

基于领域本体的分布式 UDDI 互联架构

马欣宇^{1,2}, 刘 琼¹, 钱乐秋^{1,2}

(1. 复旦大学计算机科学与工程系软件工程实验室, 上海 200433; 2. 复旦大学上海市智能信息处理重点实验室, 上海 200433)

摘要: 传统的 UDDI 工作方式存在数据同步和复制开销巨大以及无法体现 UDDI 本身领域特性的问题, 该文通过领域本体对 UDDI 的传统工作机制进行扩展, 提出了 UDDI 代理 (UDDI Broker) 的概念, 通过 UDDI 代理构建分布式的 UDDI 互联架构, 阐述了使用 UDDI 代理进行服务发布和服务请求的工作过程及相关问题。

关键词: UDDI; 领域本体; DAML-S; UDDI 代理

Distributed Architecture of UDDI Interconnection Based on Domain Ontology

MA Xinyu¹, LIU Qiong¹, QIAN Leqiu^{1,2}

(1. Lab of Software Engineering, Department of Computer Science and Engineering, Fudan University, Shanghai 200433;

2. Shanghai Key Laboratory of Intelligent Information Processing, Fudan University, Shanghai 200433)

【Abstract】 The traditional architecture of UDDI becomes more and more inappropriate to industrial reality. The cost of data synchronization and replication is huge, and domain-specific attributes of UDDI are not contained. This paper expands the traditional architecture of UDDI. A component called UDDI Broker is proposed, which contributes to the new distributed UDDI architecture. Also the mechanism of service advertisement and request according to UDDI Broker is presented.

【Key words】 UDDI; domain ontology; DAML-S; UDDI broker

1 概述

UDDI是构成Web Services的重要基础设施之一, 在整个Web服务体系内提供服务发现功能, 传统的UDDI工作方式类似于P2P网络, 各个UDDI节点之间进行数据的同步和复制^[1]。服务请求者从不同的UDDI节点进入是没有差别的, 它所面对的都是整个UDDI体系提供的服务发现功能; 服务发布者在某一个UDDI节点进行服务的发布, 然后通过数据同步和复制把服务扩散到整个UDDI体系中。这种工作方式的问题主要有以下几点: (1)各个节点之间需要持续进行数据的同步和复制, 以便维持数据的全局一致性。不仅大大增加了系统的开销, 也增加了维护UDDI数据的难度, 而且由于不同UDDI之间工作机制和数据格式的差异, 因此常常无法进行数据的同步和复制。(2)无法体现不同UDDI节点之间的领域差异。很多企业依据行业特性建立了私有的UDDI, 它所注册的服务具有领域专门的特性, 然而目前UDDI的工作方式把不同领域的服务混杂在一起存放, 直接导致服务查找的效率较低。

针对以上问题, 本文提出了UDDI代理(UDDI Broker)的概念, 各个UDDI节点在UDDI代理注册, 通过引入领域本体对UDDI节点进行分类存储, 服务发布和请求时UDDI代理使用推理机在服务本身和已注册的UDDI节点之间进行领域相关性匹配, 并把服务提交给匹配度最佳的若干UDDI节点进行实际的服务发布和请求处理。UDDI代理与UDDI节点之间关系的总体架构如图1所示。UDDI代理所接收的外部操作分为两个层次: 一个是UDDI节点的注册(UDDI Register), 负责接收UDDI节点的注册并对其进行管理; 另一个是接收服务的发布(service advertisement)和请求(service request), 负责实际的服务发现功能。

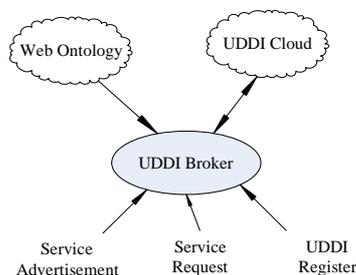


图1 总体架构

采用这种结构, 一方面不需要进行数据的同步和复制, UDDI代理把服务的发布和请求分配到各个UDDI节点上; 另一方面通过引入领域本体对UDDI节点进行分类存储, 充分利用它们之间的领域差异, 把服务进行分类后提交给领域相关的UDDI节点进行处理, 可以提高服务查找的效率。UDDI代理的内部架构见图2。

