

面向服务组合的服务语义匹配机制

张佩云¹, 黄波², 孙亚民²

(1. 安徽师范大学数学与计算机科学学院 安徽 芜湖 241000; 2. 南京理工大学计算机科学与技术学院 南京 210094)

【摘要】为提高服务组合中服务匹配速度和正确率,提出了基于服务索引的服务语义匹配机制。在面向独立服务匹配算法基础上,引入服务索引机制,提出面向服务组合的语义匹配算法,并在算法中加入了对服务输入/输出匹配的可用性分析。为进一步提高服务匹配的精确性,引入了语义匹配度概念,以实现具有多种匹配值的服务类别语义匹配,继而给出综合语义匹配值的计算方法。实验分析表明基于服务索引的语义匹配能有效地提高服务匹配效率。

关键词 服务组合; 服务索引; 语义匹配度; 服务语义匹配
中图分类号 TP393 文献标识码 A

Service Semantic Matching Mechanism for Web Services Composition

ZHANG Pei-yun¹, HUANG Bo², and SUN Ya-min²

(1. School of Mathematics and Computer Science, Anhui Normal University Wuhu Anhui 241000;

2. School of Computer Science and Technology, Nanjing University of Science and Technology Nanjing 210094)

Abstract To improve the speed and correctness of Web service matching for services composition process, a semantic matching mechanism is proposed with the help of services indexes. Based on indexes mechanism and services semantic matching algorithm for single service, a composition-oriented services semantic matching algorithm is proposed and the matching usability of service input/output is analyzed. To improve the accuracy of service matching, the semantic matching degree for services category matching is introduced, and then the computation method of overall semantic matching is provided. Experimental results show that this composition-oriented service semantic matching can enhance the efficiency of services matching based on services index.

Key words Services composition; services indexes; semantic matching degree; service semantic matching

关于Web服务及语义网技术的研究表明,自动匹配和利用Web服务以解决特定问题具有广泛的应用前景。如何匹配面向组合的Web服务以形成新的、满足不同用户需求的增值服务成为新的应用需求和研究热点^[1]。对于日益增多的Web服务,若不采取恰当的应对措施,则无法快速地从大量服务中匹配到需要的服务。因此,从大量服务中实现面向组合的语义匹配是一个重要的研究课题。

服务语义匹配研究可以分为两类:(1)面向独立服务的服务匹配;(2)面向服务组合的服务匹配。两类匹配存在一些交叉内容,但是第2类服务匹配的结果更严格,可避免误匹配导致误组合。

针对第1类匹配,文献[2-3]提出了一种语义匹配算法,给出了4种匹配,能够识别服务请求与服务发

布之间的输入/输出匹配的4种程度。但该算法的缺点是匹配分类不足、不精细。在文献[2-3]提出的服务匹配算法的基础上,文献[4]给出了服务匹配的可用性。文献[5]和文献[6]对文献[2-3]的服务匹配算法进行了扩充,前者增加了容器(container)和部分匹配(part of),后者增加了非空不相交匹配。但这些研究都是基于有限类别的匹配,没有涉及到语义匹配度。

针对第2类服务匹配,文献[7]给出了一种快速的服务发现和组合方法,该方法基于服务的句法和语义,可实现服务匹配。但该方法建立在服务索引的基础之上,未考虑到面向组合的服务匹配中前趋服务与后继服务之间的匹配。文献[8]给出了面向组合的服务匹配,主要是针对接口的参数匹配,基于文献[2-3]服务匹配中的4类匹配情况进行匹配,没有按

收稿日期: 2007-06-17; 修回日期: 2008-04-12

基金项目: 教育部高等学校博士点科研基金(20050288015); 安徽省高校省级自然科学基金重点项目(2006KJ024A)

作者简介: 张佩云(1974-),女,博士,讲师,主要从事语义网、服务计算和智能信息处理方面的研究。

照服务匹配优先级进行服务匹配的过滤。文献[9]给出了一种基于服务索引的服务匹配方法,但该文没有结合语义匹配度实现服务类别的匹配,也没有对完全匹配与包含匹配进行区分。

