

基于 RDF4S 语义服务描述模型的服务资源搜索框架^{*})

曾志浩 应 时 曹虹华

(武汉大学软件工程国家重点实验室 武汉 430072)

摘 要 语义服务资源搜索是一个满足服务请求者快速、便捷地定位可用语义服务资源的过程。一个高效、精确的基于语义的语义服务资源搜索框架和经过精心设计并实现的搜索工具是语义服务框架中的一个必要组件。提出一种基于 RDF4S 语义服务描述模型的服务资源搜索框架,详细说明了该搜索框架的结构。搜索框架根据 RDF4S 语义服务描述模型的 QoS、执行、功能和接口四种语义标注元素类型,分别采用不同的搜索匹配策略,可以很大程度上提高搜索的效率和精度。

关键词 语义服务, RDF4S, 语义服务搜索

RDF4S Based Semantic Web Service Search Framework

ZENG Zhi-hao YING Shi CAO Hong-hua

(State Key Lab of Software Engineering, Wuhan University, Wuhan 430072, China)

Abstract Semantic service search is a process of locating available semantic service recourse conveniently and accurately for users. Therefore, the semantic Web service search tool is one of the absolutely necessary components in semantic Web Service framework. The discovery and selection process finds matches between requirements and advertisements according to their semantic description. Proposed a discovery framework for semantic Web service described by RDF4S specification and introduced its structure in detail. The framework provides different search strategies based on four semantic annotation elements: QoS, execution, function and interface of RDF4S specification, which can help to enhance the efficiency and accuracy of search framework remarkably.

Keywords Semantic Web service, RDF4S, Semantic service discovery

1 引言

语义 Web 服务^[1]是语义 Web 技术与 Web 服务技术的有机结合,通过对 Web 服务进行语义标注,以支持服务的自动发现、自动组合和运行。现较成熟的语义服务描述框架有 WSDL-S^[2], OWL-S^[3]和 WSMO^[4]。RDF4S 是我们提出的一种新颖的语义服务资源描述模型,该模型定义了 QoS 语义、执行语义、功能语义和接口语义等 4 种服务资源语义标注信息。在 RDF4S 语义服务描述规范基础之上,我们提出了基于语义 Web 服务的 SPL 语义处理语言,利用该语言和 UML 2.0 活动图,可以实现一种基于语义 Web 服务的软件设计方法。

语义服务资源搜索是一个满足服务请求者快速、便捷地定位可用语义服务资源的过程。一个高效、精确的基于语义的语义服务资源搜索框架和经过精心设计并实现的搜索工具是语义服务框架中的一个必要组件。语义服务资源的搜索过程分为两个连续的阶段:服务定义阶段和服务搜索阶段。服务定义阶段通过定义服务描述规范来确定服务本身;服务搜索阶段则关心如何找出与服务请求相关的服务资源。本文提出一种基于 RDF4S 语义服务描述模型的服务资源搜索框架,详细说明了该搜索框架的结构。搜索框架根据 RDF4S 语义服务描述模型的 QoS、执行、功能和接口 4 种语义标注元素分

别采用不同的搜索匹配策略,在很大程度上提高了搜索的效率和精度。

本文第 2 节介绍了相关工作;第 3 节说明基于 RDF4S 语义服务描述模型的搜索框架结构及搜索工作过程;第 4 节介绍基于 RDF4S 语义服务描述模型的服务资源搜索框架设计;最后对全文进行总结并对下一步的工作进行展望。