服务发布和请求在UDDI代理中的处理过程为: 服务提供者使用UDDI代理公布的API(UDDI broker API)发布服务, UDDI代理把服务发布转给领域推理机(domain reasoning engine), 匹配出领域相关的UDDI节点, 然后把服务发布信息连同UDDI节点信息一并转给服务提交引擎(service submission engine), 服务提交引擎连接到具体的UDDI节点发布这个服务, 最终服务发布信息存放在若干具体的UDDI节点上。

作者简介: 马欣宇(1981 -), 男, 硕士, 主研方向: 软件工程及Web Services与SOA; 刘 琼, 硕士; 钱乐秋, 博士生导师
收稿日期: 2006-07-30 **E-mail:** xinyuppp@163.com

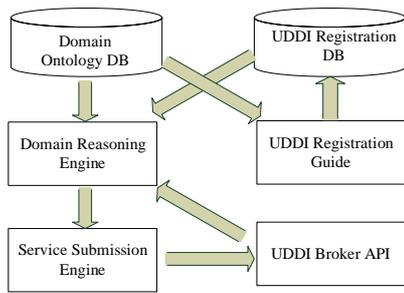


图2 UDDI Broker 内部架构

服务请求与服务发布的过程类似，服务请求者向 UDDI 代理发出请求服务的要求，UDDI 代理负责连接到领域相关的 UDDI 节点去请求服务，把返回的结果递交给服务请求者。

2 领域本体库与 UDDI 节点注册

通过 UDDI 代理为 UDDI 节点本身提供注册服务，同时利用领域本体库对 UDDI 节点进行分类存储，在处理服务的发布和请求时，领域推理机对服务本身的领域特性和 UDDI 节点的领域特性进行匹配，找出与服务领域相关的 UDDI 节点，服务只会提交给这些 UDDI 节点而不是所有的 UDDI 节点，大大提高了服务查找的效率。

UDDI 内部提供了对服务进行标准化分类的支持，它内嵌了几种分类的标准，包括 NAICS (North American Industry Classification System) 等^[1]。使用标准的 NAICS 编码^[2] 建立领域本体库，图 3 是 NAICS 编码的一个片段。领域本体库存储它的本体层次结构信息以及对应的 NAICS 标准编码。

1114	Greenhouse, Nursery, Floriculture Production
11141	Food Crops Grown Under Cover
111411	Mushroom Production
111419	Other Food Crops Grown Under Cover
11142	Nursery and Floriculture Production
111421	Nursery and Tree Production
111422	Floriculture Production

图3 领域本体层次及编码

UDDI 节点使用 UDDI 注册向导 (UDDI Registration Guide) 完成在 UDDI 代理的注册，UDDI 注册向导根据领域本体库 (Domain Ontology DB) 的指导提供领域分类的编码方式信息，供 UDDI 节点描述自己的领域特性。UDDI 节点提交领域编码的组合来描述自己的领域特性，见图 3，如一个 UDDI 节点涵盖的领域包括 Mushroom Production 和 Floriculture Production，那么它提交的编码组合就是 (111411, 111422)，下面介绍这种编码的提交规则：

基于编码的层次性，提交一个编码就意味着它的所有子领域该 UDDI 节点全都涵盖，譬如一个 UDDI 节点提交的编码组合中包括 11141 (Food Crops Grown Under Cover)，那么意味着这个编码对应领域的所有子领域，包括 111411 (Mushroom Production) 和 111419 (Other Food Crops Grown Under Cover) 该 UDDI 节点全都涵盖；如果这个 UDDI 节点实际只涵盖 Mushroom Production 这个子领域而不包括 Other Food Crops Grown Under Cover 子领域，那么它所提交的编码组合应该是 111411 而不是 11141。