本文在文献[2-3]服务匹配算法的基础上,针对面向组合的服务匹配引入服务的索引机制,促进服务查找的效率,并将服务接口的参数匹配细分成完全匹配和包含匹配两种优先级,且前者的优先级高于后者,可提高语义匹配的正确率。实现面向组合的服务语义匹配时,主要考虑服务的接口的语义匹配和基于语义匹配类别的语义匹配,关注服务匹配的语义兼容性和匹配的速度。

1 Web服务索引机制

为在大量可用的Web服务中快速匹配所需服务,需要对可用服务进行精确的前期处理,采用对服务参数和服务类别建立索引的方法实现快速查找服务。服务索引与面向组合的服务匹配之间的关系如图1所示。

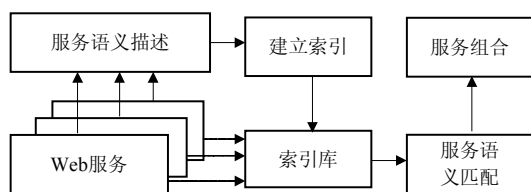


图1 服务索引与面向组合的服务匹配

图1中的“建立索引”主要针对服务的输入、输出参数和服务类别,服务索引结构如图2所示。

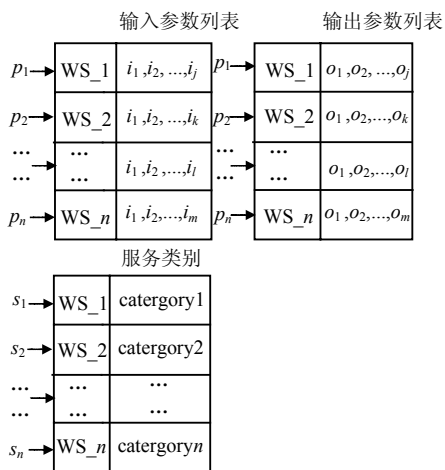


图2 服务索引结构

图2中, $p_x (x=1,2,\dots,n)$ 表示服务参数(输入参数或输出参数),通过该参数可以在索引表里进行服务检索; $WS_x(x=1,2,\dots,n)$ 表示已索引服务; $s_x(x=1,2,$

$\dots,n)$ 表示待查找的服务。索引结构可通过哈希表实现。服务索引建立后,系统进行服务匹配时可多次访问该索引,当知识库被更新时,需要重新建立服务索引,以保持数据的一致性和最新性。通过服务索引,加快了服务查找的速度,可很容易地获取具有某个输入参数、输出参数和服务类别的Web服务。

2 面向服务组合的服务语义匹配算法

因为面向独立服务的服务匹配与面向组合的服务匹配有交叉点,本文将首先分析面向独立服务的匹配算法及其不足,对其进行改进,然后在引入服务索引机制的基础上,实现面向服务组合的服务语义匹配。

2.1 面向独立服务的服务匹配算法

这类算法有很多,文献[2-3]提出的匹配算法为研究者广泛引用。该算法将匹配分为完全匹配、插拔匹配、匹配和失败4种匹配结果。为了描述匹配机理,令 $out_r[i]$ 和 $out_a[j]$ 两个概念分别对应用户请求服务(简称为 r)的一个输出参数和发布服务(简称为 a)的一个输出参数,匹配分类及其描述如表1所示。

匹配分类	匹配含义	匹配值
完全匹配(exact)	$out_r[i] \equiv out_a[j]$	1
包含匹配(subsume)	$out_r[i] \supset out_a[j]$	1/2
插拔匹配(plugin)	$out_r[i] \subset out_a[j]$	1/3
匹配失败(fail)	其他情况	0

该服务匹配算法的伪代码如下。

```

算法1 Matching( $out_r[i], out_a[j]$ )
if ( $out_r[i] \equiv out_a[j]$ ) return 1;
else if ( $out_r[i] \supset out_a[j]$ ) return 1/2;
else if ( $out_r[i] \subset out_a[j]$ ) return 1/3;
else return 0;
    
```

该算法实质是基于4种匹配值的匹配方法,主要针对独立服务的接口参数的语义匹配而非面向组合的服务匹配,因为该算法没有考虑到服务输出与服务输入之间的语义匹配关系,无法满足针对动态服务组合的服务语义匹配。

