

# 基于概率模型检测的 Web 服务组合验证

王 晶<sup>1</sup> 戎 玫<sup>2</sup> 张广泉<sup>1,3</sup> 祝 义<sup>4</sup>

(苏州大学计算机科学与技术学院 苏州 215006)<sup>1</sup> (暨南大学深圳旅游学院 深圳 518053)<sup>2</sup>

(中国科学院软件研究所计算机科学国家重点实验室 北京 100080)<sup>3</sup>

(徐州师范大学计算机科学与技术学院 徐州 221116)<sup>4</sup>

**摘 要** Web 服务组合验证对提高软件开发效率、实现服务增值具有重要意义。为了验证服务组合的有效性,提出了一种基于概率模型检测的 Web 服务组合验证方法。首先采用扩展的有限自动机模型建立 Web 服务组合模型,并将该模型转换为 Markov 模型,然后采用概率模型检测器 PRISM 验证服务组合的可靠性,最后通过实例进一步说明该方法的可行性。

**关键词** Web 服务组合,有限自动机,Markov 模型,概率模型检测

**中图分类号** TP311 **文献标识码** A

## Validation of Web Service Composition Based on Probabilistic Model Checking

WANG Jing<sup>1</sup> RONG Mei<sup>2</sup> ZHANG Guang-quan<sup>1,3</sup> ZHU Yi<sup>4</sup>

(School of Computer Science and Technology, Soochow University, Suzhou 215006, China)<sup>1</sup>

(Shenzhen Tourism College, Jinan University, Shenzhen 518053, China)<sup>2</sup>

(State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing 100080, China)<sup>3</sup>

(School of Computer Science and Technology, Xuzhou Normal University, Xuzhou 221116, China)<sup>4</sup>

**Abstract** Web service composition validation is very important for improving the efficiency of software development. This paper presented a method based on probabilistic model checking to validate the service composition. We used an extended finite automaton to represent the service composition, and then converted it into a Markov model. A probabilistic model checker PRISM was used to validate the effectiveness of the service composition. Finally, we gave an example to illustrate the feasibility of this method.

**Keywords** Web service composition, Finite automata, Markov model, Probabilistic model checking

## 1 引言

面向服务的计算(Service Oriented Computing, SOC)作为一种新的计算范型,以 Web 服务为基础,能有效地解决在动态、异构环境下的分布式应用集成问题。Web 服务是一种自治的、与平台无关的网络化构件,具有松耦合、支持动态组合和重配置等特点。由于单个 Web 服务功能有限,为了满足实际用户需求,有必要将已有的服务进行组合。服务组合就是通过基本服务之间的相互通信与协作,把相对独立及简单的服务组合成具有新功能的大粒度服务的过程<sup>[1]</sup>。

由于 Web 服务所处的网络环境是一个异构的、分布式自治的动态环境,服务组合在执行过程中可能会受到通信模式的变化、服务基础设施失效等问题的影响。Web 服务组合可能因以下情况导致失败:(1)可用服务运行失败,服务组合执行环境中的服务运行失败所引起的故障使得服务不可调用;(2)服务之间的网络连接出现中断,使得服务间无法正常通

信<sup>[2]</sup>。由于以上等因素往往会导致服务组合失效,使之无法满足用户需求,因此有必要验证服务组合的可靠性。

形式化验证技术是保证软件质量的重要途径之一。其中模型检测<sup>[3]</sup>是一种验证有限状态系统的自动化验证技术,其基本思想是用状态迁移系统表示系统的行为,用时序逻辑公式描述系统的性质。这样“系统是否具有所期望的性质”就转化为数学问题“状态迁移系统是否是公式的一个模型”。概率模型检测<sup>[4]</sup>技术用于计算系统在运行期间某事件发生的可能性,它结合了概率分析和通用模型检测技术,广泛应用于硬件和协议的验证。与通用模型检测相比,概率模型检测所分析的问题是一些典型概率问题。例如多媒体协议中,视频帧在 5ms 内被发送的可能性至少是 89%。

## 2 相关工作

Web 服务组合的验证方法主要分为以下几类:基于自动机、基于 Petri 网和基于进程代数。文献[5]提出一种基于扩

到稿日期:2011-02-16 返修日期:2011-06-21 本文受国家自然科学基金(60973149),江苏省自然科学基金(BK2011281),中国科学院计算机  
科学国家重点实验室开放课题(SYSKF0908),江苏省高校自然科学研究项目(08KJB520010,10KJB520019)资助。

