

语义 Web 服务发现研究现状与发展*

王 珏¹, 向朝参¹, 王 萌¹, 田 畅¹, 赵文栋¹, 代登坡²

(1. 解放军理工大学 通信工程学院, 南京 210007; 2. 中国人民解放军 78083 部队, 成都 610011)

摘 要: 从不同方面对语义 Web 服务发现相关技术进行综述。阐述了语义 Web 服务的基本概念和特点, 总结和分析了语义服务发现的基础——语义服务描述语言, 重点从单服务匹配和服务组合匹配两个方面对语义 Web 服务发现的关键问题——语义 Web 服务匹配近年来的研究, 进行了全面的总结和讨论, 并指出了语义 Web 服务发现研究领域的挑战和未来的研究方向。

关键词: 面向服务架构; 语义 Web 服务; 服务发现; 服务组合; 服务匹配

中图分类号: TP393

文献标志码: A

文章编号: 1001-3695(2013)01-0007-06

doi: 10.3969/j.issn.1001-3695.2013.01.002

Survey on semantic Web services discovery

WANG Jue¹, XIANG Chao-can¹, WANG Meng¹, TIAN Chang¹, ZHAO Wen-dong¹, DAI Deng-po²

(1. Institute of Communication Engineering, PLA University of Science & Technology, Nanjing 210007, China; 2. 78083 PLA Troops, Chengdu 610011, China)

Abstract: This paper presented the state-of-the-art of semantic Web service from various aspects. First of all, this paper explained the basic concepts and character of semantic Web service. And then it summarized and analyzed several kinds of service description languages which were the base of service discovery. Moreover, it comprehensively summarized and discussed service matching as the key problem of service discovery from two aspects—single service matching and service composition matching. In the end, this paper proposed the challenge and the future work of the study on semantic Web service discovery.

Key words: SOA (service-oriented architecture); semantic Web service; service discovery; service composition; service matching

0 引言

信息技术的发展带动了各类应用系统的不断涌现, 在为用户带来更多便利的同时, 也给 IT 系统的设计和开发带来了新的挑战。如何将采用不同平台和不同技术开发的独立系统进行互连与共享, 最终实现系统间的灵活配置、松散耦合、即插即用以及服务重用等功能, 从而便于用户获取服务。面向服务架构(SOA)由于实现了系统间的松散耦合, 能很好地解决异构系统互连和共享这一难题, 得到了高度的重视与广泛的应用^[1~4]。2008年Gartner公司调查发现, 在当前大型的IT系统中, 已经采用SOA的占53%, 另有25%计划在一年内采用^[5]。因此, SOA已成为各种应用系统互连和集成的一种趋势^[2]。针对不同的应用, SOA有多种实现方式, 而近年来随着因特网的不断发展, Web服务逐渐成为SOA实现技术的研究热点^[6, 7]。

Web服务是一个广义的概念, 不仅在因特网中有广泛应用, 很多其他IT系统也采用了Web服务技术。例如文献[2]提出利用Web服务技术将分布在不同地区的各种硬件设备互连, 便于设备提供的功能进行共享和重用; 文献[8, 9]提出在

无线、移动网络中利用Web服务技术对服务进行重用和组合; 文献[10]提出将Web服务技术应用于非Web组件中, 并提出了原型系统。随着Web应用的普及以及规模的扩大, Web服务数量与日俱增, 同时用户的服务需求也是千变万化, 如何根据用户的需求快速、高效地从海量的服务中发现用户所需的服务是Web服务研究的关键点和热点, 直接决定着用户获取服务的效率和性能。

Web服务发现依据服务描述方式划分, 大体经历了基于关键字、基于语法和基于语义三个阶段^[11]。由于基于关键字和语法的描述能力有限, 导致服务发现的效率低, 不便于服务的自动组合。为了提高服务描述能力和发现效率, 学术界提出了语义服务描述。语义服务描述的出现, 在增强服务描述能力的同时也给服务发现带来了新的挑战^[5, 11, 12]。针对以上存在的问题, 近年来学术界从不同侧面进行了研究, 如基于语义的服务描述语言的提出^[5, 13, 14]、基于语义的单服务匹配^[10, 15~17]与基于语义的服务组合匹配^[18~21]等。文献[22]对2006年以前发布/订阅系统的各种关键技术进行了详细的综述, 但其主要侧重于发布/订阅系统相关技术的分析; 文献[15~23]分别对2004年以前的Web服务核心技术和2008以前的服务发现技术进行了综述, 但近年来Web服务研究领域

收稿日期: 2012-05-11; 修回日期: 2012-07-16 基金项目: 国家“973”计划资助项目(2009CB3020402); 江苏省自然科学基金资助项目(BK201010103)

作者简介: 王珏(1987-), 男, 黑龙江哈尔滨人, 助理工程师, 主要研究方向为Web服务匹配(415638921@qq.com); 向朝参(1987-), 男, 博士研究生, 主要研究方向为SOA架构、无线传感网; 王萌(1983-), 男, 助教, 主要研究方向为Web服务发现; 田畅(1963-), 男, 教授, 博导, 主要研究方向为网络信息系统、无线分组网; 赵文栋(1972-), 男, 副教授, 主要研究方向为SOA体系架构、计算机网络; 代登坡(1986-), 男, 助理工程师, 主要研究方向为Web服务发现。