2 相关工作

近年来随着语义技术的成熟和语义服务资源的丰富,关于语义服务资源的搜索和发现的研究工作也取得了很大的进展。文献[5]基于 WSDL-S 规范基础提出了一种新颖的发现不同服务间语义关系的方法。通过服务的本体性描述规范来识别不同服务间前置条件和后置条件间的语义关系,并在此基础上进行服务资源的搜索和发现。文献[6]提出能够在某种具体的服务描述规范基础之上进行基于语义的匹配,并基于 DAML-S 规范提出了 4 种服务资源匹配类型。利用 DAML-S 描述规范中的服务概要规范来进行服务资源的匹配,设计了相应的匹配算法,该算法并不局限于某一特定的服务描述规范。文献[7]基于 WSMO 规范,关注语义服务资源的功能性描述,针对现有描述规范缺乏明确的语义性功能描述的问题,用一种称为“抽象状态空间 (Abstract State Space)”的包含丰富语义,并独立于服务描述规范的语义服务

^{*})国家重点基础研究发展计划 973 项目(2007CB310800),国家高技术研究发展计划 863 项目(2006AA01Z168),国家自然科学基金项目(60773006)。曾志浩 博士生,主要研究方向为 SOA、语义 Web 服务;应 时 教授,博士生导师,主要研究领域为互联网上的软件工程、软件体系结构和模式;曹虹华 博士生,主要研究领域为 SOA、语义 Web 服务。

资源模型来完成服务搜索和发现功能。基于该模型,可以得到服务间概念以及服务的功能性描述的精确数学定义。文献[8]提出了基于本体论和词汇语义相似度的服务发现方法。通过构建 Web 服务本体,给出一个明晰的 Web 服务发现对象,指出可对 Web 服务进行的几种相似度计算,为 Web 服务相似度计算、Web 服务发现提供了一种有效可行的方法。文献[9]提出了一种 QoS 有保障的 Web 服务分布式发现模型,基于 QoS 约束的服务发现问题,以及如何提高服务发现系统的可用性问题,研究 Web 服务的三维 QoS 模型并设计基于 QoS 约束的服务选择算法,提出了集成服务选择算法的 UDDI 兼容扩展模型,并运用可扩展 Kautz 图和 Bloom Filters 理论,提出了分布式 UDDI 的实现机制。

3 基于 RDF4S 语义服务描述模型的服务资源搜索框架

在参考上述相关研究工作之后,我们提出了一种基于 RDF4S 语义服务描述模型的服务资源搜索框架。显然,基于服务描述规范进行服务资源搜索,首先必须有一种服务资源描述规范语言对服务本身进行描述,而后在此基础上,服务请求者才能根据某种搜索策略得到它所需要的可用服务资源。下面我们首先将简要介绍 RDF4S 语义服务描述模型规范,再详细介绍如何基于 RDF4S 描述规范来构造服务资源搜索框架及相应的搜索过程。

3.1 RDF4S 语义标注元素

表 1 用于搜索的 RDF4S 语义标注元素

语义标注类别	标注名称	标注类型	标注示例
QoS 语义	服务 QoS	断言	<qos4s><cost>0</cost><reputation level="10"/><regulations><regulation>rd4s=0.7</regulation><regulation>soap=2.0</regulation></regulations></qos4s>
			操作 QoS
执行语义	前置操作	断言	<preoperation><logged>>true</logged></preoperation>
			后置操作
功能语义	分类	分类法 关键字	<category Name="travel" URI="NACS-category" value="products services" code="561599"/>
			服务能力
接口语义	接口	本体概念	<rd4s: interface name="OlympicInterface" ontoreference="http://example.org/categorization/products/travel">
			操作
输入	输出	本体概念	<input element="OlympicPlaceQueryInput"/>
			<output element="Gymnasium"/>
错误消息	本体概念		<fault name="Fault" ontoreference="http://www.sklse.org/2007/ontology/onto#commonFault"/>

RDF4S 语义服务描述模型中包含充分的语义标注元素。其中 QoS 语义信息用于描述语义服务的服务质量,便于提高语义服务自动发现和组合的质量;执行语义信息用于描述语义服务执行时的前置、后置操作,便于进行语义服务的组合和动态配置;功能语义信息用于描述语义服务的功能,包括服务的分类、服务提供的能力等,便于进行精确的语义服务自动发现;接口语义信息用于描述语义服务的操作、消息,便于进行语义服务自动发现和组合。QoS 语义、执行语义、功能语义和接口语义这 4 种语义标注元素都可用于服务资源的搜索过程。4 种语义标注元素的具体内容如表 1 所示。