王 晶(1988-),女,硕士生,主要研究方向为软件服务与形式化方法;戎 玫 博士,副教授,主要研究方向为软件工程与网络分布计算;  
张广泉 博士,教授,CCF 高级会员,主要研究方向为服务计算、云计算与 CPS、网络与分布式计算等,E-mail: gqzhang@suda.edu.cn(通信作  
者);祝 义 男,博士,副教授,主要研究方向为软件工程与形式化方法。

展有限自动机的方法验证组合 Web 服务,它不仅可以验证组合服务是否满足系统需求,还可以验证组合服务运行过程中是否有逻辑错误。文献[6]提出一种 Web 服务组合的正确性验证方法,即采用软件体系结构描述语言 XYZ/ADL 描述 Web 服务组合,其性质用 CTL 公式表示,最后用模型检测工具 UPPAAL 验证 Web 服务组合的正确性。文献[7]提出一种基于 Petri 网理论的 Web 服务组合建模方法,即对服务组合进行形式化建模,通过分析 Web 服务网的可达性和活性验证 Web 服务组合。文献[8]利用 Pi-演算对 Web 服务组合建立形式化模型,定义了 Pi-演算到 BPEL4WS 的概念映射,给出了基于 Pi-演算的形式化描述。虽然 Petri 网、进程代数的表达能力比有限自动机强,但其计算的空间复杂度比有限自动机高。

已有的工作大都集中于采用某种形式化方法对 Web 服务组合建模,仅考虑服务组合的绝对正确性,并没有考虑服务调用失败的情况。本文考虑因调用单个服务失败而导致组合服务失败的情况,将单个服务的可靠性引入到服务组合模型中。由于在模型中添加了概率信息,需对有限自动机模型进行扩展并转换为概率模型,相应地采用概率模型检测技术验证性质。本文提出了一种基于概率模型检测的 Web 服务组合验证方法,即采用扩展的有限自动机模型表示服务组合,并将该模型转换为 Markov 模型,然后采用概率模型检测器 PRISM<sup>[9]</sup> 验证服务组合的可靠性。最后通过实例说明了本方法的可行性。

### 3 Web 服务组合建模与转换

采用扩展的有限自动机对组合服务建模,在状态和状态间迁移添加概率表示服务执行的可靠性。可靠性是指系统在规定条件下完成规定功能的能力,用 $[0,1]$ 间的实数表示调用单个服务的可靠性。

**定义 1** 扩展的有限自动机模型  $M=(S, s_0, F, \Sigma, R)$ , 其中:

- $S$  是有穷状态集合,  $(s, \alpha) \in S$ , 其中  $s$  表示状态,  $\alpha \in [0, 1]$  表示可靠性;
- $s_0 \in S$  是初始状态;
- $F \subseteq S$  是终止状态集合;
- $\Sigma$  是字母表;
- $R \subseteq S \times P \times \Sigma \times S$  是状态间迁移关系,  $P$  表示状态间转移的概率。

将自动机中节点分为分支节点和非分支节点。分支节点表示该节点有多个后继节点,非分支节点表示该节点只有一个后继节点。非分支节点用  $(s, \alpha)$  表示,其中  $s$  表示状态,  $\alpha$  表示可靠性。分支节点(包括条件分支和并行分支)用  $s$  表示。设有  $n$  条分支,选择每条分支的概率为  $p_i (i=1, \dots, n)$ , 若属于条件分支,则  $\sum_{i=1}^n p_i = 1$ ; 若属于并行分支,则  $p_i = 1 (i=1 \dots n)$ 。

扩展的有限自动机到 Markov 模型的转换,是通过遍历扩展有限自动机模型,根据组合服务模型的执行路径进行转换。根据不同的模型特征转换为不同的 Markov 模型。若模型中没有时间信息,也没有并行分支,则用离散时间 Markov 链(Discrete-time Markov Chain, DTMC)作为分析模型;若模型中没有时间信息但有并行分支,则用马尔可夫决策过程

