离,语义距离的大小决定了两个概念间的匹配程度。两个概念的语义距离越大,其相似度越低;反之,其相似程度越大^[2]。相似度为语义距离的单调递减函数,语义距离为 0 时,其相似度为 1;语义距离为无穷大时,相似度为 0。

定义 3 语义 Web 服务。语义 Web 服务是指用 OWL 进行语义描述的 Web 服务。目前,Web 服务的语义描述方面最重要的工作是 OWL-S, OWL-S 通过 ServiceProfile、Service-Model 和 ServiceGrounding 三个方面来刻画 Web 服务,其核心语义模型主要包括服务的输入、输出、输入的前提条件以及输出的效果等。

定义 4 服务语义关联。对于两个服务 WS_i 和 WS_j ,若服务 WS_i 中存在输出参数是 WS_j 输入参数的直接或间接子类或是同一个概念,称为服务语义关联^[8],记作 $out_{WS_i} \leq in_{WS_j}$, out_{WS_i} 表示 WS_i 的一个输出参数, in_{WS_j} 表示 WS_j 的一个输入参数,并称 WS_i 为前趋服务, WS_j 为后继服务。如果这两个参数在领域本体是同一个概念,记作 $out_{WS_i} \equiv in_{WS_j}$ 。

定义 5 服务组合。一个服务组合是指能够满足用户服务目标的一个偏序发布服务序列 $(WS_1, WS_2, \dots, WS_n)^{[9]}$ 。在定义 4 的基础上,要求:① $in_req \leq in_WS_1$; ② $out_WS_n \leq out_req$; ③ $out_WS_1 \leq in_WS_2, out_WS_2 \leq in_WS_3, \dots, out_WS_{n-1} \leq in_WS_n$ (序列满足偏序关系)。其中, $in_req = \{I_1, I_2, \dots\}$ 表示请求服务的输入参数集; $out_req = \{O_1, O_2, \dots\}$ 表示请求服务的输出参数集; out_WS_{n-1} 与 in_WS_n 分别表示发布服务 WS_{n-1} 的输出参数集及 WS_n 的输入参数集。为提高服务组合的正确性,要求 $in_req \supseteq in_WS_1, out_WS_{n-1} \supseteq out_WS_n$, 且 $(in_req \cup out_WS_1 \cup \dots \cup out_WS_{n-1}) \supseteq out_WS_n$ 。

$WS_{n-1} \cup out_WS_n) \supseteq out_req。$

2 语义 Web 服务混合选择策略

2.1 Web 服务的语义匹配

采用基于本体的用户需求描述、服务功能及 QoS 描述,在此基础上提出的 Web 服务语义匹配包括三个方面的内容:① 服务类别语义匹配;② 请求服务与发布服务语义匹配;③ 前趋服务的输出参数与后继服务的输入参数的语义匹配(简称 IO 语义匹配)。对此,提出如下服务选择策略:① 基于语义匹配的服务选择。当存在单独的发布服务时,其接口参数集与用户请求服务接口参数集相同,结合服务的类别匹配、请求服务和发布服务在输入/输出接口上的语义匹配结果,选择功能和接口语义匹配最佳的发布服务;当没有找到满足用户需求的独立服务时,基于语义匹配度,选择接口之间有服务语义关联的服务;② 基于 QoS 的服务选择。当基于语义匹配的服务选择产生多个具有相同功能的候选 Web 服务时,需要针对服务的 QoS 指标来对服务加以选择。

2.2 基于语义匹配的服务选择策略

基于语义匹配度的服务选择面临着两种情况:① 经过服务类别匹配和请求服务与发布服务的接口匹配,匹配到独立服务;② 当没有独立的服务满足用户需求时,需要服务组合,包括四个方面的内容:请求服务的类别与发布服务在类别上的语义匹配;请求服务与发布服务输入参数的语义匹配;请求服务与发布服务输出参数的语义匹配;前趋服务的输出参数与后继服务的输入参数的语义匹配(称 IO 匹配)。

2.2.1 选择满足用户需求的独立服务

在匹配到满足用户需求的独立发布服务后,由于匹配结果是一个候选集,需要采用相应的服务选择策略来选择最佳匹配服务。本文基于语义匹配的选择策略是对三个方面语义匹配结果的综合:请求服务类别与该发布服务类别的语义匹配;请求服务输入接口与发布服务输入接口的语义匹配;请求服务输出接口与发布服务输出接口的语义匹配。

