

基于细胞因子网络协同进化的 Web 服务合成方法

刘再群^{1,3} 丁永生^{1,2} 胡志华¹

(东华大学信息科学与技术学院 上海 201620)¹

(东华大学数字化纺织服装技术教育部工程研究中心 上海 201620)²

(安徽师范大学生命科学学院 芜湖 241000)³

摘要 借鉴生物神经内分泌免疫系统的协同进化机制,基于细胞因子网络平台提出了 Web 服务的合成方法。在细胞因子网络调控下的生物实体代理 Web 服务,构成一个带有条件的米兰机单元,Web 服务的合成可以转换为米兰机进化过程。生物实体通过消息匹配和条件约束形成细胞因子网络,Web 服务的合成通过生物实体细胞因子网络支持。在服务协同进化过程中可以动态调整其合成的服务,完成服务的动态自组织合成和管理等工作。仿真结果表明该方法在环境动态变化时具有适应性。

关键词 生物网络,细胞因子网络,协同进化,米兰机,Web 服务

中图分类号 TP18,TP393.01 **文献标识码** A

Approach of Web Services Composition Based on Co-evolutionary Cytokine Network

LIU Zai-qun^{1,3} DING Yong-sheng^{1,2} HU Zhi-hua¹

(College of Information Sciences and Technology, Donghua University, Shanghai 201620, China)¹

(Engineering Research Center of Digitized Textile & Fashion Technology, Ministry of Education, Donghua University, Shanghai 201620, China)²

(College of Life Sciences, Anhui Normal University, Wuhu 241000, China)³

Abstract Inspired by co-evolutionary mechanism in neural-immune-endocrine system, a cytokine network platform was proposed for Web services composition. In the control of cytokine network, bio-entity delegates Web service to construct a Melay evolutionary unit with conditions. Web composition is then transferred to Melay evolutionary process. Bio-entities construct cytokine network through message matching and conditional constraints to support Web services composition. During the co-evolution process of Web services, the composed services are dynamically adjusted to finish the dynamical composition and management of Web services. The simulation results show that the approach is adaptive in dynamic environments.

Keywords Bio-network, Cytokine network, Co-evolution, Melay machine, Web service

1 引言

Web 服务是一种自描述、平台无关的计算元素,正在受到广泛的关注。它真正的潜力在于可以对简单服务进行合成,但因其复杂性,存在着大量的更新、移动等动态变化及行为交互,多数已有的研究并没有完全解决这个问题^[1,2]。语义,特别是本体的提出,为 Web 服务的自动发现、合成提供了一种新的思想。在语义服务的方式下,服务的发现、执行和合成不需要由人来完成,而是可以由 Agent 自动完成用户提出的复杂任务要求。

下一代互联网和服务计算框架^[1]研究指出:未来的服务存在着服务访问控制的分布性、自动的控制策略、服务变化的

自适应、自进化等主要特征。利用 Agent 自主性设计符合这种特征的服务合成方法正在受到广泛的重视^[3],但存在着 Agent 的通信、协商、产生、消亡、迁移、稳定性等问题。对 Web 服务合成如何实现新的管理框架,是当前要解决的问题。服务管理和合成的一个有益借鉴来自同为复杂巨系统的生物组织系统。通过细胞因子和受体组成生物网络通用语言,免疫系统和神经内分泌系统构成了自组织、自协作的分布服务网络^[4]。

细胞因子网络表现出的重要特性有自组织、聚集现象和协作行为。受细胞因子网络协同进化机制的启发^[5],本文给出 Web 服务合成问题基于带条件米兰机的全局协同进化模型。在生物网络计算框架^[2,6-8]的基础上,将生物实体(带有生

到稿日期:2008-10-13 返修日期:2008-11-26 本文受国家自然科学基金重点项目(60534020),国家自然科学基金(70701009),教育部科技创新工程重大项目培育资金项目(706024),上海市国际科技合作基金项目(061307041),上海市人才发展资金,上海市领军人才后备人选专项资金资助。

刘再群(1975-),男,博士研究生,副教授,从事生物系统建模、生物网络结构等研究,E-mail:liuzaiqun@126.com;丁永生(1967-),男,博士,教授,博士生导师,从事智能系统、网络智能、DNA 计算、人工免疫系统、生物网络结构、生物信息学、数字化纺织服装、智能决策与分析等研究;胡志华(1977-),男,博士研究生,从事免疫计算、分布式计算等研究。