3.2 基于 RDF4S 语义服务描述模型的服务资源搜索框架结构

基于 RDF4S 语义服务描述模型的搜索框架包括 6 个主要部分:资源获取模块、RDF4S 语义服务资源快照、语义索引库、断言及关键字索引库、概念索引库和资源匹配模块。

资源获取模块负责将部署在互联网中的符合 RDF4S 服务资源描述规范的服务资源描述文件提取出来。提取的服务资源描述文件经过 RDF4S 解析后,一方面形成 RDF4S 语义服务资源快照,另一方面进行概念索引后按表 1 中给出的语义标注元素分别建立索引表供资源匹配模块使用。

RDF4S 语义服务资源快照保存已获取的语义服务资源描述文件,最终返回给用户的搜索结果也将从中获取。语义索引库保存资源获取模块建立的概念索引表。概念索引库保存由本体概念爬虫根据推理规则将领域本体交由推理机处理后生成的概念索引。断言及关键字索引库保存服务资源描述文件中包含的断言和分类关键字。语义索引库、概念索引库以及断言及关键字索引库为资源匹配模块提供语义信息,支撑基于 RDF4S 描述规范的语义匹配过程。

资源匹配模块负责将用户的查询条件按照表 1 中列出的语义标注元素进行解析,并对解析后的查询条件分别与对应的索引项进行匹配。

图 1 所示为本文提出的基于 RDF4S 语义服务描述模型的服务资源搜索框架结构示意图。

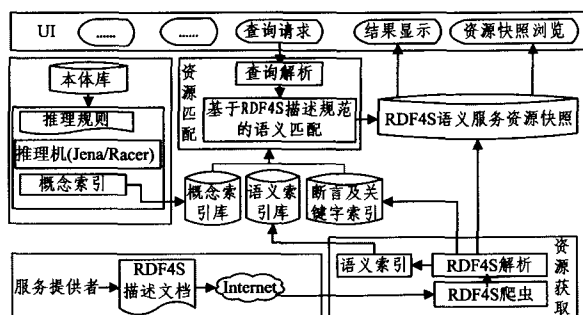


图 1 基于 RDF4S 语义服务描述模型的服务资源搜索框架示意图

3.3 基于 RDF4S 语义服务描述模型的服务资源搜索过程

图 1 中资源匹配模块负责完成语义服务资源的语义匹配操作。资源匹配根据 RDF4S 描述规范分成 3 个部分:断言匹配、关键字匹配和语义匹配。搜索引擎对用户的查询条件进行三种类型的匹配操作后,以 RDF4S 语义服务资源快照的形式返回给用户,以使用户对服务集进行进一步的操作。图 2 为基于 RDF4S 语义服务描述模型的服务资源搜索过程示意图。