考虑到权重时,基于语义匹配度的服务选择公式如下:

$$M_s(req, ad) = \sum_{i=1}^3 M_i W_i / \sum_{i=1}^3 W_i \tag{1}$$

式中, M_s 为语义匹配值(req 为用户请求服务, ad 为已发布的服务); M_i 为第 i 个指标对应的语义匹配值; W_i 为第 i 个指标对应的权重; $i=1, 2, 3$ 分别表示请求服务与发布服务在服务类别、输入参数集、输出参数集上的语义匹配。

需要注意的是,此处 M_2 和 M_3 分别对应服务接口集合的语义匹配度,而不是指单个参数间的匹配度, $M_s(req, ad)$ 是利用各指标对应的语义匹配值及其相应的权重的综合结果。

当存在多个功能都满足的单个服务时,按服务的 QoS 值来选择适当的服务。

2.2.2 选择用于服务组合的服务

当不存在或没有匹配到满足用户需求的独立服务时,则需要选择满足 IO 匹配的服务,以实现服务组合。

选择参与服务组合的服务时,基于语义匹配的选择策略涉及三方面的内容:请求服务与发布服务在输入参数上的语义匹配;请求服务与发布服务在输出参数上的语义匹配;前趋服务的输出参数与后继服务的输入参数的语义匹配。这三类匹配的表示及涵义如表 1 所示。

表 1 Web 服务匹配函数

Tab. 1 Matching Functions of Web Services

函数	功能	参数 1	参数 2
InputMatching(in_req, in_ad_1)	服务请求输入集与发布服务 ad_1 输入集的语义匹配	in_req : 请求服务的输入参数集	in_ad_1 : 与请求服务的输入有语义关联的发布服务 ad_1 的输入参数集
OutputMatching(out_req, out_ad_n)	服务请求输出集与发布服务 ad_n 输入集的语义匹配	out_req : 请求服务的输出参数集	out_ad_n : 与请求服务的输出有语义关联的发布服务 ad_n 的输出参数集
IOMatching(out_ad_i, in_ad_{i+1})	前趋服务与后继服务集合接口参数的匹配	out_ad_i : 第 i 个发布服务的输出参数集	in_ad_{i+1} : 第 $i+1$ 个发布服务的输入接口参数集

2.3 基于 QoS 的服务选择策略

QoS 描述了一个服务满足消费者需求的能力,提供有保证的 QoS 是 Web 服务在商业应用中获得成功的关键^[10]。本文主要采用国际标准组织所给的指标 QoS(cost, time, availability, reliability)。在下一步研究中,将逐步增加对其他

QoS 指标的支持。基于 QoS 指标的服务选择计算公式如下:

$$\begin{aligned} M_{Q_1} &= \min(\sum_{i=1}^2 Q_i W_i / \sum_{i=1}^2 W_i) \\ M_{Q_2} &= \max(\sum_{i=3}^4 Q_i W_i / \sum_{i=3}^4 W_i) \end{aligned} \tag{2}$$

式中, M_{Q_1} 、 M_{Q_2} 为两类 QoS 匹配值; Q_i 为第 i 个指标的 QoS 匹配值; W_i 为第 i 个指标的 QoS 权重; $i=1$ 表示服务价格; $i=2$ 表示服务的执行时间; $i=3$ 表示服务的可用性; $i=4$ 表示服务的可靠性。

涉及的 QoS 指标不同, 计算方法也不同。本文将 QoS 指标归类为效益型和成本型, M_{Q_1} 对应的是最小服务价格和执行时间匹配值之和, 属于成本型 QoS; M_{Q_2} 是最大服务可用性和可靠性匹配值之和, 属于效益型 QoS。当存在功能相同的服务需要选择时, 将根据 M_{Q_1} 和 M_{Q_2} 的值对 Web 服务功能相似的服务进行选择。

2.4 结合语义匹配与 QoS 的服务混合选择策略及算法

采取基于语义匹配度和 QoS 指标的服务混合选择策略, 在服务选择的过程中, 需兼顾服务的语义匹配信息和 QoS 信息, 并通过对两者的综合评价而得到最佳服务。Web 服务的综合选择根据语义匹配度和基于 QoS 的指标值及权值来决定。

语义 Web 服务组合组件由原子服务和控制两部分组成, 控制结构如顺序、并行、选择、循环等。语义 Web 服务组合可通过类 BNF 范式的符号定义, 代数操作符的语法如下:

$$S ::= X \mid S_c \odot S_{c+1} \mid S_c \oplus S_{c+1} \mid S_c \diamond S_{c+1} \\ \mid \mu S \mid S_c \parallel S_{c+1}$$