物功能设计的移动 Agent)作为自动机单元,通过消息匹配能力(亲和度,即由服务需求的匹配强度、服务质量相对评定和服务信任因素组成的综合评定值)和条件约束,建立满足用户要求的服务合成进化过程,建立实体网络,完成服务的合成。仿真表明这种方法具有良好的性能和对环境的适应性。

2 Web 服务合成的米兰机模型

Web 服务是一个有状态的动作序列,可以看作一个七元组,如式(1)所示:

$$B=(U, Y, Q, O, \delta, F, QC) \quad (1)$$

其中, U 是一系列 Web 服务输入消息的集合; Y 是一系列 Web 服务输出消息的集合; Q 是一系列 Web 服务的有限状态集合(状态表示 Web 服务交互序列中的历史记录和条件判断记录); O 是起始状态; $\delta: Q \times U \times Y \times QC \rightarrow Q$ 是一个转移函数,即从给定状态根据输入/输出匹配和条件转移到另一个状态, $U \times Y$ 是字母表; $F \subset Q$ 是一系列服务的接受有限状态集合,即用户能够与 Web 服务结束交互的状态集合; $QC: Q \times C$ 是一系列条件的集合。

在语义服务环境中, Web 服务信息描述采用 OWL-S 本体语言,服务的内容由 Service profile 类功能属性和非功能属性两部分组成,实现了功能和属性、质量参数的描述。服务的合成可以转换为语义支持下的 Web 服务米兰机演化过程。用户的 Web 服务请求可以描述为一个抽象 OWL-S 定义形式。OWL-S 的过程模型包括 sequence, split, Split + join, choice, unordered, if-then-else, repeat-while 和 repeat-until 结构。这些结构可以映射到自动机的过程,即可以通过自动机的演化满足用户的服务需求^[9]。Web 服务的合成是一种基于消息的会话及匹配关联的过程,服务合成是封闭的(合成的服务仍然是一个米兰机),可以通过消息会话完成更高层次的合成。

3 基于细胞因子网络的 Web 服务合成方法

在生物网络^[6-8]平台上通过引入细胞因子网络服务层建立细胞因子网络平台,生物实体之间通过细胞因子网络连为一体。借鉴米兰机的模型,生物实体代理 Web 服务,可以描述为米兰机。每个生物实体有消息匹配接口,接口中含有 Web 服务的相关信息,同时接收用户服务请求进行匹配。它通过特定演化规则完成 Web 服务的合成方法,实现了自组织和自进化的功能。按照这种方式,合成模式转换为生物实体之间的匹配和动态管理问题。

3.1 细胞因子协同进化网络平台

如图 1 所示,细胞因子网络建立在生物网络平台之上。在生物网络平台之上,细胞因子网络由两层构成:一层是细胞因子网络模拟器,它提供对细胞因子网络的基本服务的支持,包括监控、能量管理、生命周期管理、细胞因子网络实体近邻关系管理、定时服务和系统管理等等,这些服务在生物网络之上对生物实体进行管理。该层可以被认为是通用的细胞因子网络服务层。之上的另外一层是细胞因子网络的机制设计层,重点体现基于细胞因子网络的神经系统、内分泌系统和免疫系统的协同进化机制的设计。这些机制中涉及的实体,映射为生物网络中的生物实体,是生物实体在细胞因子网络下的扩展。显然,细胞因子网络提供了在这些实体之间建立协

同进化机制的平台,并且这些机制是受到生物体神经-内分泌-免疫 3 大系统的机制的启发。本文研究基于细胞因子网络的 Web 服务合成。这种合成关系向下映射为细胞因子网络;而 Web 服务则依次映射为细胞因子网络中的实体(免疫细胞、神经细胞、内分泌细胞)和生物网络中的生物实体。在生物网络中,强调生物实体的独立能力,以及生物实体之间的通信。而在细胞因子网络中,生物实体具有免疫、神经或内分泌细胞的特征,并且通过细胞因子网络协同进化功能连为一体。