又出现了很多新的进展,如基于语义的服务描述语言的提出^[12]以及基于语义的服务组合匹配^[24]等。本文主要对近几年基于语义的Web服务发现研究进行梳理,总结和分析了目前研究的热点和主要研究方向,为语义Web服务的进一步发展提供技术支持和帮助。

1 语义Web服务的基本概念

1.1 Web服务的定义和组成

Web服务在学术界和工业界都还没有一个统一的定义,不同组织机构和研究人员的理解都有所不同,其中一种常见的定义是:Web服务是一种可以被描述和发现,可通过可扩展标记语言(extensible markup language, XML)、统一标记语言和标准协议在各种网络中获得的软件应用系^[5]。简而言之,Web服务是一种采用SOA的软件应用系统。

如图1所示,Web服务架构由服务提供者(provider)、服务请求者(client)和服务注册代理(register)三部分组成,服务发现过程主要包含服务描述、服务发布和服务匹配三步。

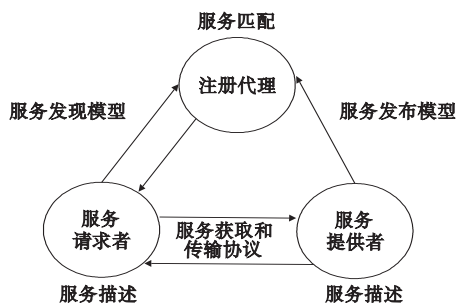


图1 Web服务基本架构

1.2 Web服务的特点和优点

Web服务由于采用SOA,具有松散耦合、粗粒度以及位置和传输协议透明等特点,并具有以下优点:

- a) 可将不同系统(采用不同平台、语言、技术以及开发公司开发的)互连,面向所有用户共享。
- b) 服务可重用,可将不同系统集成。
- c) 服务的更新和实现对用户透明。
- d) 即插即用,系统的传输协议和端口地址对用户透明。

语义Web服务的出现和发展,使Web服务发现的研究呈现出新的活力,下面将从服务描述、服务匹配两个方面分别对最近几年语义Web服务发现研究的基础和关键问题进行分析 and 总结。

2 基于语义的服务描述

服务描述是指对服务请求者的需求和服务提供者发布的服务进行统一的描述,以便于服务发现。它是语义Web服务发现的基础,影响服务发现算法的设计,其发展过程大体可以分为基于关键字、基于语法和基于语义描述三个阶段。

2.1 基于关键字和语法的描述

基于关键字的服务描述是最早提出的,同时也是最简单的一种服务描述方法。它通过几个关键字来描述服务,描述方法及相应的匹配机制都很简单,但其服务描述能力不强、描述不准确、灵活性差,并且基于该描述的服务匹配精度不高。基于语法的描述是在基于关键字服务描述的基础上发展起来的,主要以服务描述语言(Web services description language, WSDL)为代表。WSDL是由万维网联盟(World Wide Web Con-

sortium, W3C)推荐,采用XML格式表示的一种基于语法的描述语言,其最新的版本是W3C在2007年推荐的WSDL 2.0。WSDL 2.0服务描述由抽象描述(abstract section)和具体描述(concrete section)两部分组成,前者主要描述服务的功能,如输入、输出以及先决条件等;后者主要将抽象描述映射到具体的服务上,由两部分组成:描述如何获取服务,如采用什么协议等,以及描述去何地获取服务,如获取地址等^[5]。

基于关键字和语法的描述的研究相对成熟,但其描述能力有限、灵活性和扩展性较差,例如,由于这两种方法不能描述南京和白下区两者之间的关系,不能很好地解决其匹配问题,人们将更多的目光转向服务描述能力更强的语义描述。

2.2 基于语义的服务描述

基于语义的服务描述主要是利用本体表述领域内的知识,使服务描述带有语义;同时,采用描述逻辑进行推理,使服务的语义能被机器所“理解”,便于服务的自动发现和组合^[5]。针对语义服务描述,相关组织在2004年—2007年之间相继提出和推荐了几种基于语义的服务描述语言^[25],主要有SAWSDL(semantic annotations for WSDL and XML schema)、OWL-S(ontology Web language for services)、SWSF(semantic Web services framework)、WSMO(Web service modeling ontology)等。下面介绍和比较两种常见的语义服务描述语言SAWSDL和OWL-S。

SAWSDL是W3C在2007年最新推出的基于语义的服务描述语言,它是对WSDL的语义扩展,其前身是WSDL-S。SAWSDL采用本体对WSDL中各服务描述元素进行语义标记,使之与本体进行关联和对等,从而通过本体使服务描述带有语义信息,便于服务推理、匹配和组合。

OWL-S是W3C在2004年推荐的基于语义的服务描述语言,它是以OWL为基础的描述Web服务语义的一种描述语言。OWL-S由三部分组成:服务轮廓(service profile)、处理模型(process model)和服务背景(service grounding)。服务轮廓描述服务做什么,如服务的名字、提供者、输入、输出以及先决条件等;处理模型描述服务如何做,包括如何调用、执行以及监管服务等;服务背景描述如何获取服务,如服务的通信协议、消息格式以及服务的端口地址等。