其中, X 表示一个原子服务或空服务 (即一个服务没有执行任何操作); $S_c \odot S_{c+1}$ 表示一个组合服务是顺序执行 S_c 和 S_{c+1} 后形成的; $S_c \oplus S_{c+1}$ 表示一个组合服务是执行 S_c 或 S_{c+1} 后形成的 (不可兼或); $S_c \diamond S_{c+1}$ 表示一个组合服务是按 S_c 和 S_{c+1} 的顺序执行或按 S_{c+1} 和 S_c 的顺序执行 (即无序执行), 等价于 $S_c \odot S_{c+1} \oplus S_{c+1} \odot S_c$; μS 表示一个组合服务是由 S 循环执行 μ 次; $S_c \parallel S_{c+1}$ 表示一个组合服务是由 S_c 和 S_{c+1} 并发执行。

针对以上控制结构, 结合语义匹配和 QoS 指标, 选择满足用户需求的独立服务的混合选择算法流程图如图 1 所示。图 1 中, 算法 chooseServiceforComposition(req, ad) 用于实现面向组合服务的服务选择, 该算法是结合语义匹配和 QoS 指标的组服务混合选择算法, 其流程图如图 2 所示。图 2 中主要考虑了顺序、无序执行、选择、循环四种控制结构, 其中顺序和无序执行被合并考虑。从算法 2 可以看出 QoS 与语义匹配二者之间的综合影响。通过对服务的语义匹配度和 QoS 值的综合考虑, 减少了候选服务数目, 提高了服务选择的准确率。

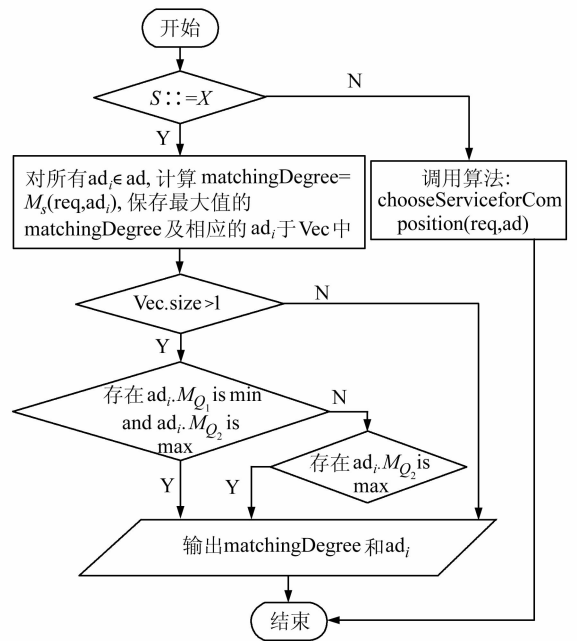


图 1 chooseService(req, ad) 的流程图
Fig. 1 Flow Chart of chooseService(req, ad)

3 实例及实验分析

3.1 实例分析

选择用于服务组合的服务的算法实例如图 3 所示。图中, WS_1 、 WS_2 、 WS_3 与 WS_4 是服务组合中的四个抽象服务; $WS_{i,j}$ ($i=1, \dots, 4; j=1, \dots, n$) 是具体候选服务; \rightarrow 表示控制流, 在图 3 中体现的是一种顺序结构; \rightarrow 表示语义匹配选择及 QoS 选择。

按定义 5 及 Web 服务语义匹配获取各处抽象服务 i 的候选服务集 $WS_{i,j}$ ($j=1, \dots, n$), 对同一个抽象服务而言, 由于服务语义匹配后所获得的候选服务集的服务功能相似, 但 QoS 的值会有所不同, 因此按图 2 中的两类 QoS 值的大小选择最佳 QoS 性能的服务。

3.2 实验分析

实验环境为 Intel Pentium 5、512 M 内存及 Windows XP, 算法用 Java 语言实现。

将本文提出的混合选择算法 (hybrid service selection method, HYSSM) 与基于句法匹配的服务选择 (syntax-based service selection method, SYSSM) 及 Seog-Chan Oh 提出的基于语义匹配的服务选择算法 (semantic-based service selection method, SESSM)^[7] 相比较。针对候选服务数分别为 10、15、20、25、30、35、40 的服务选择流程进行仿真, 每个候选服务拥有 4 个 QoS 指标