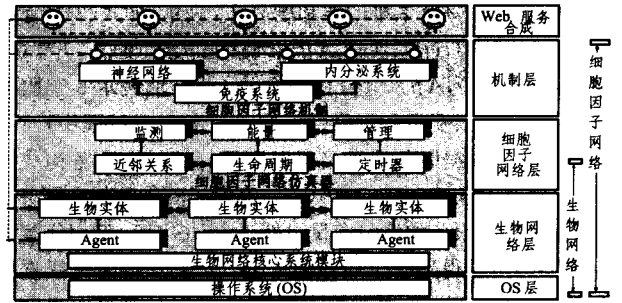


图 1 基于生物网络的细胞因子网络平台

3.2 细胞因子生物实体的结构

生物实体是一种自治的移动 Agent,用于代理 Web 服务与完成合成。生物实体由属性、行为和功能组成。属性描述关于生物实体的信息,包括全局标志、代理服务状态、能量、过程状态等标志;行为实现相关操作,包括能量交换、生物操作(迁移、死亡、复制/再生等)、信息发布、数据接口操作;功能实现生物实体代理的服务运算,包括与服务相关的处理操作。细胞因子生物实体是细胞因子网络应用中的最小组件,是一个能量驱动的服务单元。它可以接收服务请求消息,通过消息匹配方法完成服务序列演化构建。细胞因子生物实体在细胞因子网络环境中进行了扩展,在属性、功能和行为上增加了细胞因子网络的协同进化状态属性、网络关系管理功能,以及通过细胞因子网络进行促进和抑制的行为能力。

3.3 生物实体的消息匹配与算法

生物实体代理 Web 服务,它能观察其他实体的输入、输出动作(即接收到的消息)。在观察的基础上,根据用户需求的描述,通过匹配算法驱动其进化,构造一个可接受的协作行为序列。对于实体间的消息匹配和规则,本文采用了多元综合评估的方式,即对 3 个主要消息匹配的影响因素进行综合度量,根据亲和度的强度选择亲和度匹配最大的实体组成网络。亲和度是衡量免疫组件之间匹配程度的测度。服务合成是一系列实体通过最强匹配协同合成的群体。亲和度测量的影响因素有结构性匹配(服务消息的匹配)强度以及 2 个非结构性匹配,即服务质量 $QoS^{[10]}$ 和信任关系度量。服务匹配强度的计算是一个综合评定指标。下面给出 3 个参数及亲和度的计算方法。在细胞因子网络中,网络节点由扩展的生物实体构成,而生物实体之间则体现为多值的因子关系。细胞因子网络定义为 $CN=(E, R)$,其中 E 是生物实体集合, $R \subset E \times E$ 是生物实体之间的因子联系,本文中定义为多值集合。

a) 服务消息的匹配是服务接口的结构性匹配,表示为 $Intface(e_i, e_j)$,其中 $e_i, e_j \in E$ 。如果在 Web 服务的语义上,两者在接口上构成匹配,则 $Intface(e_i, e_j)=1$,否则 $Intface(e_i, e_j)=0$ 。对于一个消息序列 $WM=(C_1, C_2, \dots, C_n)$,其中

C_i 是服务接口的描述。通过服务的结构性接口匹配,消息序列 WM 可以转化为服务的接口序列,定义为 WS。WS = $(e_1, e_2, \dots, e_{n-1})$, 其中 $Interface(e_i, e_{i+1}) = 1, 1 < i < n-2$ 。显然, $Interface(e_i, e_{i+1})$ 由消息序列中的接口 C_i 和 C_{i+1} 决定。而在细胞因子网络中,这一部分生物实体之间的亲和度强度 ($Aff_{e \sim e}(e_i, e_j)$) 定义为 1, 即

$$(Aff_{e \sim e}(e_i, e_j) = 1) \rightarrow (Interface(e_i, e_j) = 1)$$

b) 服务质量(QoS)是对 Web 服务质量的相对度量。在细胞因子网络中,Web 服务质量映射为生物实体之间协作的服务度量,通过一个多维的亲和度 Aff_Q 表示,如式(2)所示:

$$Aff_Q(e) = (Aff_{Q.t}(e), Aff_{Q.c}(e), Aff_{Q.a}(e)) \quad (2)$$

其中, $Aff_{Q.t}(e)$, $Aff_{Q.c}(e)$ 和 $Aff_{Q.a}(e)$ 分别表示 3 个服务参数: 服务时间、费用和可靠性。假设服务合成过程中,对于服务接口 C_i 根据服务消息接口匹配得到的候选服务构成一个服务集合 $WS_i = \{e_{i1}, e_{i2}, \dots, e_{ik}\}$, 而对于其中的每个服务都可以得到一个服务的质量向量 $Aff_Q(e), e \in WS_i$, 则对于候选服务集合,可以得到一个质量矩阵 Q , Q_j 表示第 j 个候选服务,而 Q_{jh} 表示第 j 个候选服务的第 h 个服务质量向量的分量。通过 SAW^[11] 方法可以将决定服务质量向量转换为相对的标量数值。由于服务时间和费用属于否定性属性,使用式(3)量化。而可靠性属于肯定性属性,通过式(4)量化。