(Markov Decision Process, MDP)作为分析模型。

扩展的有限自动机  $M$  的初始状态对应 Markov 模型的初始状态,Markov 模型的终止状态是  $M$  的终止状态与表示服务调用失败的 fail 状态的并集。 $M$  中非分支节点对应 Markov 模型中带两条出边的节点,一条表示执行成功转到  $M$  中下一个状态,另一条表示执行失败转到 fail 状态; $M$  中分支节点对应 Markov 模型中一个节点,并在分支上添加相应的概率信息。

**定义 2**(扩展的有限自动机  $M$  到 Markov 模型的转换)

若  $M$  中状态  $s$  属于非分支节点,则  $s$  对应 Markov 模型中状态  $s'$ 。  $s'$  有两条出边,一条表示服务执行成功转移到下一个状态  $s''$ ,选择该边的概率为  $\alpha$ ; 另一条表示执行失败转移到 fail 状态,选择该边的概率为  $1-\alpha$ 。若  $M$  中状态  $s$  属于分支节点,有后继状态  $s_i (i=1, \dots, n)$ , 则  $s$  对应 Markov 模型中状态  $s', s'$  有后继状态  $s'_i (i=1, \dots, n)$ , 若属于条件分支,则  $\sum_{i=1}^n P(s', s'_i) = 1$ ; 若属于并行分支,则  $P(s', s'_i) = 1 (i=1 \dots n)$ 。

Web 服务组合由单个 Web 服务通过基本组合结构组合而成。这些基本组合结构包括顺序、选择、循环、并行等。根据定义 2, 将基本组合结构的扩展有限自动机模型转换为对应 Markov 模型。

(1) 顺序: 顺序组合是指单个 Web 服务按业务逻辑依次执行,只有等前一个 Web 服务完成后方可执行下一个服务。顺序结构如图 1 所示,组合服务 WS 由 WS1 和 WS2 组成。

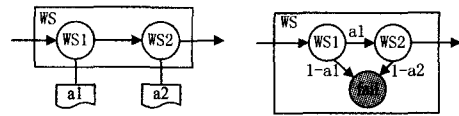


图 1 顺序组合及对应 Markov 模型

(2) 选择: 选择组合是指根据选择条件选择满足条件的 Web 服务分支,执行该 Web 服务分支(见图 2)。

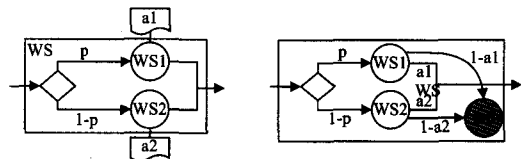


图 2 选择组合及对应 Markov 模型

(3) 循环: 满足条件的前提下重复执行一个或多个 Web 服务。转换为 Markov 模型时,将循环路径展开,转换方法类似顺序组合(见图 3)。

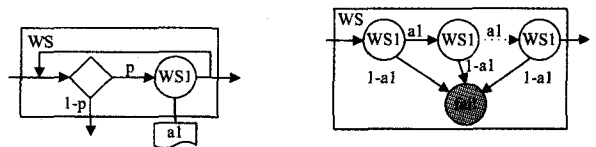


图 3 循环组合及对应 Markov 模型

(4) 并行: 同时执行两个或两个以上 Web 服务,这些服务是并行执行的(见图 4)。

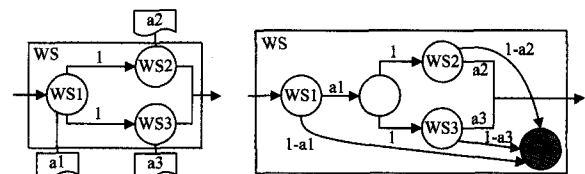


图 4 并行组合及对应 Markov 模型

## 4 Web 服务组合的验证

概率模型检测技术结合了概率分析和通用模型检测技术,目前较为成功的概率模型检测器 PRISM 由英国伯明翰大学开发。针对不同的系统特征,有以下不同的概率模型:离散时间 Markov 链 DTMC,用于表示概率选择;连续时间 Markov 链(Continuous-time Markov Chain, CTMC),用于表示连续时间和概率选择;马尔可夫决策过程 MDP 用于表示概率和不确定选择。概率性质规约包括概率计算树逻辑 PCTL<sup>[10]</sup>和连续随机逻辑 CSL。