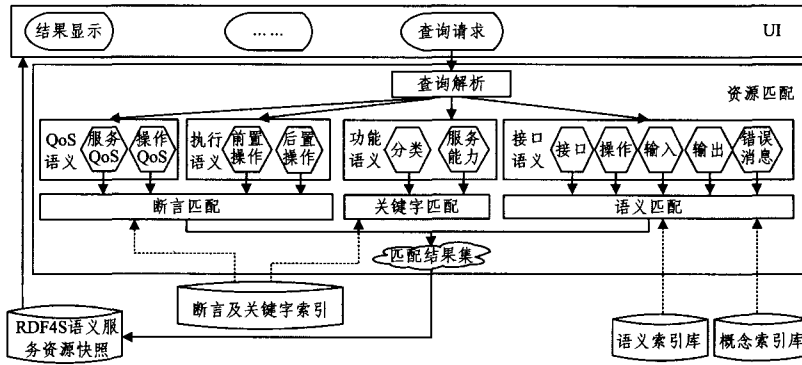


图2 基于RDF4S语义服务描述模型的服务资源搜索过程示意图

4 基于RDF4S语义服务描述模型的服务资源搜索框架设计

4.1 RDF4S语义服务资源快照

RDF4S服务资源快照类似于传统搜索工具的网页快照。资源快照用于保存RDF4S语义服务资源描述文件。为了区分保存的资源描述文件，必须为每个资源描述文件添加一个唯一标识。资源描述文件经过RDF4S解析后按照语义标注元素分类后进行保存，如图3所示。

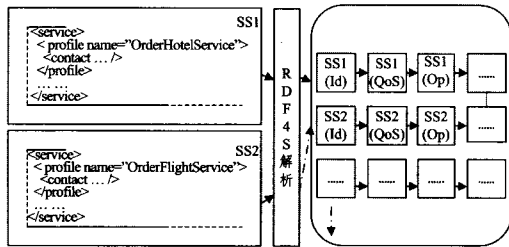


图3 RDF4S语义服务资源快照

4.2 概念索引库

概念索引库是参考领域本体库，根据推理机制，由推理机为服务相关概念建立的等价概念和子类概念集合倒排索引表。概念索引库建立过程如图4所示。

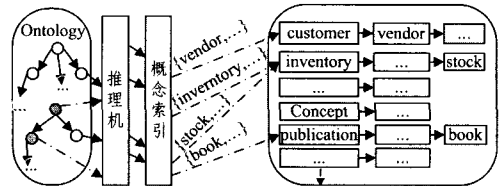


图4 概念索引库

4.3 服务资源语义索引库

服务资源语义索引库是支撑服务资源匹配的关键部分，索引库的结构必须符合资源匹配处理时的需要。在语义匹配处理过程中，它将参考概念索引库的内容实现服务语义匹配。

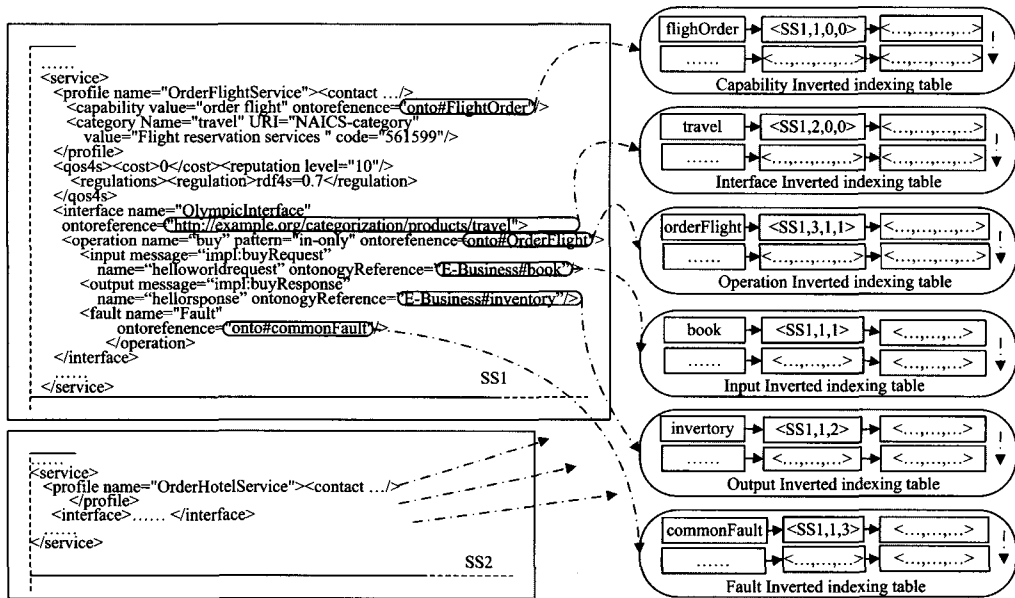


图5 服务资源语义索引库

服务资源语义索引库共6个倒排索引表。根据RDF4S规范，分别是服务能力标注索引表、接口标注索引表、操作标注索引表、输入标注索引表、输出标注索引表和错误消息标注索引表。为了便于实现，在实际设计时可前3个概念索引