注册向导对 UDDI 节点最终提交的编码组合进行检查，一方面删除冗余，如 UDDI 节点提交的编码组合中同时包括 11141 和 111419，那么注册向导会把 111419 这个编码删除，只保留 11141 这一个编码；另一方面进行合并，如 UDDI 节点提交的编码组合中同时包括 111411 和 111419，那么注册向导会实际存储 11141 这个编码。

3 服务描述与服务领域相关性匹配

使用 DAML-S^[3] 来描述服务，它可以提供关于服务更为详细的领域信息，并且基于它和 DAML+OIL 的内在联系，还可以充分利用语义网所能提供的本体信息。目前已有一些相关工作研究如何把 DAML-S 内嵌于已有的 UDDI 框架中，文献 [4] 提供了一种借助 UDDI 中 tModel 的扩展机制把 DAML-S Service Profile 映射到 UDDI Representation 的方法，借助 DAML-S 所能表达的领域分类信息来进行服务与 UDDI 节点的领域相关性匹配，指导 UDDI 代理对服务的实际指派。

DAML-S Service Profile 中的功能属性 ServiceCategory 可以用来描述服务的领域特性，它的实例包括以下几个^[3]：categoryName, taxonomy, value 和 code，分别用来指出所使用的分类方法名称、模式以及具体的分类编码值等。图 4 给出了描述一个服务发布的片段，侧重展示了如何表示它所使用的分类标准以及如何标识自己的领域特性，使用 NAICS02 来描述服务的领域特性，所使用的编码片段见图 3。使用 DAML-S Service Profile 描述一个服务请求的方法类似。

```

<profile:Profile rdf:ID="Mushroom Production">
  <profile:serviceName>Mushroom Production</profile:serviceName>
  <profile:providedBy>...</profile:providedBy>
  <ServiceCategory>
    <profile:categoryName>NAICS02</profile:categoryName>
    <profile:taxonomy
      rdf:resource="http://www.census.gov/epcd/naics02/
      naicod02.htm"/>
    <profile:code>111411</profile:code>
  </ServiceCategory>
</profile:Profile>

```

图4 描述服务发布的领域特性

UDDI 代理接收到一个服务发布或请求后，把它转给内部的领域推理机进行处理，领域推理机取出服务使用的分类编码值，然后到 UDDI 注册数据库中遍历全部的 UDDI 节点注册信息，对它们双方的编码值进行领域相关性的匹配。由于领域本体自身所具有的层次关系，因此可以利用这种关系度量领域相关性，具体的匹配结果有以下几种：

(1) 服务的领域编码值与某个 UDDI 节点注册的领域编码组合中的某一个值完全相同，这种情况视为完全匹配，这个 UDDI 节点会被选择进行实际的服务发布或者服务请求。

(2) 服务的领域处在某个 UDDI 节点注册领域的子领域层次上，由于 UDDI 节点注册时的要求，这个 UDDI 节点的领域应该涵盖所有的子领域，即涵盖服务的领域，因此这种情况也视为完全匹配，这个 UDDI 节点会被选择进行实际的服务发布或者服务请求。

(3) 服务的领域处在某个 UDDI 节点注册领域的父领域层次上，这种情况下可以简单地根据一个比例来衡量匹配的程度。见图 3 的编码片段，假定服务的领域是 11142，而某个 UDDI 节点注册的领域是 111421 (其他可能的领域编码都不是 11142 的子领域)，那么这个 UDDI 节点有 1 个领域是服务领域的子领域，而服务领域一共有 2 个子领域，因此度量匹配程度为 1/2，即 50%。可以给服务提供者或请求者以设定匹配程度阈值的权限，由它们设定匹配过滤的最低程度，如把匹配程度设定为 75%，那么所有最高匹配程度低于 75% 的 UDDI 节点都不会参与最终的服务实际处理。

(4) 服务的领域与 UDDI 节点注册时的任意一个领域都没有上下的继承关系，这时候视为不匹配，这个 UDDI 节点不会参与服务最终的实际处理。

(下转第 53 页)