值,三个算法分别运行 10 次,平均精确性的对比结果如图 4 所示。

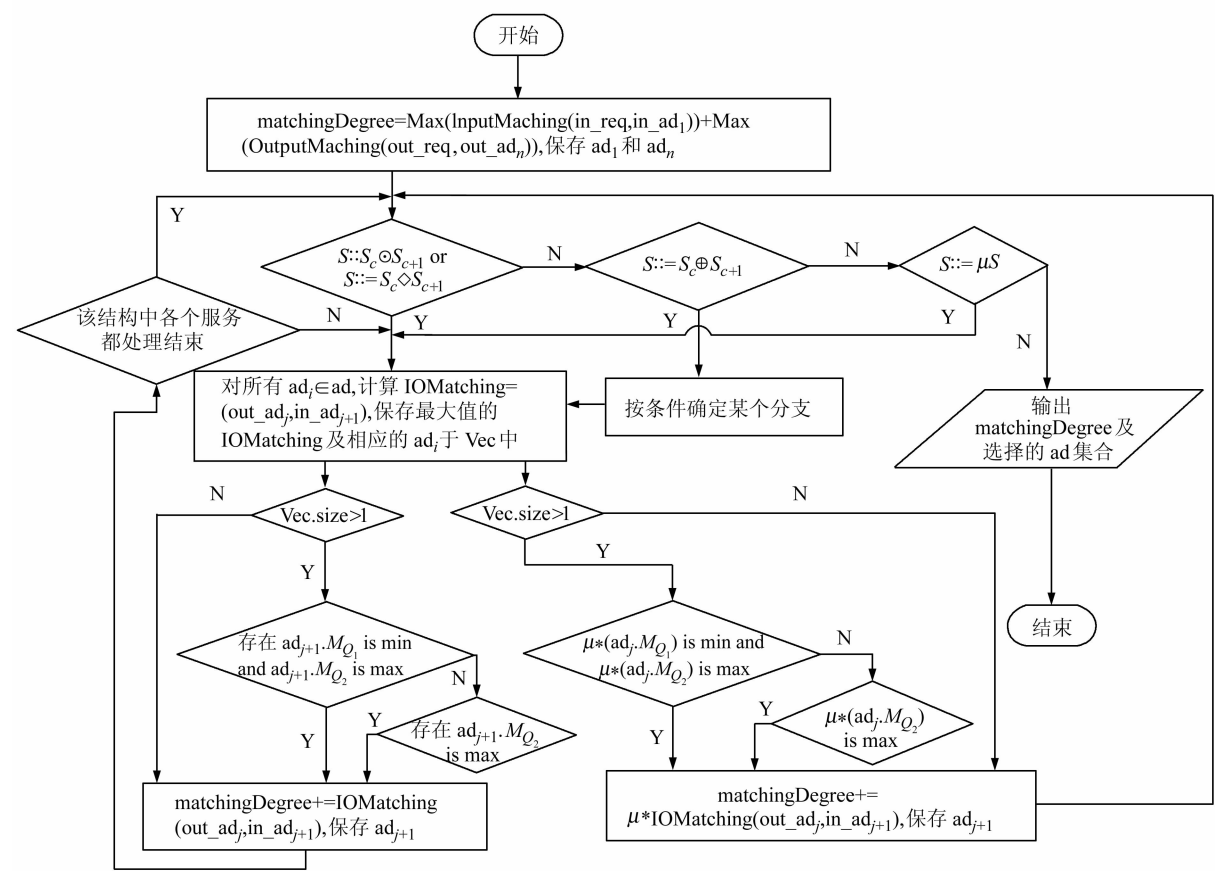


图 2 chooseServiceforComposition(req,ad) 的流程图

Fig. 2 Flow Chart of chooseServiceforComposition(req,ad)

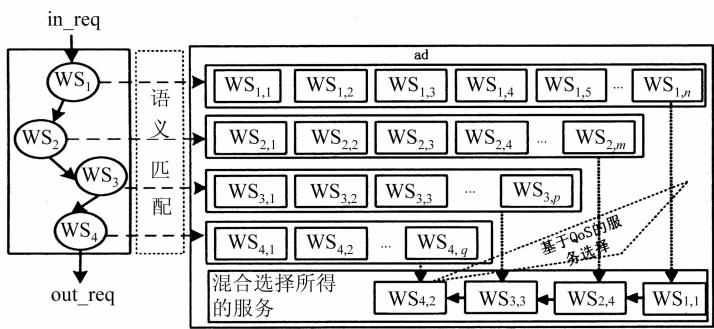


图 3 选择用于服务组合的服务的实例分析

Fig. 3 Instance Analysis of Choosing Services for Web Services Composition

实验结果表明,SYSSM 没有考虑到服务的语义,具有很大的局限性,其算法性能比其他两个算法差,这与该算法只进行数据匹配及关键词匹配有关。SESSM 算法的性能随着测试集波动较大,这与该算法只涉及服务的功能而没有涉及到服务的 QoS 约束有关,因为服务的 QoS 指标决定了服务的重要程度,虽然语义匹配的结果比基于句法的匹配结果要好,但由于缺乏 QoS 的支持,致使其性能波动较大,影响了服务选择的结果。HYSSM