SAWSDL与OWL-S之间既有区别又有联系。由于SAWSDL是以WSDL为基础,利用本体对语义进行标记和关联,它对基于语法的描述WSDL能很好地进行兼容和继承;而OWL-S由于是在OWL的基础上发展起来的,对WSDL的支持和兼容性差。但是它们之间也存在一定的联系,OWL-S的主要组成部分服务轮廓、处理模型和服务背景分别与SAWSDL中服务抽象描述和具体描述部分对应和等同。

现有的几种基于语义的服务描述语言已具有较强的语义描述能力,基本上能满足当前服务的语义描述要求。近年来,相比于语义服务描述语言的研究,它的实际应用研究比较少,工程中实际服务的语义描述还存在很多难题需要解决,例如,如何构建复杂、庞大的领域知识本体;如何构建服务语义模型对实际服务进行描述等。因此,近年来对语义服务描述语言的研究相对较少,而具有一定应用背景的服务语义模型及应用研究开始引起了人们的关注^[13, 20, 25-27]。文献[13]针对战场目标探测和跟踪这一特定应用领域,提出了丰富航迹语义模型(rich semantic model of track)来描述战场环境中的航迹语义信息,对航迹的各个要素进行描述,从而便于航迹的自动识别、跟踪和信息融合,为战场信息的分发、处理及战场信息系统的互

操作研究提供了新的思路。

3 基于语义的服务匹配

对于 Web 服务而言,它的关键问题是服务请求者如何能够准确地获取所需服务,即服务发现^[11],它是 Web 服务系统架构的重要组成部分,也是实现服务调用的前提。Web 服务发现的核心和关键是 Web 服务匹配。按照服务匹配个数可将语义服务匹配分为单服务匹配与服务组合匹配。单服务匹配主要利用本体概念间语义关系来推理和计算两服务之间的语义相似度,根据服务之间的语义相似度进行匹配。它按照服务语义相似度的计算方法可分为基于概念本体逻辑关系匹配和基于概念本体相似度匹配。服务组合利用服务的可重用性,将多个服务进行组合以满足用户复杂的服务需求,为用户提供更加丰富和强大的服务。它以单服务匹配为基础,按照服务组合方式可分为 top-down、bottom-up 以及混合式服务组合匹配,其所有的语义服务匹配关系如图 2 所示。下面将分别从单服务匹配与服务组合匹配两个方面对近几年语义服务匹配相关研究进行分析和总结。

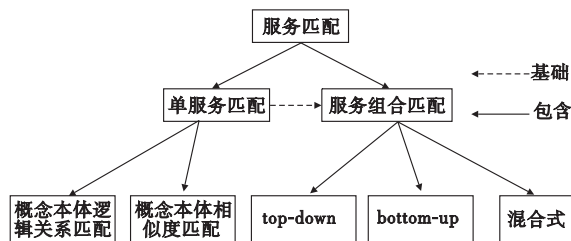


图2 语义Web服务匹配关系

3.1 基于语义的单服务匹配算法

基于语义的单服务匹配的一般思想是:首先分别对要匹配的两服务描述中各个对应属性值的本体概念计算匹配程度,然后求和计算总的服务匹配程度^[10,15,16]。如式(1)所示。其中: S_r 和 S_p 分别表示要匹配的服务请求和发布的服务; c_{ri} 和 c_{pi} 分别表示要匹配的两服务描述中第 i 个属性的本体概念; n 表示服务描述中属性概念的总个数。

$$\text{match}(S_r, S_p) = \sum_{i=1}^n \text{match}(c_{ri}, c_{pi}) / n \quad (1)$$

由式(1)可知,单服务匹配的关键是计算两概念的语义匹配程度 $\text{match}(c_{ri}, c_{pi})$,可分为基于概念的本体逻辑关系^[10,28]和基于概念的本体相似度^[15~17]两种。

基于概念本体逻辑关系计算语义匹配程度的思想是根据两概念本体的逻辑关系进行逻辑关系推理匹配,如概念本体逻辑关系继承(inheritance)、被包含(hasPart)和等价(equivalent)^[10]等。文献[28]通过对服务的输入、输出、先决条件和执行结果(input output preconception effect, IOPE)四个属性的语义描述来进行服务描述,利用描述逻辑概念包含和一致性检测原理进行推理,根据各属性概念本体间的关系对服务进行匹配,不同关系的概念间的匹配结果等级不同,如精确匹配(exact match)、绝对匹配(perfect match)和一般匹配(common match)等。这种方法的缺点是匹配精度不高,返回的匹配结果分成不同的等级,粒度较粗,不便于准确地比较各个服务间的匹配程度。文献[10]为了解决这个问题,将匹配结果进行量化,给每个等级的匹配结果设置一定的分数,然后计算整个服务的匹配分数,这样更加准确地度量服务的匹配程度,提高了服务匹配的精度和性能。