表合并成一个索引表，索引表中索引项是一个四元组(服务标识, 标注类型, 操作数量, 概念位置标识)。其中服务标识为服务唯一标识编号。标注类型用于匹配具体语义标注类型，可取值为1, 2, 3。其中1表示服务能力标注, 2表示接口标注, 3

表示操作标注。操作数量用于说明服务共包含几个操作,可取值为 $0, \dots, n$ (n 为操作的最大数量)。0表示该服务没有操作标注,当标注为服务能力标注或接口标注时,操作数量元素取值也为0。概念位置标识只在操作数量元素非零时有意义,具体表示该索引概念在某个具体服务中属于哪个操作。例如,概念 flightOrder 连接四元组 $\langle ss1, 1, 0, 0 \rangle$,表示服务 ss1 的服务能力标注包含概念 flightOrder;概念 orderFlight 连接四元组 $\langle ss1, 3, 2, 1 \rangle$,表示服务 ss1 有两个操作,其中第一个操作标注包含概念 orderFlight。

出于提高搜索速度的考虑,在实际设计时可将后3个概念索引表合并成一个索引表。索引表中索引项是一个三元组 \langle 服务标识,概念位置标识,标注类型 \rangle 。其中服务标识为服务唯一标识编号。概念位置标识用于区分概念属于服务描述中的哪个操作。标注类型用于匹配具体语义标注类型,可取值为1,2,3。其中1表示输入标注,2表示输出标注,3表示错误消息标注。例如概念 book 连接三元组 $\langle ss1, 1, 1 \rangle$,表示服务 ss1 的第一个操作的输入标注包含概念 book;概念 inventory 连接三元组 $\langle ss1, 2, 2 \rangle$,表示服务 ss1 的第二个操作的输出标注包含概念 inventory。服务资源语义索引库内部结构如图5所示。

4.4 断言及关键字索引

RDF4S 语义服务描述模型中 QoS 语义的服务级别 QoS 和操作级别 QoS 以及执行语义的前置操作和后置操作是通过断言进行声明的。因此,不同于有本体概念标注的语义,在资源匹配处理时必须基于断言进行匹配。与语义索引库的建立过程类似,同样建立倒排索引供资源匹配处理使用。

功能语义中的服务分类是通过分类法建立的。因此,对功能语义的服务分类进行资源匹配时可以只考虑传统的关键字匹配,而无需其它特别处理。对功能语义的服务分类涉及到的分类法关键字进行倒排索引后,建立服务分类关键字索引供资源匹配处理使用。

结束语 为使用户能快速、精确地搜索已发布的 RDF4S 语义服务资源,本文提出一种基于 RDF4S 语义服务描述模型的服务资源搜索框架,详细说明了该搜索框架的结构。整个基于 RDF4S 语义服务描述模型的搜索框架主要由资源获取

模块、RDF4S 语义服务资源快照、服务资源概念索引库、断言及关键字索引库、概念索引库和资源匹配模块6部分组成。RDF4S 语义服务资源的匹配主要分为断言匹配、关键字匹配和语义匹配。搜索框架根据 RDF4S 语义服务描述模型的 QoS、执行、功能和接口4种语义标注元素分别采用不同的搜索匹配策略,可以在很大程度上提高搜索的效率和精度。下一步的研究工作将围绕两方面进行:一方面将基于 Eclipse 开发平台实现上述基于 RDF4S 语义服务描述模型的服务资源搜索框架;另一方面将继续研究在资源匹配过程中不同匹配策略,以提高 RDF4S 语义服务资源搜索框架的搜索效率和精确度。