2.2 基于索引面向组合的服务匹配

引入服务索引并对算法1进行改进,实现面向组合的服务语义匹配,包括服务参数匹配及服务类别匹配,具体描述为:

(1) 匹配1 请求服务的输入参数与组合服务的输入间的语义匹配;

(2) 匹配2 前趋服务的输出与后继服务的输入之间的语义匹配;

(3) 匹配3 请求服务的输出参数与组合服务的输出之间的语义匹配;

(4) 匹配4 基于语义匹配度的服务类别的语义匹配。

上述4种匹配中, 前3种具有相同的本质, 都需满足参数间的完全匹配(exact)或是包含关系(subsume)^[10], 但3者在服务索引列表中进行查找时有所区别。对于匹配1, 根据请求服务的输入参数在服务索引输入列表中查找; 对于匹配2, 根据前趋服务的输出参数在服务索引的输入列表中查找; 对于匹配3, 根据请求服务的输出参数在服务索引输出列表中查找; 对于匹配4, 根据服务类别在服务类别索引列表中查找。以匹配2为例, 前趋服务(pre-service, 简写为pre)的输出与后继服务(sub-service, 简写为sub)的输入之间的匹配如图3所示。

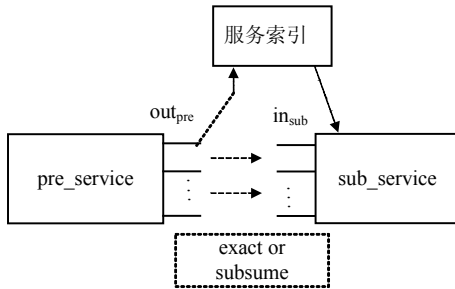


图3 基于服务索引的前趋服务与后继服务语义匹配

图3中, out_{pre} 和 in_{sub} 分别表示前趋服务的输出参数集和后继服务的输入参数集。为减少误匹配所导致的误组合, 两个参数间的语义匹配要求满足算法1的完全匹配或包含(subsume)关系。算法伪代码算法2所示。

算法2 IOMatching(out_{pre})

输入: 一个前趋服务的输出参数集 out_{pre} 。

输出: IO语义匹配度及后继服务等。

步骤1 若 out_{pre} 中参数个数为0, 即 $|out_{pre}|=0$, 则退出该算法, 否则转步骤(2)。

步骤2 对 out_{pre} 中的输出参数 $out_{pre}[i]$, 在服务输入索引中查找满足 $out_{pre}[i] \equiv in_{sub}[j]$ 的后继服务 $sub_service$, 若找到, 转步骤3; 否则转步骤4。

步骤3 计算 $usability = |out_{pre}| / |in_{sub}|$, 若 $usability \leq 1$;

(1) 计算两个参数间的语义匹配度 $IOMatch[i] = Matching(out_{pre}[i], in_{sub}[j])$;

(2) 对于 out_{pre} 中的其余的参数 $out_{pre}[k]$, 在 $sub_service$ 对应的输入服务列表中进行类似于 $out_{pre}[i]$ 的处理, 即寻找与 $out_{pre}[k]$ 完全匹配或包含关系的 $in_{sub}[l]$;

(3) 重复(2), 直到 out_{pre} 中所有参数都处理完, 转步骤5。

步骤4 在 本体树中查找满足 $out_{pre} \supset onto_k$ 关系的 $onto_k$, 若不存在该 $onto_k$, 则退出该算法, 否则在服务输入索引中查找满足 $onto_k \equiv in_{sub}[j]$ 的后继服务; 找到满足条件的后继服务, 转步骤3; 否则转步骤4。

步骤5 计算权语义匹配度:

$$result = \frac{\sum_{i=1}^n IOMatch[i] * W[out_{pre}[i]]}{\sum_{i=1}^n W[out_{pre}[i]]}$$

步骤6 将 $result$ 、 out_{pre} 、 in_{sub} 及 $usability$ 封装在 $vector$ 中, 返回 $vector$ 。