$$Aff_{Q(i,j)} = \begin{cases} \frac{Q_j^{\max} - Q_{jh}}{Q_j^{\max} - Q_j^{\min}}, & Q_j^{\max} - Q_j^{\min} \neq 0 \\ 1, & Q_j^{\max} - Q_j^{\min} = 0 \end{cases} \quad (3)$$

$$Aff_{Q(i,j)} = \begin{cases} -\frac{Q_j^{\max} - Q_{jh}}{Q_j^{\max} - Q_j^{\min}}, & Q_j^{\max} - Q_j^{\min} \neq 0 \\ 1, & Q_j^{\max} - Q_j^{\min} = 0 \end{cases} \quad (4)$$

其中, $Q_j^{\max} = \max_i(Q_{ij})$, $Q_j^{\min} = \min_i(Q_{ij})$ 。从而,可以得到服务的相对矩阵 $P = [P_{ij}]$ 。另外,对于 3 个服务质量向量的分量,定义相应的权重向量 $W = (w_1, w_2, w_3)$, $\sum_{h=1}^3 w_h = 1$ 。在本文中,假定对于所有服务,使用相同的权重向量。对于服务 e_i , 质量的相对评价值为 $Aff_{|Q|}(e_i)$, 根据式(5)计算。 $Aff_{|Q|}(e_i)$ 越大,说明其服务质量越好,在当前服务选择决策中,它被选择的可能性就越大。

$$Aff_{|Q|}(e_i) = \sum_{h=1}^3 P_{ih} w_h \quad (5)$$

c) 服务信任的度量。信任机制建立生物实体之间的一种面向安全连接的亲和度测度。它通过建立和改变生物服务实体之间的信任值来更新和加强亲和度网络。服务合成完成后,使用者根据服务过程返回一个支付(defray)消息,该消息包括一个协作记录、一个表示用户对请求命中的评价价值 R 。 R 值可以是一个奖励,也可以是一个惩罚,它表示用户对接收到的请求命中的偏爱程度。这个消息沿着开始传递的路径传播,可将它调节用来传递初始发现请求关系的信任值。作为奖励时,信任值增加;作为惩罚时,信任值减少。信任通过生物实体之间的信任亲和度表示,对生物实体 e_i 的信任亲和度记为 $Aff_t(e_i)$, 通过式(6)计算:

$$Aff_t(e_i) = Aff_t(e_i) + \Delta$$

$$\Delta = \begin{cases} (1 - Aff_t(e_i)) \times R^2, & R \geq 0 \\ Aff_t(e_i) \left(1 - \frac{1}{1+R}\right), & R < 0 \end{cases}$$

$$R \in [-0.5, 1] \quad (6)$$

通过式(6)计算的 $Aff_t(e_i)$ 是属于 $(0, 1]$ 的一个数值。 Δ

计算信任的调整值。考虑到评价的有效性和公正性,并反映主观评价的特征,式(6)对 Δ 做了优化处理, $Aff_t(e_i)$ 下降很快,而上升很慢。

d) 综合 3 个方面的参数因素,表示消息的匹配能力通过细胞因子网络中生物实体之间的多亲和度计算,如式(7)所示:

$$Aff(e_i) = \begin{cases} (1-\alpha)Aff_{|Q|}(e_i) + \alpha Aff_t(e_i), & e_i \in WS \\ 0, & e_i \notin WS \end{cases} \quad (7)$$

其中, $\alpha \in [0, 1]$ 是影响亲和度的权重系数。为了加速服务的选择,系统设定亲和度阈值 $\theta \in [0, 1]$, 如果 $Aff(e_i) \geq \theta$, 则认为服务符合匹配的条件。如果符合条件的服务有多个,一种是根据最大匹配原则选择,这样涉及排序并需要获得整个列表;另一种是满意原则,一旦发现符合条件的服务,则选择它。本文采用生物服务实体的亲和度最大匹配原则,选取亲和度值最大实体对象建立一个有序的服务合成序列。但服务质量是否满足用户的最初要求,需要对整个序列进行评定,如合成后新服务的综合费用和响应时间等。实体会累加统计性能指标,最后根据序列结构完成突现的总体性能计算。通过上面的论述可知,Web 服务合成是一系列实体通过消息的亲和度匹配演化的群体。针对用户的不同服务请求,系统通过亲和度算法动态合成实体,并根据环境的变化调整生物实体,过程处于一种动态的自进化状态。