基于概念本体相似度计算语义匹配程度的思想是:根据概念本体逻辑关系树提取影响两概念本体之间相似度的因素,利用这些因素定义相似度函数。不同文献采用的因素和构造函数的形式各有不同,没有一个统一的定义,如文献[15]采用两概念本体的最短距离和到最近同一根节点的距离两个因素构造了一个以自然数为底的指数函数;文献[16]增加了到树根节点的距离等因素构造了一个以小于1的常数 t 为底的指数函数。虽然它们考虑的因素和采用的函数形式各异,但它们的本质和原理是相同的。在众多的相似度函数中,什么样的相似度函数度量概念本体之间的相似程度最准确?还没有一个较成熟的方法去比较和证明各种不同的相似度函数描述概念本体之间的匹配程度的优劣性。

基于概念本体逻辑关系和基于概念本体相似度这两种计算方法各有优缺点和难点,比较结果如表1所示。

表1 两种计算概念语义匹配程度方法的比较

方法	优点	缺点	难点
基于概念本体逻辑关系	方法简单、直观	逻辑推理较复杂,匹配资源耗费大	如何改进逻辑推理,对服务的匹配精度和匹配复杂度进行折中
基于概念本体相似度	计算和匹配耗费低	相似度函数较难定义,相似度函数的好坏较难评价	最优相似度函数如何定义,如何证明和比较各个相似度函数的优劣

当前,在研究如何进行基于语义的单服务匹配的同时,也对如何提高服务匹配的速度和降低匹配的复杂度进行了研究^[17,25,27]。文献[17]在文献[16]基于概念本体相似度的服务匹配基础上,通过预先建立和维护概念近似度数据表和近似概念倒位索引表,分别为每个概念本体保存相近的概念本体及所对应的服务,采用近似概念搜索算法使服务匹配在目标概念附近搜索,提高了服务匹配的效率和速度。文献[25,27]在基于概念本体逻辑关系的语义服务匹配基础上,研究如何在服务匹配精度和匹配复杂度之间进行折中。文献[27]提出了一种新的逻辑推理匹配机制,当系统所剩下的资源不能继续进行服务匹配时则停止匹配,从而在满足系统资源条件下,尽可能地服务匹配,提高服务匹配精度。然而,由于各个属性的匹配对整个服务匹配的重要程度不一样,而本方法没有考虑各个属性匹配权重这一问题,导致很多决定服务匹配的关键属性由于资源有限没能进行匹配,降低了服务匹配的准确度。文献[25]在文献[27]的基础上,通过采取减少一些不必要或不重要的推理、在推理过程中根据当前系统的资源来决定推理是否继续以及根据各个属性的匹配权重进行排序,权重高的属性优先匹配等一系列机制改进 Tableaux 推理算法,使在资源有限的情况下尽可能提高服务匹配的精度,较好地解决了文献[27]中未考虑各属性匹配权重的问题;但由于采用的是部分匹配方法,匹配结果的置信度随着匹配程度的不同而不同,因此其缺点是未给出匹配结果的置信度。

3.2 基于语义的服务组合匹配

由于在 SOA 中服务功能划分较细,而用户所需求服务的粒度较粗,单个服务很难满足用户的全部需求,因此为了更加充分地利用已发布的服务,更好地满足用户的服务需求,利用服务的可重用性对细粒度的服务进行组合^[29]。服务组合思想很早就被提出,伴随着 SOA 概念的产生而产生,如 SOA 在 1996 年被提出的初衷就是为了便于服务间的组合和集成^[30],然而服务组合匹配研究最近几年才被广泛关注^[15,23]。虽然学

术界和工程界近年来对服务组合匹配展开了大量的研究,但服务组合匹配还没有一个明确的、统一的分类方式。文献[31]提出将服务组合分为 top-down、bottom-up 和混合式三种;文献[32]将服务组合分为静态组合、半自动组合和自动组合。因为第一种分类方法更加常见,所以下面将分别根据第一种分类的三种方式对服务组合匹配当前研究进行分析和总结。

1) Top-down 服务组合匹配

Top-down 服务组合匹配是指根据用户的服务需求,按照服务的执行流程和各功能的逻辑划分,先构建一个由多个子服务功能组成的功能组合流程,如图3所示,然后针对每个子服务功能的需求从发布的服务中选择具体的服务,最后按照功能组合流程结构将子服务进行组合^[31]。这种方法的优点是简单、易实现,在构建功能组合流程成功后,组合服务匹配就转换为子服务匹配,即单服务匹配;但其缺点是在创建功能组合流程时,只考虑了用户服务需求的功能组合流程和逻辑划分,没有考虑已发布的具体服务,因此不能很好地利用已发布的服务,从而导致服务匹配的成功率不高,不能很好地满足用户的服务请求。

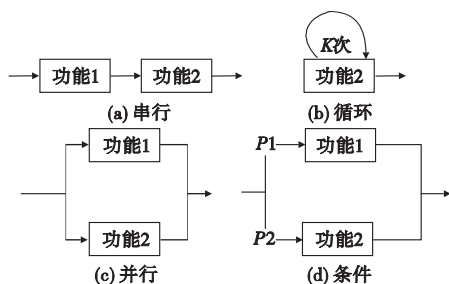


图3 Top-down功能组合流程的四种结构

Top-down 服务组合匹配方式简单,易于工程实现,研究和应用较多,其组合流程的结构复杂多样,通常包括串行、并行、循环和条件四种最基本的结构^[33],如图3所示。