算法2中, 对于前趋服务的 out_{pre} 中一个元素 $out_{pre}[i]$, 在服务索引的输入列表中遍历后继服务的输入集 in_{sub} , 搜索与元素 $out_{pre}[i]$ 完全匹配或被 $out_{pre}[i]$ 包含的 $in_{sub}[j]$ 所对应的后继服务, 当在该后继服务中存在与 out_{pre} 的所有参数完全或包含关系的输入参数时, 认为匹配成功。最终的前趋与后继语义匹配值将由 out_{pre} 中各个元素权重及各个元素与 $in_{sub}[j]$ 的匹配值综合。算法2中采用了匹配优先级, 完全匹配优先级高于包含关系, 将优先得到处理, 如步骤2中的“ $out_{pre}[i] \equiv in_{sub}[j]$ ”体现了完全匹配, 步骤4中的“ $out_{pre} \supset onto_k$ ”体现了包含关系。采用匹配优先级可提高服务匹配的正确性。算法2考虑了服务参数匹配的可用性, 即要求前趋服务的输出参数的个数限定为少于或等于后继服务的输入参数个数, 以提高匹配的正确性和减少组合复杂度。

2.3 基于索引的服务类别语义匹配

算法1只考虑了4种匹配情况, 没有考虑到更细致的语义匹配关系, 不能实现更精确的服务类别匹配。为此, 引入匹配度概念, 将4类分类机制扩展成无限分类, 以表示服务类别的任意值匹配。服务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。

定义1 服务参数的语义相似度为^[11]:

$$MatchingDegree(o_1, o_2) = 1 - d(o_1, o_2)$$

式中 $d(o_1, o_2)$ 为 o_1 、 o_2 两个服务参数概念之间的

语义距离, 且 $d(o_1, o_2)$ 取值于 $(0,1)$ 。

进行服务类别的语义匹配, 即计算 $\text{Matching Degree}(\text{cat}_{\text{pre}}, \text{cat}_{\text{sub}})$ 的值, 需要从服务类别索引中快

$$S = \frac{W_1 \times \text{IOMatching}(\text{out}_{\text{pre}}) + W_2 \times \text{MatchingDegree}(\text{cat}_{\text{pre}}, \text{cat}_{\text{sub}})}{W_1 + W_2}$$

式中 $\text{IOMatching}(\text{out}_{\text{pre}})$ 的值来自算法 2; $\text{Matching Degree}(\text{cat}_{\text{pre}}, \text{cat}_{\text{sub}})$ 的值来自定义 2; W_1 和 W_2 分别是各自的权重。

3 实验分析

为验证算法和策略, 采用如下实验环境: CPU 为 Pentium(R)4, 2.8 GHz, 内存 512 MB; 操作系统为 Windows XP。在 JDK1.5、Jbuilder9 和 Tomcat 5.5 软件环境下, 从服务匹配正确率和时间响应率上对关键词匹配、算法 1 和算法 2 进行比对, 实验结果如图 4 所示。从图 4 可知, 基于关键字服务匹配虽然在响应时间上比较好, 但是由于没有语义支持, 存在漏配, 其匹配正确率不高。算法 1 提出的服务匹配方法虽然采用了语义匹配机制, 在服务匹配的正确率上比前者好, 但是匹配到的服务不能适合于服务组合。算法 2 由于基于服务索引和面向服务组合, 正确率和响应时间相对于前两者都有所提高。由正确率和时间响应的结果可知, 由于基于服务索引的建立, 为后期的多次服务语义匹配提供了检索空间, 节省了查找时间, 从实验结果可知具有较好的服务匹配效率。

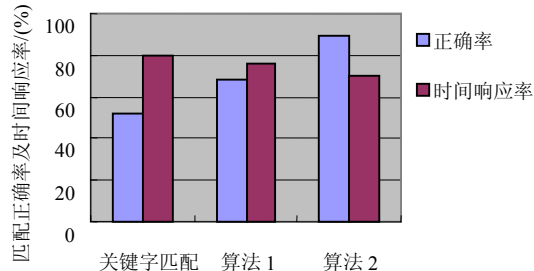


图4 服务匹配正确率及时间响应率比较

4 结论及展望

Web 服务以其特有的优势而其有广泛的应用前景。现有的面向组合的服务语义匹配存在着一些不足, 诸如缺少相应的辅助机制提高服务匹配的效率, 使得服务匹配和组合难以快速实现。文中基于服务索引机制和语义匹配度, 给出了相应的语义服务匹配算法, 该方法可提高服务匹配速度和精确度, 为动态服务组合奠定了良好的基础。服务索引的引入会增加建立服务索引的计算时间的耗费及需要一定

