基于服务簇的 Web 服务绑定方法

宁玉辉[1],杨 栋[2],杜玉越[1]

(1. 山东科技大学 信息科学与工程学院,山东 青岛 266590;
2. 中国人民解放军 71939 部队,山东 济南 250014)

摘 要:针对较大粒度服务簇的服务绑定效率优化问题,提出了一种 Web 服务绑定方法。通过改进服务簇的逻辑 Petri 网模型,量化服务参数集,缩小相似度计算总量,只需对用户需求参数和服务参数进行运算,计算复杂度为 $O(km)$,提高了服务发现效率;给出服务绑定的动态替换方法,匹配服务过程无需查找本体树,能够提高运算效率;解决了网络环境的变化造成的服务响应失效问题,提高服务响应的自适应性。

关键词:服务簇;服务绑定;服务替换;逻辑 Petri 网;仿真实验

中图分类号:TP311 **文献标志码**:A **文章编号**:1672-3767(2014)04-0094-05

A Web Service Binding Method Based on Service Clusters

Ning Yuhui[1], Yang Dong[2], Du Yuyue[1]

(1. College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao, Shandong 266590, China; 2. PLA Troop 71939, Jinan, Shandong 250014, China)

Abstract: To raise the efficiency of service binding based on service clusters with large granularity, the new method of service binding was introduced. The logic Petri net (LPN) model of service clusters was improved and the parameter set was quantized to reduce the total times of similarity computing and improve service discovery. The computation complexity is $O(km)$. Moreover, service substitution was proposed to solve the problem of service failure without searching concepts from ontology tree. The adaptation of service response is enhanced.

Key words: service cluster; service binding; service substitution; logic Petri net; simulation experiment

Web 服务是一种由 URL(uniform resource locator)标识的软件,因具备网络调用及分布式网络服务等优点,近年来发展迅速[1]。随着 Web 服务开发及应用环境的不断优化,Web 服务数量、种类不断增多,如何高效精准地在海量 Web 服务中选取目标服务成为研究热点。另外,在现有 Web 服务应用环境下,当用户需求发生细微改变或由于网络环境的变化造成当前响应服务失效时,很难快速寻找替代服务,服务响应的自适应性不高。

为解决上述问题,胡强等[2]提出了服务聚类技术。在进行面向用户需求的服务查找前,先进行服务聚类,将功能相同的服务集合在一起作为一个整体提供服务。典型的服务聚类方法,如 Nayak 等[3]将关键词在服务描述文档中出现的频率作为服务相似度,提出了扩展 Web 服务的语义描述,并引入了异构 Web 服务分组的思想。在上述方法基础上,层次聚类算法又应用于相似 Web 服务的聚类过程[4]。针对网格服务,基于服务输入输出及功能本体相似性,提出了一种基于本体聚类的服务发现方法[5]。为提高服务聚类的内聚

收稿日期:2014-03-13

基金项目:国家自然科学基金项目(61170078,61202016);教育部高等学校博士点基金博导类项目(20113718110004);青岛市科技计划基础研究项目(13-1-4-116-jch);山东科技大学科研创新团队支持计划项目(2011KYTD102)

作者简介:宁玉辉(1981—),男,山东德州人,博士研究生,主要从事分布式系统、Web 服务等方面的研究.
杜玉越(1960—),男,山东聊城人,教授,博士生导师,CCF 高级会员,主要从事 CSCW、软件工程、形式化技术、Petri 网等方向的研究,本文通信作者. E-mail: yydu001@163.com

性,孙萍等[6]从服务的功能相似和过程相似两个层面,对服务进行聚类研究。同时,在计算相似度的基础上,提出基于用户需求[7]和用户经验[8]等因素的聚类方法。

上述文献涉及的主要聚类方法是相似度计算。在相似度计算时,查找本体树[9]是必经过程。但本体树结构复杂,且概念节点较多,服务聚类的计算复杂度较高。文献[10]提出一种基于逻辑 Petri 网的 Web 服务簇模型,整合服务参数集,建立服务簇与全部服务的映射关系,将相似度计算控制在服务簇内进行,提高了服务发现效率。然而,对于粒度较大的服务簇,服务数量较多,参数匹配规模大,服务绑定时相似度计算复杂度较高,服务绑定效率还需进一步提高。针对这一问题,本研究在文献[10]的服务簇模型基础上,提出一种 Web 服务绑定及替换方法,改进服务簇的逻辑 Petri 网[11-13]模型,量化服务参数集,使相似度计算仅在量化用户需求时进行,进一步提高服务发现效率。另外,通过给出服务绑定的动态替换方法,解决服务失效问题。