2) Bottom-up 服务组合匹配

Bottom-up 服务组合匹配是指在无服务功能组合流程图的情况下,根据用户的服务需求,从发布的服务中选择多个服务进行自动组合,使组合的服务满足用户的需求^[31]。这种组合匹配方法的优点是充分地利用了已发布的服务;缺点是在服务组合时没有考虑到用户需求服务的逻辑功能组合流程,产生的最优服务组合通常难以满足用户服务的要求,并且由于组合结构多样、原子服务数量大,使组合匹配实现的难度大^[34]。

当前, bottom-up 服务组合匹配一般只能进行简单的服务组合,组合结构比较简单,如串行、并行等。文献[10]根据各个服务之间输入和输出的语义匹配程度,选择出与用户所请求服务最匹配的多个服务作为组合候选子服务,然后用户根据服务功能需求将各候选子服务进行手动组合连接。这种服务组合匹配方法类似于现在互联网上广泛使用的 Yahoo Pipes,但它避免了 Yahoo Pipes 输入、输出类型和流程结构有限,以及只能联网使用等缺陷,同时,利用语义使服务组合匹配更加灵活,功能更加强大。文献[30]假设服务组合匹配的各候选子服务已经确定,在无服务功能组合流程图的情况下,将各候选子服务组合以满足用户服务的需求。首先,它通过计算各服务之间输入和输出的关联程度建立服务之间的依赖关系矩阵,然后分别按照服务被依赖的服务数和依赖的服务数进行排序,最后基于这两个排序结果进行服务组合。这种方法是一种启发式算法,它认为某个服务如果被依赖的服务数越多,其服务在组合中排列越靠前;如果它依赖其他服务数越多,则在组合中排

列越靠后。文献[5,10]都采用了 bottom-up 服务组合匹配方法,根据各个服务之间输入和输出的关系进行服务组合匹配,但存在的主要问题是服务组合匹配的能力比较弱,文献[10]只能选择出部分候选子服务,需要用户自己去补充、选择和连接这些候选子服务;文献[5]则回避了子服务选择问题,并且其组合结构仅限于简单的串行和并行。这两种组合方法因为只根据服务之间输入和输出的简单关系进行服务组合匹配,所以匹配的服务可能不满足用户的需求,这也是 bottom-up 服务组合匹配很难满足用户需求的重要原因之一。如何更加准确地描述服务之间的关系是 bottom-up 服务组合匹配的难点和关键点。

为了提高 bottom-up 服务组合匹配的速度和准确度,文献[5,20]提出在服务组合匹配前预先建立服务之间的关系。文献[20]利用服务形式化的概念分析,将服务依据它们之间的关系组织成格型结构(lattice structure),后代服务可以继承祖辈服务的所有属性,可以代替祖辈实现所需的服务功能,因此它便于服务组合匹配,尤其是当组合服务中的某个服务出现故障或者不可获得时,可快速发现可替代的服务。文献[35]借鉴网页中超链接思想,将服务之间的输入和输出关系显示地描述在服务中,但它存在的问题是这种显示关系是静态的,不支持动态的调整和自动描述,灵活性差。

Bottom-up 组合匹配方式虽然在服务组合匹配中能力比较弱,但它适合于服务挖掘(service mining)。当用户对所需组合服务的目标、结构以及现存的服务等信息不清楚时,同时随着现有服务规模日益扩大,如何充分地利用已有的服务,主动地挖掘出新的服务供用户选择,这就需要服务挖掘^[19]。服务挖掘是指在无用户组合服务的目标等先验知识的情况下,根据已发布的服务进行主动组合,挖掘出新的、有用的服务。它的难点和关键问题是:a)如果服务任意组合,则服务的个数将呈指数上升,产生组合爆炸,如何避免服务组合爆炸,降低服务组合的资源消耗;b)由于服务任意组合,因此,如何对组合的服务进行评估,将有价值和感兴趣的服务保留,将与原服务功能相近的或无价值的服务丢弃。文献[19]针对这两个问题,通过设定用户感兴趣的领域,从而在该领域的本体库中进行服务组合,使服务组合的规模随着服务个数的增长呈线性增长,抑制了服务的组合爆炸,并且通过客观评价(objective evaluation)、人机交互式推测表示(interactive hypothesis formulation)、仿真运行(simulation)以及主观评价(subjective evaluation)四步对组合服务的有用性(useful)和感兴趣的程度(interesting)进行有效评估。但该方法的缺点是服务评估的算法较复杂,评估时间比较长。

3) 混合式服务组合匹配

混合式服务组合匹配是指将 top-down 和 bottom-up 两种组合匹配方式的优点相结合,先根据用户的服务需求和服务的逻辑功能构建一个由几个较粗粒度的子功能组成的服务功能组合流程,然后对各个子功能从已发布的服务中进行 bottom-up 服务组合匹配。文献[28]采用混合式服务组合匹配,其提出的背景是当前服务越来越多,较难发现和充分利用所需要的服务,因此构建了一个类似于社会关系网的服务网(SN)来描述各服务之间的关系,便于服务发现。它具有三大功能,一是作为集中式注册中心;二是建立和维护各个服务之间的关系信息,便于服务发现和组合;三是对已存在的服务采用文献[19]的方法进行服务挖掘,以扩充发布的服务集,挖掘更多新的服务。这样既利用了 bottom-up 服务组合丰富了发布的服务集,