### 4.1 性质描述

采用 PCTL 描述组合服务待验证的性质。

定义 3 PCTL 的语法定义如下:

- 状态公式  $\varphi ::= \text{true} \mid a \mid \varphi \wedge \varphi \mid \neg \varphi \mid P_{\sim p}[\psi]$
- 路径公式  $\psi ::= X\varphi \mid \varphi U \varphi \mid \varphi U^{\leq t} \varphi$

式中,  $a$  是原子命题,  $\sim \in \{<, \leq, >, \geq\}$  是一个关系操作符,  $p \in [0, 1]$  是概率,  $t \in \mathbb{N}$ 。

PCTL 的语义可用满足关系  $\models$  表示。对于 Markov 模型中状态  $s$ , 状态公式  $\varphi$ ,  $s \models \varphi$  表示  $\varphi$  在状态  $s$  为真。对于 Markov 模型中路径  $\omega$ , 路径公式  $\psi$ ,  $\omega \models \psi$  表示  $\psi$  在路径  $\omega$  上为真。

定义 4 PCTL 的语义定义如下:

- $s \models a \Leftrightarrow a \in L(s)$ ;
- $s \models \varphi_1 \wedge \varphi_2 \Leftrightarrow s \models \varphi_1$  且  $s \models \varphi_2$ ;
- $s \models \neg \varphi \Leftrightarrow s \not\models \varphi$ ;
- $s \models P_{\sim p}[\psi] \Leftrightarrow \text{Prob}(s, \psi) \sim p$ , 其中  $\text{Prob}(s, \psi) = \text{Pr}_s\{\omega \in \text{Path}_s \mid \omega \models \psi\}$ ;
- $\omega \models X\varphi \Leftrightarrow$  路径  $\omega$  的下一个状态满足  $\varphi$ ;
- $\omega \models \varphi_1 U \varphi_2 \Leftrightarrow \exists k \geq 0$ , 使得  $\omega(k) \models \varphi_2$  且对  $\forall i < k$ , 有  $\omega(i) \models \varphi_1$ ;
- $\omega \models \varphi_1 U^{\leq t} \varphi_2 \Leftrightarrow \exists i \leq t$ , 使得  $\omega(i) \models \varphi_2$  且对  $\forall j < i$ , 有  $\omega(j) \models \varphi_1$ 。

### 4.2 Web 服务组合验证

验证组合服务的可靠性,主要关注两种性质:(1)  $P_{\sim p}[\text{true} U (\text{system\_state} = \text{success})]$ , 即从初始状态到达组合成功状态的概率是否满足阈值  $p$ ;(2)  $\text{system\_state} = \text{"a certain service incocation"} \Rightarrow P_{\sim p}[\text{true} U (\text{system\_state} = \text{success})]$ , 即从服务组合中某个状态到达成功状态的概率是否满足阈值  $p$ 。

根据上述步骤,将转换得到的 Markov 模型和性质规约输入概率模型检测器 PRISM,验证组合服务是否满足性质。若满足,则继续验证下一个组合服务;否则,替换服务,直到组合服务满足性质。

## 5 实例分析

Web 服务组合“旅游代理服务”由服务序列  $[WS_0, WS_1, WS_2, WS_3, WS_4, WS_5, WS_6, WS_7]$  组成。旅游开始时,执行订票、订旅馆服务,同时可搜索旅游地点,计算景点到旅馆的距离。若距离大于某个值,则乘出租车去景点,否则租用自行车。图 5 为旅游代理服务的扩展有限自动机模型。

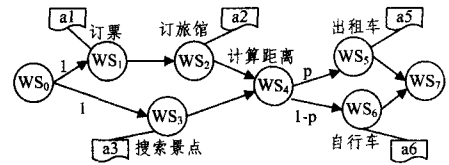


图 5 组合服务“旅游代理服务”