速查找前趋服务的类别 cat_{pre} 和后继服务的类别 cat_{sub} 。面向组合的前趋服务与后继服务的综合语义匹配值为:

的存储空间来存储索引, 但是从长远目标来看, 这种前期的时间耗费和空间投入是有意义的, 可以为服务匹配和服务组合节省大量时间。

本文研究工作得到安徽师范大学博士引进启动资金的资助, 在此表示感谢。

参 考 文 献

- [1] 刘书雷, 刘云翔, 张帆, 等. 一种服务聚合中 QoS 全局最优服务动态选择算法[J]. 软件学报, 2007, 18(3): 646-656.
LIU Shu-lei, LIU Yun-xiang, ZHANG Fan, et al. A dynamic Web services selection algorithm with QoS global optimal in Web services composition[J]. Journal of Software, 2007, 18(3): 646-656.
- [2] PAOLUCCI M, KAWAMYRA T, PAYNE T R, et al. Semantic matching of web services capabilities[C]// Proceedings of the First International Semantic Web Conference on the Semantic Web (ISWC). Seattle Washington, USA: Springer-Verlag, 2002: 333-347.
- [3] SRINIVASAN N, PAOLUCCI M, SYCARA K. Semantic Web service discovery in the OWL-S IDE[C]// Proceedings of the 39th Annual Hawaii International Conference on System Sciences. Kauai Hawaii, USA: IEEE Computer Society, 2006, 6: 1-10.
- [4] NAMGOONG H, CHUNG M, KIM K, et al. Effective Semantic web services discovery using usability[C]// The 8th International Conference On Advanced Communication Technology. Phoenix Park, Korea: IEEE Computer Society, 2006: 2199-2203.
- [5] MANIKRAO U S, PRABHAKAR T V. Dynamic selection of web services with recommendation system[C]// Proceedings of the International Conference on Next Generation Web Services Practices. Seoul, Korea: IEEE Computer Society, 2005: 22-26.
- [6] LI L, HORROCKS I. A software framework for matchmaking based on semantic web technology[J]. International Journal of Electronic Commerce, 2004, 8(4): 39-60.
- [7] XU Bin, LI Tao, GU Zhi-feng, et al. SWSDS: quick web service discovery and composition in SEWSIP[C]// Proceedings of the 8th IEEE International Conference on E-Commerce Technology and the 3rd IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services (CEC/EEE '06). Palo Alto, USA: IEEE Computer Society, 2006: 71-73.
- [8] BUDAK ARPINAR I, ALEMAN-MEZA B, ZHANG R, et al. Ontology-driven web services composition platform[C]// Proceedings of the IEEE International Conference on

- E-Commerce Technology. San Diego, USA: IEEE Computer Society, 2004: 146-152.
- [9] AIELLO M, PLATZER C, ROSEENBERG F, et al. Web service indexing for efficient retrieval and composition[C]//Proceedings of the 8th IEEE International Conference on E-Commerce Technology and the 3rd IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services (CEC/EEE '06). Palo Alto, USA: IEEE Computer Society, 2006: 63-65.
- [10] BLAKE M B, CHEUNG W, JAEGER M C, et al. WSC-06: the web service challenge[C]//Proceedings of the 8th IEEE International Conference on E-Commerce Technology and the 3rd IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services (CEC/EEE '06). Palo Alto, USA: IEEE Computer Society, 2006: 62-63.
- [12] 朱海平. 基于概念图的语义搜索[D]. 上海: 上海交通大学, 2006.
- ZHU Hai-ping. Semantic search by matching conceptual graphs[D]. Shanghai: Shanghai Jiaotong University, 2006.