便于用户选择,又可以根据用户的服务需求进行 top-down 服务组合匹配。

混合式服务组合匹配实际工程应用价值比较大,它既可以用户的需求进行服务组合匹配,又尽可能扩充和挖掘新的服务,但仍还有很多理论和工程实践问题需要进一步研究和解决。例如,如何将 top-down 服务组合和 bottom-up 服务组合很好地结合,以提高服务组合匹配的性能;如何将服务按照逻辑功能划分成粗粒度的子服务使混合式服务组合性能最佳等。

由本文前面的分析和总结可知,语义 Web 服务组合的方式各有优缺点(表2),其适用范围各不相同,应根据不同的应用场景和需求选择合适的组合算法。

表2 两种计算概念语义匹配程度方法的比较

服务组合方式	优点	缺点	难点
top-down	匹配简单、易于工程实现	未考虑已发布的各个具体服务,灵活性差	如何使构建的工作流考虑已发布的服务
bottom-up	能够充分地重用已发布的服务,适合于服务挖掘	组合复杂高,并且组合结果较难满足用户的实际需求	如何降低服务组合的时间复杂度和提高组合结果的精度
混合式	既能满足用户的需求又能较好地重用已发布的服务	方法设计较复杂	如何使 top-down 和 bottom-up 更好地结合以提高服务组合的性能

服务组合匹配的研究不仅具有较大的理论意义,而且具有广阔的应用前景,目前很多基于 SOA 的工程系统中应用了服务组合,便于用户获取更多、更丰富的服务,同时,部分对实际工程中如何进行服务组合进行研究。例如,文献[36]提出如何将分布在不同地方、用网络连接的各种硬件设备提供的功能进行服务组合;文献[37]则针对在工程实践中,服务组合匹配已经完成,解决如何对接口不同的服务进行组合这一问题。

4 需要研究的问题

虽然近些年对基于语义的 Web 服务发现的研究取得了丰硕的成果,但距离语义 Web 服务的目标——计算机能根据服务的语义自动地、智能地进行服务发现、组合和处理,存在一定的差距,还有很多值得进一步研究的问题。

1) 基于语义的服务描述

虽然基于语义的服务描述语言已被几大组织提出和推荐,并且具有较强的描述能力,但是如何利用这些描述语言对某一具体工程领域的服务进行语义描述还有待进一步研究。其难点在于:a)既要考虑服务描述的能力,又要考虑服务描述的复杂度,更要考虑实际工程服务描述的要求和特点;b)如何检验和评估服务描述的好坏。

2) 基于语义的服务匹配

在单服务匹配中,虽然已提出了很多方法^[10,16,21,24],但存在匹配复杂度高、相似度函数较难确定等问题。因此,如何对服务本体逻辑推理进行改进,在服务的匹配精度和匹配复杂度之间进行折中;如何定义一个更准确描述语义本体之间相似度的函数;如何评价和比较各种不同相似度函数的性能;如何对单服务匹配算法进行改进,以提高服务匹配的速度。虽然目前大量的语义 Web 服务匹配算法被提出,但还存在很多问题值得进一步研究。

a) 匹配时间复杂度以及服务的查全率和查准率是衡量服务匹配算法的三个重要指标,但是这三个指标是互相制约的,

很难同时达到最优。因此,如何根据实际应用的需求对这三个指标进行很好的折中是当前研究的一个难点。

b) 由于影响概念间相似度计算的因素有很多,如概念在本体中的深度、概念间的密度和距离等。因此,很难提出一个相似度计算函数适合所有不同的应用场景和需求;同时,如何评价一个相似度计算函数的优劣也是当前研究的一个难点。

c) 语义 Web 服务匹配算法不仅需要考虑服务之间的参数匹配,还要考虑服务执行时间以及服务费用等非功能属性匹配。因此,合理地结合功能性匹配算法和非功能性匹配算法以提高服务匹配算法的精度和效率是当前研究的一个热点。