混合选择算法在各个测试集合里面一直保持较好的性能,在 QoS 指标的支持下,找到满足语义需要的服务,较好地满足了用户的需求。

4 结 语

本文提出的基于语义匹配和 QoS 的服务混合选择策略,兼顾语义匹配度和 QoS 值进行服务的选择,减少了候选服务的数目,提高了服务选择

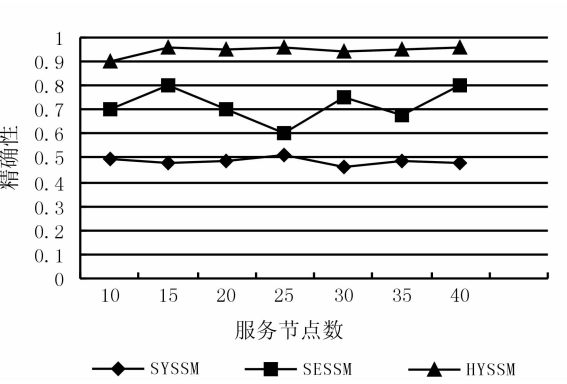


图 4 三种算法的服务选择精确性比较
Fig. 4 Precision Comparison of Services Selection of Three Algorithms

的精确度,为产生增值效果的动态服务组合奠定了良好的基础,促进了 Web 服务的应用。进一步的研究是提高服务选择精确度的途径,并逐步增加其他 QoS 指标对服务选择的支持。

参 考 文 献

[1] Snoeck M, Lemahieu W, Goethals F, et al. Events as Atomic Contracts for Component Integration[J]. Data & Knowledge Engineering, 2004, 51(1): 81-107

[2] 代钰,杨雷,张斌,等. 支持组合服务选取的 QoS 模型及优化求解[J]. 计算机学报, 2006, 29(7): 1 167-1 178

[3] Zhu Junwu, Wang Jiandong, Li Bin. Web Services

Selection Based on Semantic Similarity[J]. Journal of Southeast University (English Edition), 2006, 22(3): 297-301

[4] 陈彦萍,李增智,郭志胜,等. Web 服务组合中基于服务质量的服务选择算法[J]. 西安交通大学学报, 2006, 40(8): 897-900

[5] Serhani M A, Dssouli R, Hafid A, et al. A QoS Broker Based Architecture for Efficient Web Services Selection[C]. IEEE International Conference on Web Services, Orlando, USA, 2005

[6] Grønmo R, Jaeger M C. Model-Driven Semantic Web Service Composition[C]. The 12th Asia-Pacific Software Engineering Conference, Taipei, 2005

[7] Oh S C, Kil H, Lee D, et al. Algorithms for Web Services Discovery and Composition Based on Syntactic and Semantic Service Descriptions[C]. The 8th IEEE International Conference on E-Commerce Technology, Palo Alto, California, USA, 2006

[8] The OWL Services Coalition. OWL-S: Semantic Markup for Web Services [OL]. <http://www.daml.org/services/owl-s/1.0/>, 2003

[9] 李曼,王大治,杜小勇,等. 基于领域本体的 Web 服务动态组合[J]. 计算机学报, 2005, 28(4): 644-650

[10] 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

第一作者简介:张佩云,博士生,讲师。主要研究方向为语义网、Web 服务和智能信息处理。
E-mail: njustzpy@yahoo. com. cn

A Hybrid Selection Approach for Web Services
Based on Semantic Matching and QoS

ZHANG Peiyun¹ HUANG Bo¹ SUN Yamin¹

(1 School of Computer Science & Technology, Nanjing University of Science & Technology,
200 Xiaolingwei, Nanjing 210094, China)

Abstract: An approach for Web service semantic matching is proposed, two kinds of services selecting strategies of based on semantic matching and QoS are focused. The services selection approach towards the individual Web service and the services for composition is analyzed. The hybrid selection strategy is put forward and the corresponding algorithm is proposed. The results show that the hybrid selection approach can deal with different requests, increase the correctness and improve the efficiency of services composition.

Key words: Web services; semantic matching; semantic matching selection; QoS selection

About the first author: ZHANG Peiyun, Ph. D candidate, lecturer, majors in semantic Web, Web services and intelligent information processing.
E-mail: njustzpy@yahoo. com. cn