1 基本定义

逻辑 Petri 网(LPN,logic Petri net)是一种高级 Petri 网[14],能够方便地描述批处理及传值不确定系统。

定义 1 六元组 $LPN=(P,T,F,f_I,f_O,M)$为一个逻辑 Petri 网,当且仅当以下条件成立。

1)P 是一个有限库所集。

2)$T=T_D\cup T_I\cup T_O$是一个有限变迁集,且 $T\cup P\neq\emptyset$,$P\cap T=\emptyset$,$\forall t\in T_I\cup T_O:{}^\cdot t\cap t^\cdot=\emptyset$,其中:

①T_D表示传统变迁;

②T_I表示逻辑输入变迁,且$\forall t\in T_I$,逻辑输入库所受到逻辑输入表达式 $f_I(t)$的约束;

③T_O表示逻辑输出变迁,且$\forall t\in T_O$,逻辑输出库所受到逻辑输出表达式 $f_O(t)$的约束。

3)$F\subseteq(P\times T)\cup(T\times P)$是一个有向弧集。

4)I 是一个从逻辑输入变迁到逻辑输入表达式的映射,即$\forall t\in T_I$,$I(t)=f_I(t)$。

5)O 是一个从逻辑输出变迁到逻辑输出表达式的映射,即$\forall t\in T_O$,$O(t)=f_O(t)$。

6)$M:P\to\{0,1\}$是一个标识函数,且$\forall p\in P$,$M(p)$表示 p 中的托肯(token)数。

7)变迁引发规则:

①$\forall t\in T_D$,变迁引发规则与传统 Petri 网的变迁引发规则相同;

②$\forall t\in T_I$,t 是使能的,当且仅当 $f_I(t)|_M=.T.$,其中,$.T.$表示逻辑真,$.F.$表示逻辑假.如果 $M[t>M'$,则$\forall p\in{}^\cdot t:M'(p)=0$;$\forall p\in t^\cdot:M(p)=0,M'(p)=1$;且$\forall p\notin{}^\cdot t\cup t^\cdot:M'(p)=M(p)$;

③$\forall t\in T_O$,t 是使能的,当且仅当$\forall p\in{}^\cdot t:M(p)=1$。若 $M[t>M'$,则$\forall p\in{}^\cdot t:M'(p)=M(p)-1$;$\forall p\in{}^\cdot t\cup t^\cdot:M'(p)=M(p)$;且$\forall p\in t^\cdot$,$M$ 必须满足 $f_O(t)|_{M'}=.T.$,即 $t^\cdot$ 必须满足在 M'时的逻辑输出表达式 $f_O(t)$。

Web 服务是一个包含若干操作、若干输入输出的由 URL 标识的软件。

定义 2 一个 Web 服务可以描述为一个七元组 $Wservice=(Description,Id,Ops,I,O,Links,R)$,其中:$Description$ 为 Web 服务的简单描述;Id 是对服务的标识,唯一确定一个 Web 服务;Ops 表示服务提供的操作;I 是服务的输入参数集,并且每一个输入参数都用本体进行标注;O 表示服务的输出参数集,并且每一个输出参数都用本体进行标注;$Links$ 为服务的端口链接地址集;R 表示 I 与 O,I 与 $Links$,O 与 $Links$ 的映射关系。

2 基于 LPN 的服务簇模型

文献[10]提出一种基于逻辑 Petri 网的 Web 服务簇模型。

定义 3 一个服务簇为一个七元组 $LSC=(CN,CD,CS,I,O,f_I,f_O)$,其中:$CN$ 表示 Web 服务簇的名称;CD 为服务簇功能语义描述;CS 是服务簇中 Web 服务的集合;I 是服务簇的输入参数集;O 是服务簇的输出参数集;f_I 为基于输入参数集 I 的逻辑表达式,用来描述服务簇的输入参数限制;f_O 为基于输出参数集 O 的逻辑表达式,用来描述服务簇的输出参数限制。

定义 3 整合了服务参数集,使相似度计算控制在服务簇内进行。定义 4 则给出一种改进的服务簇定义,把各种参数提前进行基于相似度计算的量化操作,体现服务簇模型的结构特性。