编辑 熊思亮

(上接第817页)

参 考 文 献

- [1] HAYKIN S. Cognitive radio: Brain-empowered wireless communications[J]. IEEE Journal on Selected Areas in Communications, 2005, 23(2): 201-220.
- [2] HEREDIA-URETA H, CRUZ-PEREZ F A, ORTIGOZA-GUERRERO L. Capacity optimization in multiservice mobile wireless networks with multiple fractional channel reservation[J]. Vehicular Technology, IEEE Transactions on, 2003, 52(6): 1519-1539.
- [3] ANDREWS M, KUMARAN K, et al. CDMA data QoS scheduling on the forward link with variable channel conditions[R]. Bell Laboratories Technical Report, April 2000.
- [4] SHAKKOTTAI S, STOLYAR A L. Scheduling for multiple flows sharing a time-varying channel- the exponential rule[R]. Bell Laboratories Technical Report, Dec 2000.
- [5] VISWANATH P, TSE D N C, LARROIA R. Opportunistic beamforming using dumb antennas[J]. Information Theory, IEEE Transactions on, 2002, 48(6): 1277-1294.
- [6] ELI S, GERALD C. WRAN channel modeling. [2005- 07-12]. http://www.ieee802.org/22/Meeting_documents/2005_July/22-05-0055-01-0000_WRAN_Channel_Modeling-ES-GC.doc.
- [7] CAI X, GIANNAKIS G B. A two dimensional channel simulation model for shadow fading processes[J]. IEEE Transactions on Vehicular Technology, 2003, 52(6): 1558-1567.
- [8] BAUGH C R, HUANG J. Traffic model for 802.16 TG3 MAC/PHY simulations[DB/OL]. [2001-03-02]. http://www.ieee802.org/16/tg3/contrib/802163c-01_3_0r1.pdf.
- [9] WEISS T A, JONDRAL F K. Spectrum pooling an innovative strategy for the enhancement of spectrum efficiency[J]. Communications Magazine, IEEE, 2004, 42(3): S8-14.
- [10] BENKO J, CHEONG Y C, CORDEIRO C. A PHY/ MAC proposal for IEEE 802.22 WRAN systems part 2: The Cognitive MAC[DB/OL]. [2006-02-23]. http://www.ieee802.org/22/Meeting_documents/2006_Mar/22-06-0003-03-0000_ETRI-FT-I2R-Motorola-Philips-Samsung-Thomson_MAC_Spec.doc.

编辑 漆 蓉

(上接第903页)

- [3] SZU H H, TELFER B A, KADAMBE S L. Neural network adaptive wavelets for signal representation and classification[J]. Optical Engineering, 1992, 31(9): 1907-1916.
- [4] ZHANG Q, BENVENISE A. Wavelet network[J]. IEEE Trans on Neural Network, 1992, 3(6): 889-898.
- [5] SZU H H, TELFER B, KADAMBE B. Neural network adaptive wavelets for signal representation and classification[J]. Optical Engineering, 1992, 31(9): 1907-1916.
- [6] ZHANG J, WALTER G, MIAO Y, et al. Wavelet neural networks for function learning[J]. IEEE Transactions on Signal Processing, 1995, 43(6): 1485-1497.
- [7] CAO L, HONG Y. Predicting chaotic time series with wavelet networks[J]. Physic D, 1995, 21(5): 85-97.
- [8] 马 光. 共轭梯度法训练小波神经网络[J]. 信息技术, 2004, 28(9): 54-60.
- MA Guang. Training arithmetic of wavelet neural network conjugate grads algorithm[J]. Information Technology, 2004, 28(9): 54-60.
- [9] ZHANG J, WALTER G, MIAO Y, et al. Wavelet neural network for function learning[J]. IEEE Transactions on Signal Processing, 1995, 43(6): 1485-1497.
- [10] Zhang Qing-hua. Using wavelet network in nonparametric estimation[J]. IEEE Trans Neural Network, 1997, 8(2): 227-236.
- [11] COHEN M, A, GROSSBERG S. Absolute stability of global pattern formation and parallel memory storage by competitive neural networks[J]. IEEE Trans Syst Man Cybern, 1983, SMC-13: 815-826.

编辑 熊思亮