4 仿真结果及性能分析

4.1 仿真设计

在基于细胞因子生物网络平台上进行了 Web 服务合成方法的仿真。该平台设计了 Java 编写的软件接口、公用结构体、仿真器,并支持即插即用和灵活的操作,仿真对原生物网络平台增加了服务数据接口处理和亲和度计算功能,实现了 Web 服务合成的仿真操作。同时增加细胞因子网络机制层,支持细胞因子网络的生物实体多亲和度的计算。由于没有标准的测试数据集,本文采用随机生成的 Web 服务之间的消息匹配度数据集和服务质量的数据集。同时为了简化运算,对 $Aff_t(e_i)$ 做如下规定:初值规定为 1, 否则为 0。本文生成了包含 1000 个不同资源的多数据集,随机抽取了 100, 200, 300, 400, 500 个服务来评价方法性能。将生成的测试数据分配给生物实体,用来测试系统的服务合成特性。仿真主要讨论这种自进化方法的性能、自适应性和抗毁性。

本仿真实验采用了 Pentium 4 处理器(2.4GHz 和 512M RAM)的计算机,实验设置具有 432 个(18×24)节点的网络结构。假设每个虚拟机的网络节点上都运行有仿真平台,网络平台上的资源花费能量相同。访问的服务请求频率不随时间变化,设定用户的网络服务需求是 10 次/s。系统设置生物实体能量为 10000, 节点延迟为 0.01s, 仿真时间为 100min。

仿真中对自进化 Web 服务突现性能的影响因素有多个方面,这里主要以消息匹配的亲和度权重系数 $\alpha \in \{0, 0.5, 0.8, 1\}$ 和 $\theta \in \{0.6, 0.7, 0.8, 0.9, 1\}$ 的变化来讨论系统的性能和适应性。评估主要是以下 3 个方面:服务合成的响应时间、每次服务合成的跳数和自适应值。

4.2 结果和分析

服务响应时间和每次服务请求的跳数分别是 Web 服务

合成质量、费用评定的主要指标。第一个实验仿真分析方法的适应性,仿真结果如图2所示。这里设置 $\theta=0.8$,生物实体为200个。为了对比合成方法的性能,特别设计了随机非生物实体(不带生物功能的移动 Agent)的情况与 $\alpha \in \{0, 0.5, 0.8, 1\}$ 的对比数据。从图2的仿真结果可以看出,非生物实体响应时间参数的趋势没有变化。生物实体的相应数据在最初较高,达到700ms;在200ms以后,数据指标开始下降至300~400ms。这是因为生物实体在开始阶段要建立相应的亲和度关系,许多资源和时间都浪费在盲目的寻找上。同时,生物操作开始时存储能量,进行迁移、变异和交叉等生物操作以完成服务合成。随着时间的推移,生物实体形成了关系的聚类,很快响应提供服务,费用也变少了。同时可以看到, $\alpha=0.5$,反应时间最少。对服务合成跳数的数据分析也可得到同样的结论。以上仿真实验结果表明,Web服务合成模型自适应、自进化地满足用户的需求,生物实体关系网络所构建的Web服务合成方法具有自适应、自进化等功能特点。

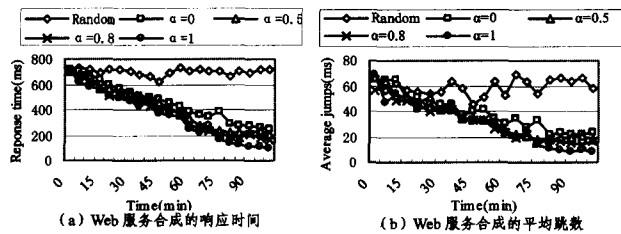


图2 仿真结果

第二个仿真实验是服务合成的成功率随 θ 和 α 的变化情况,如表1所列。仿真分别设置了 θ 和 α 。当 θ 较小时, α 的变化对成功率的影响很小;随 θ 增大, α 的变化影响变大,系统的信任关系起到了很大作用。 $\alpha=0$ 时,系统拥有较高的成功率,但其费用也较高。本实验数据表明, α 的正确选择加强了信任关系的反馈和调节,对系统的性能起到了提高的作用。