定义 4 改进型服务簇描述为一个八元组 $LSC\_AD=(CN,RL,CS,I_{ad},O_{ad},f_{Iad},f_{Oad},R)$，其中：$CN$ 表示 Web 服务簇的名称；RL 为服务簇中概念量化规则描述；CS 是服务簇中 Web 服务的集合；I_{ad} 是服务簇的输入参数量化集；O_{ad} 是服务簇的输出参数量化集；f_{Iad} 为基于输入参数量化集 I_{ad} 的逻辑表达式，用来描述服务簇的输入参数限制；f_{Oad} 为基于输出参数量化集 O_{ad} 的逻辑表达式，用来描述服务簇的输出参数限制；R 表示 CS 与 I_{ad}，CS 与 O_{ad} 的映射关系。

输入输出参数的量化如算法 1 所示。

算法 1 基于相似度计算的参数量化算法

输入：服务簇 LSC 的输入参数集 $I=\{I_1,I_2,\cdots,I_m\}$，输出参数集 $O=\{O_1,O_2,\cdots,O_n\}$，本体树 $Otree$。

输出：输入参数量化集 I_{ad}，输出参数量化集 O_{ad}。

Step 1 $I_{ad}=\emptyset$，$O_{ad}=\emptyset$。

Step 2 利用 $Otree$ 标注 I 及 O 中元素。

Step 3 选取 $Otree$ 中概念 P_t 为参考点。

Step 4 遍历 I。

①假设当前元素为 I_k，且 $k\in\{1,2,\cdots,m\}$。I_k 的量化值 $I_{k\_ad}=(d_{max}-d(I_k,P_t))/(d_{max}-d_{min})$。

②把 I_{k_ad} 移入输入参数量化集 I_{ad}。

Step 5 遍历 O。

①假设当前元素为 O_k，且 $k\in\{1,2,\cdots,n\}$。O_k 的量化值 $O_{k\_ad}=(d_{max}-d(O_k,P_t))/(d_{max}-d_{min})$。

②把 O_{k_ad} 移入输出参数量化集 O_{ad}。

Step 6 输出 I_{ad}，O_{ad}。

其中，$d(I_k,P_t)$ 是两个概念之间基于本体树的本体距离，为本体树中连接最短路径的边数；d_{max} 及 d_{min} 是本体树中所有概念间距离的最大值及最小值[10]。下面给出改进型服务簇 $LSC\_AD=(CN,RL,CS,I_{ad},O_{ad},f_{Iad},f_{Oad},R)$ 的 LPN 模型的定义。

定义 5 $LPN=(P,T,F,f_I,f_O,M)$ 为一个服务簇 LPN 模型，其中：$P=\{p_1,p_2,\cdots,p_{n+m}\}$ 对应 LSC_AD 的 I_{ad} 和 O_{ad}；$T=t_1\in T_I\&T_O$ 对应 LSC_AD 中 CS 的 Ops；F 对应 LSC_AD 中的 R；f_I 对应 LSC_AD 中的 f_{Iad}；f_O 对应 LSC_AD 中的 f_{Oad}；M 标识了当前的库所 P 的状态。

由定义 5 建立的基于 LPN 的服务簇模型如图 1 所示，其中，$p_1\sim p_n$ 代表 n 个量化参数输入；$p_{n+1}\sim p_{n+m}$ 代表 m 个量化参数输出；t_1 表示输入与输出的 Web 服务触发动作。

3 基于服务簇的服务绑定及动态替换

本节给出用户驱动的基于服务簇的服务绑定方法。用户需求是对所要查找服务的一种描述，一般包含多个原子需求。用户需求提炼为输入输出参数，如定义 6 所示。

定义 6 一个用户需求为一个四元组 $Usrdemand=(UN,I_{ad},O_{ad},R)$，其中，$UN$ 表示用户需求的名称，唯一标识一个用户需求；I_{ad} 是用户需求的输入参数量化集；O_{ad} 是用户需求的输出参数量化集；R 表示 I_{ad} 与 O_{ad} 的映射关系。

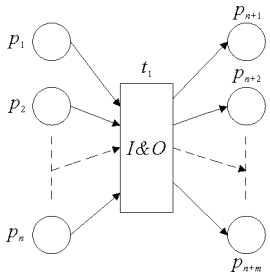

图 1 服务簇的 LPN 模型

Fig. 1 The LPN model of a service cluster

用户需求参数集的量化算法与算法 1 相同。算法 2 为基于服务簇的服务绑定算法。

算法 2 服务绑定算法

输入：服务簇 $LSC\_AD=(CN,RL,CS,I_{ad},O_{ad},f_{Iad},f_{Oad},R)$；用户需求 $Usrdemand=(UN,I_{ad},O_{ad},R)$。