在服务组合匹配中,bottom-up 组合匹配是一种服务的自动组合方式,更能满足人们的需求,但目前 bottom-up 组合匹配方法^[5,10]的服务组合能力比较弱,并且只能支持简单的串行或者并行组合。因此,如何支持更加灵活、多样的服务组合结构,如何更加准确地描述服务之间的关系以提高服务匹配的准确度,如何提高 bottom-up 组合匹配的性能,使组合的服务更加符合用户的需求等问题还需要研究。Bottom-up 组合匹配适合于服务挖掘,虽然文献[19]已解决了服务挖掘的部分问题,但还存在评估复杂度高缺点,如何更加有效进行服务挖掘和评估还有待研究;同时,将服务挖掘与 top-down 服务组合匹配相结合可以提高服务组合匹配的性能,具有较大的研究意义和前景,但是如何很好地结合有待进一步的研究。虽然服务组合的理论研究比较多,方法也相对成熟,但服务组合在实际工程中的实现还存在很多问题,还值得研究,如语义 Web 服务组合中的复合语义、事务处理、复杂服务质量模型以及服务组合当中的安全问题;如怎样实现服务的自动组合并满足用户的服务需求、怎样自动发现服务等,如文献[38]为了提高情报监视侦查系统(intelligence surveillance and reconnaissance systems,ISR)快速适应多变任务和环境的能力,利用插件技术在工程中实现了服务组合,并提出了原型系统。虽然它在工程中实现了服务组合的原型系统,但需要人工选择和组合服务,仅实现了简单的、静态的服务组合;文献[4]基于 SOA 提出了资源中介服务(resource brokering service,RBS),为用户动态地、按需地分配各种资源,提高了资源利用的灵活性和实时性,同时采用了服务组合技术,便于多种资源的组合利用。但它只提出了相关的概念和设计架构,还存在一些实际工程问题正在研究中,如服务自动组合和执行算法、语义服务描述等。因此,服务组合的工程实现还有待进一步的研究。

5 结束语

近年来语义 Web 服务成为 SOA 研究的一个最主要的领域^[7],其应用规模和范围日益扩大,不仅局限于互连网中传统的 Web 服务,还可以将不同硬件设备互连,便于其功能共享和重用,如 RBS^[4]、ISR^[38]、信息栅格(information grid)等。在语义 Web 服务的实现和研究中,服务发现是关键和难点。工业界和学术界都从不同方面对此进行研究,并从理论和应用两个方面提出了各种原型系统和实际应用产品。本文从服务描述、服务匹配两个方面,分析了当前语义 Web 服务发现的研究现状,重点讨论了当前的研究热点——语义服务匹配,既讨论了各种匹配方法提出的背景,又分析了它们之间的关系,还比较了各自的优缺点。最后提出了这一研究领域中有待解决的关键问题和未来的研究方向。

参考文献:

[1] Bin IBRAHIM M N, Bin HASSAN M F. A survey on different interoper-

- erability frameworks of SOA systems towards seamless interoperability [C]//Proc of International Symposium in Information Technology. Washington DC: IEEE Computer Society, 2010: 1119-1123.
- [2] GUINARD D, TRIFA V, KARNOUSKOS S *et al.* Interacting with the SOA-based Internet of things: discovery, query, selection, and on-demand provisioning of Web services [J]. *IEEE Trans on Services Computing* 2010, 3(3): 223-235.
- [3] ZHANG Xiu-guo, LIU Hong-bo. Formal verification for CCML based Web service composition [J]. *Information Technology Journal*, 2011, 10(9): 1692-1700.
- [4] FORD M, LJUNGBERG M, VAN HOOK D J *et al.* Resource brokering service: automatic plan composition and execution [C]//Proc of Military Communications Conference. Washington DC: IEEE Computer Society 2010: 2040-2045.
- [5] ZHENG Xiao-lin, LIN Zhen, YANG Yan-bo *et al.* Research on context-aware Web service composition based on the fluent calculus [J]. *Advances in Information Sciences and Service Sciences* 2011, 3(7): 62-74.
- [6] GERIC S. The potential of service-oriented architectures [C]//Proc of the 32nd International Conference on Information Technology Interfaces. 2010: 471-476.
- [7] PAPAIOGLOU M P, TRAVERSO P, DUSTDAR S *et al.* Service-oriented computing: state of the art and research challenges [J]. *Computer* 2007, 40(11): 38-45.
- [8] YANG Xu, BOUGUETTAYA A. Semantic access to composite M-services [C]//Proc of Advances in Database Systems. Berlin: Springer, 2009: 81-103.
- [9] YANG Xu, BOUGUETTAYA A, LIU Xu-min. Efficient access to composite M-services [C]//Proc of IEEE International Conference on Web Services. Washington DC: IEEE Computer Society, 2009: 381-388.
- [10] NGU A H H, CARLSON M P, SHENG Q Z *et al.* Semantic-based mashup of composite applications [J]. *IEEE Trans on Services Computing* 2010, 3(1): 2-15.
- [11] 员红娟, 叶飞跃, 李霞, 等. 基于语义的 Web 服务发现核心技术研究 [J]. *计算机应用* 2006, 26(11): 2661-2666.
- [12] 马于涛, 张海粟, 刘玉超, 等. 一种 Web 服务综合描述模型 [J]. *电子与信息学报*, 2012, 34(3): 549-556.
- [13] HAYES-ROTH R, BLAIS C. A rich semantic model of track as a foundation for sharing beliefs regarding dynamic objects and events [J]. *Intelligent Decision Technologies* 2008, 2(1): 53-72.
- [14] CARDOSO J, BARROS A, MAY N *et al.* Towards a unified service description language for the Internet of services: requirements and first developments [C]//Proc of IEEE International Conference on Services Computing. Washington DC: IEEE Computer Society 2010: 602-609.
- [15] 张孝国, 黄广君, 郭洪涛. 基于语义的 Web 服务发现技术研究 [J]. *计算机应用* 2008, 28(4): 881-883.
- [16] 邱田, 李鹏飞, 林品. 一个基于概念语义近似度的 Web 服务匹配算法 [J]. *电子学报* 2009, 37(2): 429-432.
- [17] 邱田, 胡晓惠, 李鹏飞, 等. 基于 OWL_S 的服务发现语义匹配机制 [J]. *电子学报* 2010, 38(1): 42-47.
- [18] 张莹, 黄厚宽, 杨冬, 等. 基于 Chord 的带有 QoS 的语义 Web 服务发现方法研究 [J]. *电子与信息学报* 2009, 31(3): 711-715.
- [19] ZHENG G, BOUGUETTAYA A. Service mining on the Web [J]. *IEEE Trans on Services Computing* 2009, 2(1): 65-78.
- [20] 周相兵, 马洪江, 常桂娟. 基于 Topic Maps 的语义 Web 服务组合引擎研究 [J]. *计算机应用* 2010, 30(1): 162-166.
- [21] GHARZOULI M, BOUFAIDA M. A distributed P2P-based architecture for semantic Web services discovery and composition [C]//Proc of the 10th Annual International Conference on New Technologies of Distributed systems. Washington DC: IEEE Computer Society 2010: 315-320.
- [22] 马建刚, 黄涛, 汪锦岭, 等. 面向大规模分布式计算发布订阅系统核心技术 [J]. *软件学报* 2006, 17(1): 134-147.
- [23] 岳昆, 王晓玲, 周傲英. Web 服务核心支撑技术: 研究综述 [J]. *软件学报* 2004, 15(3): 428-442.
- [24] WANG Hai, LI Zeng-zhi, FAN Lin. Capability matchmaking of semantic Web services with preconditions and effects [J]. *Journal of Southeast University: English Edition* 2009, 25(4): 464-467.
- [25] 仲梅, 宋顺林. 一种语义 Web 服务的多层次匹配方法 [J]. *计算机应用*, 2007, 27(1): 100-204.
- [26] LIANG Ping. An approach to semantic re-composition of Web services for service-oriented architecture [C]//Proc of the 2nd International Conference on Computer Engineering and Technology. Washington DC: IEEE Computer Society 2010: 519-522.
- [27] BOUILLET E, FEBLOWITZ M, FENG Han-hua *et al.* A folksonomy-based model of Web services for discovery and automatic composition [C]//Proc of IEEE International Conference on Services Computing. Washington DC: IEEE Computer Society 2008: 389-396.
- [28] WANG Hui, FENG Zhi-yong, SUI Yang *et al.* Service network: an infrastructure of Web services [C]//Proc of IEEE International Conference on Intelligent Computing and Intelligent Systems. Washington DC: IEEE Computer Society 2009: 303-308.
- [29] FOSTER H, UCHITEL S, MAGEE J *et al.* An integrated workbench for model-based engineering of service compositions [J]. *IEEE Trans on Services Computing* 2010, 3(2): 131-144.
- [30] ZOU Ying, HUA Xiao, NIGUL L *et al.* Workshop on automatic service composition [C]//Proc of Conference of the Center for Advanced Studies on Collaborative Research. Riverton: IBM corp, 2009: 378-380.
- [31] TIAN Yong-hong, SU Yi-la, ZHUANG Xu-fei. Research on service identification methods based on SOA [C]//Proc of the 3rd International Conference on Advanced Computer Theory and Engineering. Washington DC: IEEE Computer Society 2010: 27-31.
- [32] SHALIL M, WALKER D W, GRAY W A. A framework for automated service composition in service-oriented architectures [C]//Proc of the 1st European Semantic Web Symposium. Berlin: Springer 2004: 269-283.
- [33] ZHENG Hui-yuan, ZHAO Wei-liang, YANG Jian *et al.* QoS analysis for Web service composition [C]//Proc of IEEE International Conference on Services Computing. Washington DC: IEEE Computer Society 2009: 235-242.
- [34] 李鑫, 程渤, 杨国纬, 等. 一种基于事件的 Web 服务组合方法 [J]. *软件学报* 2009, 20(12): 3101-3116.
- [35] 顾志峰, 李涓子, 胡建强, 等. Web 服务之间数据关联的建模与应用 [J]. *计算机学报*, 2008, 31(8): 1309-1318.
- [36] GALATOPOULLOS D G, KALOFONOS D N, MANOLAKOS E S. A P2P SOA enabling group collaboration through service composition [C]//Proc of the 5th International Conference on Pervasive Services. New York: ACM Press, 2008: 111-120.
- [37] TILSNER M, FIECH A, SPECHT T. Integrating heterogeneous user interfaces in service oriented Web applications [C]//Proc of the 2nd Canadian Conference on Computer Science and Software Engineering. New York: ACM Press 2009: 73-81.
- [38] VIGGH D H E M, WEED C, CHAN D M T *et al.* Composable applications using service encapsulation (CAUSE) [C]//Proc of Military Communications Conference. Washington DC: IEEE Computer Society 2010: 2046-2051.
- [39] 刘华文, 申春, 杨冬, 等. 语义 Web 服务基础技术研究综述 [J]. *吉林大学学报: 信息科学版* 2010, 28(1): 47-54.