表1 不同阈值下的成功率

θ	$\alpha=0$	$\alpha=0.5$	$\alpha=0.8$	$\alpha=1$
0.6	94.5	92.7	91.8	90.6
0.7	86.1	80.1	80.5	78.9
0.8	81.1	74.6	75.9	74.9
0.9	65.5	62.1	59.6	57.4
1.0	49.8	48.1	48.1	47.1

第三个仿真实验是考虑在服务计算环境中,服务和网络很有可能变得不可靠(即不能发挥资源发现的作用)。本文仿真了这种环境下发现方法的自适应性和抗毁性。在实验中,随机选择1%的信息服务生物实体和1%网络节点并将它们设置为不可靠,设计了一个评估指标自适应性 D (可靠和不可靠状态情况下响应时间的相对值)来评估Web服务突现方法的特性,如式(8)所示:

$$D = \frac{N_{res-time}^{unreliable} - N_{res-time}^{reliable}}{N_{res-time}^{reliable}} \quad (8)$$

其中: $N_{res-time}^{unreliable}$ 和 $N_{res-time}^{reliable}$ 代表不可靠和可靠状态情况下的响应时间。这个实验设置了生物实体数量为100,200,300,400和500等情况。从表2的结果看,生物实体的数量越大,亲和度网络的性能越好,系统的自适应性和抗毁性也越好。

表2 适应D值

θ	$\alpha=0$	$\alpha=0.5$	$\alpha=0.8$	$\alpha=1$
100	5.2	5.4	5.2	4.8
200	4.7	3.6	4.3	4.5
300	4.6	3.4	3.3	3.8
400	4.3	3.5	5.2	5.4
500	4.0	3.2	3.0	3.1

结束语 Web服务是一个正在迅速发展的研究领域,服务动态管理复杂性的特点与生物系统相似。本文借鉴细胞因子网络的协作机理,结合生物网络、Agent、Web服务和语义的有关原理,设计了面向Web服务合成的协同进化智能模型,利用生物实体消息匹配的多亲和度值构成关系亲和度细胞因子网络。仿真的良好特性显示这种方法可以为下一代Web服务动态合成、自进化适应等设计提供一种新颖的研究途径和思路。进一步的工作是扩充Web服务突现方法的功能模块,完善智能管理调节设计,提高系统的并行效率和语义匹配效率,探讨其在复杂系统中的应用。

参考文献

- [1] Bultan T, Xiang F, Hull R. Convehation specification: a new approach to design and analysis of e-service composition[C]// Proceeding of the 12th International WWW Conference, 2003
- [2] 孙宏彬,丁永生. 基于生物网络的 e-service 自进化合成方法研究[J]. 计算机应用研究, 2008, 25(4): 999-1003
- [3] Hai Z-G. The future interconnection environment [J]. IEEE Computer, 2005, 38(4): 27-33
- [4] Vargas P A, Muioli R, Castro L N D. Artificial homeostatic system: a novel approach advances in artificial life[C]// Proceeding of the 8th European Conference, 2005
- [5] Liu Z-Q, Ding Y-S, Zhang X-F. Bio-entity inspired from immune TSTS model for distributed intrusion detection system[J]. Dynamics of Continuous, Discrete and Impulsive Systems-Series B: Applications & Algorithms, 2006, 13(8): 46-50
- [6] 丁永生,任立红. 一种基于免疫突现计算的生物网络结构的设计[J]. 控制与决策, 2003, 18(2): 185-189
- [7] Gao L, Ding Y-S, Hao Y. An adaptive social network-based computation platform as grid middleware system[J]. International Journal of Intelligent Systems, 2004, 19(10): 859-884
- [8] Sun H-B, Ding Y-S, Hao K-R. Web service emergent system based on mechanism of bio-systemic intelligence[J]. Dynamics of Continuous Discrete and Impulsive Systems Series B: Application & Algorithms, 2006, 13(8): 6-10
- [9] Berardi D, Glacomo G D, Lenzenrini M. Synthesis of underspecified composite e-services based on automated reasoning[C]// Proceeding of the 2nd International Conference on Service Oriented Computing, 2004
- [10] Zeng L-Z, Benatallah B, Ngu A H H. QoS-aware middleware for services composition[J]. IEEE Trans on Software Engineering, 2004, 30(5): 311-327
- [11] Cardoso J, Sheth A. Semantic e-workflow composition[J]. Intelligent Information Systems, 2003, 21(3): 191-225