参考文献

- [1] Burstein M, Bussler C, Zaremba M, et al. A Semantic Web Services Architecture. *IEEE Internet Computing*, 2005, 9 (5): 72-81
 - [2] WSDL-S, W3C Member Submission on Web Service Semantics. <http://www.w3.org/Submission/WSDL-S/>
 - [3] Martin D, Burstein M, Hobbs J, et al. OWL-S Semantic Markup for Web Services. W3C submission 2004. <http://www.w3.org/Submission/OWL-S/>
 - [4] Roman D, Keller U, Lausen H, et al. Web Service Modeling Ontology. *Applied Ontology*, 2005, 1(1): 77-106
 - [5] Lin L, Arpinar I B. Discovery of Semantic Relations between Web Services//*International Conference of Web Services*. 2006; 357-364
 - [6] Paolucci M, Mawamura T, Payne T R, et al. Semantic Matching of Web Service Capabilities//*First International Semantic Web Conference 2002*. 2002; 333-347
 - [7] Keller U, Lausen H, Stollberg M. On the Semantics of Functional Descriptions of Web Services//*3rd European Semantic Web Conference (ESWC)*. 2006; 605-619
 - [8] 吴健, 吴朝晖, 李莹, 等. 基于本体论和词汇语义相似度的 Web 服务发现. *计算机学报*, 2005, 28(4): 595-602
 - [9] 郭得科, 任彦, 陈洪辉, 等. 一种 QoS 有保障的 Web 服务分布式发现模型. *软件学报*, 2006, 17(11): 2324-2334
-
- [3] SCTE. Data - Over - Cable Service Interface Specifications - Radio Frequency Interface Specification. 2002; 20-22
 - [4] Fellows D, Jones D. DOCSIS cable modem technology[J]. *Communications Magazine*, IEEE, 2001, 39(3)
 - [5] Sdralia V, Holcombe M. Ranging schemes for fast dynamic recovery of DOCSIS networks[J]//*Proc. 9th IEEE International Conference on Networks (ICON'01)*. Bangkok, Thailand, 2001
 - [6] Sdralia V. Optimized recovery of DOCSIS networks using reserved persistent ranging[J]//*Proc. IEEE Globecom 2001*. San Antonio, TX, USA, 2001
 - [7] Lin Y D, Yin W M, Huang C Y. An Investigation into HFC MAC Protocols[J]; *Mechanisms, Implementation, and Research Issues*. *IEEE Communication Surveys*. <http://www.comsoc.org/pubs/surveys>. 3rd Quarter 2000
 - [8] Carroll M D. Aligning the Initial Maintenance Intervals of Cable Modem Upstream Channels[J]. *IEEE Communication Magazine*, September 2003

(上接第 106 页)

见惯的编码方式进行一些改进,在不破坏网络编码整体规范性的前提下,缩短数据包的长度无疑是提高通信效率的一种方法。本文提出的优化 TLV 编码的方法,虽然增加了一些编码和解码的复杂性,缩小了可以一次编码的参数类型的数量,但却可以有效地缩短网络上大量数据包的长度。随着处理器速度的提高,数据处理复杂性带来的影响正在减弱,而网络上与日俱增的数据量,以及有限的网络时间和带宽才是约束网络系统性能提高的瓶颈,所以本文提出的这种对 TLV 编码优化的方法有着广泛的应用前景。

参考文献

- [1] ITU-T Study Group. OSI networking and system aspects-Abstract Syntax Notation One, X. 690. 2002; 3-4
- [2] CableLabs. Data-Over-Cable Service Interface Specifications Radio Frequency Interface Specification. SP-RFIV1. 1-I06-001215. 2000; 82-85