输出：Web 服务的 URL 链接集 $Slinks$。

Step 1 $Slinks=\emptyset$。

Step 2 遍历 $Usrdemand.O_{ad}$。

1)假设当前元素为 $O_{ad}.O_{adn}$，由 $Usrdemand.R$ 得到与 $O_{ad}.O_{adn}$ 关联的 I_{ad} 的元素集 SI_{ad}。

2)遍历 $LSC\_AD.O_{ad}$。

a)假设当前元素为 $O_{ad}.O_{adm}$,设定变量集 $x=\{x_1,x_2,\cdots,x_k\}$,其中所有元素值赋值为 $+\infty$。令 $x_m=|LSC\_AD.O_{ad}.O_{adm}-Usrdemand.O_{ad}.O_{adn}|$。

b)设变量集 $x=\{x_1,x_2,\cdots,x_k\}$ 中最小值为 x_t,则由 $LSC\_AD.R$ 得到与 $LSC\_AD.O_{ad}.O_{adt}$ 关联的 $LSC\_AD.I_{ad}$ 的元素集 SSI_{ad}。

c)遍历 SI_{ad}。

①假设当前元素为 $S_{ad}.S_{adm}$. 遍历 SSI_{ad};

②假设当前元素为 $SSI_{ad}.SSI_{adn}$,设定集合 $S=\varnothing$,变量集 $x=\{x_1,x_2,\cdots,x_k\}$,其中所有元素值赋值为 $+\infty$。令 $x_n=|S_{ad}.S_{adm}-SSI_{ad}.SSI_{adn}|$;

③设变量集 $x=\{x_1,x_2,\cdots,x_k\}$ 中最小值为 x_t,则把 SSI_{adt} 放入集合 S 中。

d)遍历 S。

①假设当前元素为 S_k,给定集合集 $W=\{W_1,W_2,\cdots,W_m\}$,其中所有集合值赋值为 $\varnothing$,由 $LSC\_AD.R$ 得到与 S_k 关联的 $Wservice$ 的服务集 W_k;

②设集合集 W 中个数最多的 $Wservice$ 为 Ws_i,由 $Wservice.R$ 得到一个服务链接 URL,并放入 $Slinks$ 中。

Step 3 输出 $Slinks$。

由算法 2,可返回用户所需的服务地址。匹配服务过程无需查找本体树,提高了运算效率。

算法 3 服务绑定的动态替换算法

输入:服务簇 $LSC\_AD=(CN,RL,CS,I_{ad},O_{ad},f_{Iad},f_{Oad},R)$;需替换的服务链接地址 $Link_1$。

输出:可替代的链接地址 $Link_2$

Step 1 由链接地址 $Link_1$ 可映射到 Web 服务 $Wservice_1$;

Step 2 由 $Wservice_1.R$,得到 $Link_1$ 对应的输出 O_i;

Step 3 由算法 1 对 O_i 进行量化运算;

Step 4 令 O_i 为算法 2 的输入,得到新的链接地址 $Link_2$;

Step 5 输出 $Link_2$。

算法 3 只需执行 1 次本体树的查找算法,能够快速提供可替换的服务绑定,解决了网络环境变化造成的服务响应失效问题,提高服务响应的自适应性。

4 仿真实验及分析

1)Web 服务的定义

根据实验需要和定义 2 对 Web 服务的形式化描述,自定义 150 个 Web 服务,前 3 项 Web 服务如表 1 所示。

表 1 Web 服务集

Tab. 1 Web service set

Description	*id*	*Ops*	*I*	*O*	*Links*	*R*
成绩查询	1	查询	学号	成绩	1090	Null
天气查询	2	查询	城市	天气	1091	Null
机票预定	3	订票	区间	机票	1092	Null

2)服务簇的生成

从这 150 个 Web 服务中提取输入输出参数,经过参数合并后,得到参数集。输入参数集为{学号;城市;区间;…;班级},输出参数集为{成绩;天气;机票;…;ID 号}。按照文献[9]中的方法建立基于词汇语义的本体树 *Otree*。利用 *Otree* 对所有服务进行标注。按照文献[10]中的方法建立服务簇。由定义 4、算法 1 及定义 5 得到服务簇的 LPN 模型。

3)服务绑定及服务替换

给定用户需求:输入学号,输出成绩。由定义 6 得到用户需求的输入集{学号},输出集{成绩}。由算法 1 量化参数集,得到输入集{23},输出集{106}。由算法 2 进行服务查找,得到绑定的服务链接。此链接失效后,由算法 3 得到可替换的服务链接。