通过统计的方法,得到非分支节点服务执行成功的概率以及分支节点选择每条分支的概率。具体如表 1 所列。

表 1 状态及迁移的概率信息

	值		值		值
a1	0.96	a3	0.98	a5	0.99
a2	0.97	p	0.5	a6	0.95

按照定义 2,将扩展有限自动机模型转换为 Markov 模型,如图 6 所示。

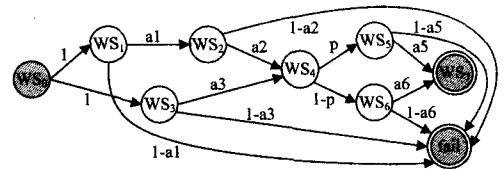


图 6 “旅游代理服务”的 Markov 模型

转换得到的 Markov 模型可分为两种:DTMC 和 MDP。其中,DTMC 不考虑时间和并行情况;MDP 不考虑时间但考虑并行情况。DTMC 将模型中并行部分压缩为一个节点,DTMC 可用于评价整个服务组合的可靠性或部分服务组合的可靠性。

状态	PCTL 公式	结果
WS0	$P[\text{true} U (\text{system\_state} = \text{success})]$	0.927
WS4	$WS_4 \Rightarrow P[\text{true} U (\text{system\_state} = \text{success})]$	0.97

MDP 可用于计算组合模型中并行部分的可靠性。在 DTMC 中将并行部分压缩成一节点,不能表示并行部分内部节点的可靠性。

状态	PCTL 公式	结果
WS1	$WS_1 \Rightarrow P_{\max}[\text{true} U (\text{system\_state} = \text{success})]$	0.951
WS1	$WS_1 \Rightarrow P_{\min}[\text{true} U (\text{system\_state} = \text{success})]$	0.903

对组合服务“旅游代理服务”,服务组合到达成功状态的概率大于 90%,则可认为服务组合是成功的;到达 fail 状态的概率大于 10%,则可认为服务组合是失败的,需对服务进行替换。由图 7 可知,到达成功状态的概率为  $0.927 > 90\%$ ,到达 fail 状态的概率小于 10%,服务组合满足用户需求。

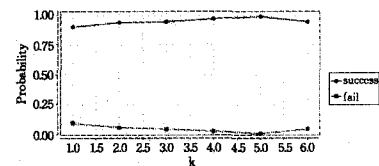


图 7 每个状态到达成功状态和 fail 状态的概率

**结束语** 随着 Internet 的快速发展,Web 服务以其良好的连接异构平台能力和实现互操作的优势受到工业界和学术界的共同关注。由于单个 Web 服务提供的功能有限,需通过服务组合来实现服务的增值。而 Web 服务所处环境的异构性、动态性等特征可能导致服务组合失败,Web 服务组合的可靠性验证对提高软件开发效率具有重要意义。本文提出了

一种基于概率模型检测的 Web 服务组合验证方法,即采用扩展的有限自动机模型表示服务组合,并将其转换为 Markov 模型。然后采用概率模型检测器 PRISM 验证服务组合;最后通过实例进一步说明了方法的可行性。本文所采用的 Markov 模型不能表示时间信息。下一步工作将考虑扩展分析模型,使之能表示服务组合中的时间信息,并进行实验验证。

## 参 考 文 献

- [1] 范贵生,虞慧群,陈丽琼,等. 基于 Petri 网的服务组合故障诊断与处理[J]. 软件学报,2010,21(2):231-243
- [2] Pleisch S, Schiper A. Approaches to fault-tolerant and transactional mobile agent execution—An algorithmic view[J]. ACM Computing Surveys,2004,36(3):219-262
- [3] 林惠民,张文辉. 模型检测:理论、方法与应用[J]. 电子学报,2002,30(12A):1907-1912
- [4] Kwiatkowska M, Norman G, Parker D. Advances and Challenges

(上接第 100 页)

测试结果表明,基于关键字的服务发现方法的查准率和查全率都较低;本文方法的查全率达到 80%左右,查准率达到 90%以上。因此,利用本文方法,能够提高需求和服务之间的语义匹配精度,缩小搜索范围,能够找到最符合用户需求的、高质量的服务资源来构造网络化软件。

**结束语** 本文提出一种网络化软件的按需服务发现方法。首先构建了需求和服务的 3 类领域本体,使得领域内的需求和服务的相关概念统一起来,实现了知识共享,为网络化软件的需求和服务的匹配提供了语义支持;接着重点讲述了基于本体图的概念语义相似度度量与分析,并以此为基础,建立了按需服务“基本+功能+质量”三阶段语义匹配模型;最后阐述了网络化软件的按需服务发现方法,并以物流领域的具体实例进行了性能测试。下一步工作将集中于探讨建立结构更加合理、知识更加丰富的领域本体,同时考虑如何将本文所提方法和模型应用于实际环境中。

## 参 考 文 献

- [1] He Ke-qing, Wang Jian, Liang Peng. Semantic Interoperability Aggregation in Service Requirements Refinement [J]. Journal of Computer Science and Technology,2010,25(6):1103-1117
- [2] 何克清,彭蓉,李兵,等. 网络化软件[M]. 北京:科学出版社,2008
- [3] 吴朝晖,邓水光,吴健. 服务计算与技术[M]. 杭州:浙江大学出版社,2009
- [4] Paolucci M, Kawamura T, et al. Semantic Matching of Web Services Capabilities[C]//Proceedings of the 1st International Semantic Web Conference. 2002:333-347
- [5] Maximilien E M, Singh M P. A framework and ontology for dynamic Web services Selection [J]. Internet Computing,2004,8(5):84-93
- [6] Soydan B A, Singh Munindar P. A DAML-based repository for QoS-aware semantic Web service selection[C]//Proceedings of the IEEE International Conference on Web Services, California, USA,2004:368-375
- [7] El Haddad J, Manouvrier M, Rukoz M. TQoS: Transactional and

of Probabilistic Model Checking [C]//Communication Control and Computing. Allerton, IEEE press, 2010:1691-1698

- [5] 雷丽晖,段振华. 一种基于扩展有限自动机验证组合 Web 服务的方法[J]. 软件学报,2007,18(12):2980-2990
- [6] 何亚丽,戎玫,张广泉. 一种基于 UPPAAL 的 Web 服务组合模型检测方法[J]. 计算机科学,2010,37(11):122-125
- [7] 陈丁剑,吴健,马满福,等. 基于 Petri 网的 Web 服务组合建模[J]. 计算机科学,2006,33(5):128-130
- [8] 辜希武,卢正鼎. 基于 Pi-演算的 BPEL4WS Web 服务组合形式化模型[J]. 计算机科学,2007,34(3):69-74
- [9] Kwiatkowska M, Norman G, Parker D. PRISM 2. 0: a tool for probabilistic model checking[C]//International Conference on the Quantitative Evaluation of Systems (QEST), Enschede, The Netherlands; IEEE Computer Society. 2004:27-30
- [10] Hansson H, Jonsson B. A logic for reasoning about time and reliability [J]. Formal Aspects of Computing,1994,6(5):512-535

QoS-aware Selection Algorithm for Automatic Web Service Composition [J]. IEEE Transactions on Services Computing, 2010,3(1):73-85

- [8] 吴健,吴朝晖. 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
- [9] 杨放春,苏森,李祯. 混合 QoS 模型感知的语义 Web 服务组合策略[J]. 中国科学(E辑:信息科学),2008,38(10):1697-1716
- [10] 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
- [11] Web Service Challenge [EB/OL]. <http://ws-challenge.georgetown.edu/wsc09/>
- [12] Jiang W, Zhang C, Huang Z, et al. Qsynth: A tool for QoS-aware automatic service composition[C]//Proceedings of the IEEE International Conference on Web Services. Miami, FL, USA, 2010
- [13] Bartalos P, Bielikovaa M. Semantic Web service composition framework based on parallel processing[C]//Proceedings of the 11th IEEE Conference on Commerce and Enterprise Computing. Vienna, Austria, 2009
- [14] 孙海霞,钱庆,成颖. 基于本体的语义相似度计算方法研究综述[J]. 现代图书情报技术,2010(1):51-56
- [15] Antoniou G, Van Harmelen F. A Semantic Web Primer [M]. The Mit Press, 2004:200-208
- [16] Lord P W, Stevens R D, Brass A, et al. Investigating Semantic Similarity Measures Across the Gene Ontology: The Relationship Between Sequence and Annotation [J]. Bioinformatics, 2003,19(10):1275-1283
- [17] 潘谦红,王炬,史忠植. 基于属性论的文本相似度计算[J]. 计算机学报,1999,22(6):651-655
- [18] Knappe R, Bulskov H, Andreassen T. On Similarity Measures for Content-based Querying[C]//Proceedings of the 10th International Fuzzy Systems Association World Congress. 2003:400-403
- [19] 喻坚,韩燕波. 面向服务的计算—原理和应用[M]. 北京:清华大学出版社,2006
- [20] 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