4)比较分析

文献[10]给出的基于 LPN 的服务簇模型,整合了服务参数,把相似度匹配限制在服务簇内。本研究进

化了服务簇模型，通过对服务簇内的参数进行基于本体相似度的量化处理，提高计算效率。

用文献[10]中的方法进行服务查找。根据用户需求，得出输入输出参数集，设参数集元素个数为 k，服务簇内服务参数个数为 m，则需要进行 km 次相似度计算。设一次相似度计算复杂度为 n，则进行一次服务绑定的计算复杂度为 $O(kmn)$，用本文提出的方法进行服务绑定，只需要进行用户需求参数和服务参数的运算，则计算复杂度为 $O(km)$，因此，本文方法减小了服务绑定的时间复杂度。

5 结束语

在文献[10]基于 LPN 的服务簇模型基础上，给出了改进型服务簇的定义。通过对服务簇中参数进行基于相似度计算的量化处理，进化了服务簇构架，在服务绑定中无需再进行基于语义的相似度计算。针对服务失效问题，提出服务绑定的替换方法。通过实验及比较分析验证了上述方法在提高绑定效率及服务响应的自适应性方面的优越性。下一步工作是对本体树的结构进行优化，建立形式化模型，提高在本体树概念查找阶段的效率。

参考文献：

[1]李刚，孙红梅，李智. 资源受限 Web 服务[J]. 软件学报，2010，33(2)：215-220.
Li Gang，Sun Hongmei，Li Zhi. Resource-constrained Web service [J]. Journal of Software，2010，33(2)：215-220.

[2]胡强，杜玉越. 面向服务簇的服务体系结构及服务发现[J]. 计算机应用，2013，33(8)：2163-2166.
Hu Qiang，Du Yuyue. Service architecture and service discovery oriented to service clusters[J]. Journal of Computer Applications，2013，33(8)：2163-2166.

[3]Nayak R，Lee B. Web service discovery with additional semantics and clustering. [C]// IEEE/WIC/ACM International Conference on Web Intelligence. Silicon Valley，Nov. 2-5，2007：555-558.

[4]张景雨，余雪丽，付丰科. 利用聚类优化语义 Web 服务发现[J]. 计算机工程与应用，2009，45(34)：139-143.
Zhang Jingyu，Yu Xueli，Fu Fengke. Using service clusters to facilitate semantic Web service discovery[J]. Computer Engineering and Application，2009，45(34)：139-143.

[5]Sudha R，Thamarai S. Semantic grid service discovery approach using clustering of service ontologies[C]// IEEE/WIC/ACM International Conference on Web Intelligence. Hong Kong，Nov. 14-17，2006：1-4.

[6]孙萍，蒋昌俊. 利用服务聚类优化面向过程模型的语义 Web 服务发现[J]. 计算机学报，2008，31(8)：1340-1353.
Sun Ping，Jiang Changjun. Using service clustering to facilitate process-oriented semantic Web service discovery[J]. Chinese Journal of Computers，2008，31(8)：1340-1353.

[7]Liu X Z，Huang G，Mei H. Discovering homogeneous Web service community in the user-centric Web environment [J]. IEEE Transactions on Service Computing，2009，2(2)：167-181.

[8]Han S，Wang H Y，Cui L Z. A user experience-oriented service discovery method with clustering technology [C]//2nd International Symposium on Computational Intelligence and Design. Changsha，Dec. 12-14，2009：64-67.

[9]吴健，吴朝晖，李莹. 基于本体论和词汇语义相似度的 Web 服务发现[J]. 计算机学报，2005，28(4)：595-602.
Wu Jian，Wu Zhaohui，Li Ying. Web service discovery based on ontology and similarity of words[J]. Chinese Journal of Computers，2005，28(4)：595-602.

[10]邓世阳，杜玉越. 基于逻辑 Petri 网的 Web 服务簇模型[J]. 计算机应用，2012，32(8)：2328-2332.
Deng Shiyang，Du Yuyue. Logic Petri net based model for Web service cluster[J]. Journal of Computer Applications，2012，32(8)：2328-2332.

[11]Du Y Y，Qi L，Zhou M C. A vector matching method for analyzing logic Petri nets[J]. Enterprise Information Systems，2011，5(4)：449-468.

[12]Du Y Y，Jiang C J. Formal representation and analysis of batch stock trading systems by logical Petri net workflows[J]. Lecture Notes in Computer Science，2002，2495：221-225.

[13]Du Y Y，Guo B Q. Logic Petri nets and equivalency[J]. Information Technology Journal，2009，8(1)：95-100.

[14]蒋昌俊. Petri 网的行为理论及其应用[M]. 北京：高等教育出版社，2003：52-95.

（责任编辑：吕文红）