

面向用户服务质量需求的服务选择方法研究

崔纪鹏^{1,2}, 马炳先^{1,2}, 张正明^{1,2}

(1. 山东网络环境智能计算技术重点实验室, 山东 济南 250022;

2. 济南大学 信息科学与工程学院, 山东 济南 250022)

摘要:服务选择问题是服务组合中的一个关键问题,尤其在服务选择时应充分考虑服务质量因素。该文研究了基于用户服务质量需求的服务选择方法,通过对各项 QoS (quality of service) 因素分别加权,将面向用户服务质量需求的服务选择问题转化成单目标优化问题。在充分考虑服务实际质量和用户主观需求的基础上,给出了 QoS 组合权重的计算模型;并在候选服务集的基础上,依据用户的服务质量需求,定义了用户满意度函数,进一步利用遗传算法解决服务选择问题。实验结果表明,文中提出方法具有明确的用户需求指向性和选择灵活性,能够最大程度上满足用户需求。

关键词: Web 服务; 服务质量; 服务选择; 遗传算法

中图分类号: TP31

文献标识码: A

文章编号: 1673-629X(2012)12-0038-05

Study on Service Selection Method Oriented to User QoS Requirements

CUI Ji-peng^{1,2}, MA Bing-xian^{1,2}, ZHANG Zheng-ming^{1,2}

(1. Shandong Provincial Key Laboratory of Network Based Intelligent Computing, Jinan 250022, China;

2. School of Information Science and Engineering, University of Jinan, Jinan 250022, China)

Abstract: Service selection is a key issue in service composition, and the quality of service should be taken full account of, especially. As to service selection oriented to user QoS requirements, a method is given to transform it into a single objective optimization problem by weighting each QoS factor. Considering both actual quality of service and users' preferences, propose a combination weight computing model; also, according to users' QoS requirements, define the concept of degree of users' satisfaction, further by using the genetic algorithm to solve service selection problem. The method is proved to be more effective in meeting users' requirements, both on demand directivity and dynamic selecting flexibility.

Key words: Web service; quality of service; service selection; genetic algorithm

0 引言

Web 服务提供了一种在不同平台之间进行应用层功能自动整合集成和自动化处理所需要的技术架构,为功能整合集成和信息交换处理提供基础^[1]。但是单个 Web 服务所提供的功能较为单一,因此为了实现完整的业务功能,满足用户的不同需求,需要把分散的 Web 服务组合起来,形成具有增值价值的 Web 服务组合。

实际应用中的 Web 服务往往在功能上满足某一

特定需求而性能各异,如何在这些服务中选择合适的服务是服务组合中的关键问题^[2]。目前,解决 Web 服务组合中服务选择问题的方法主要是通过 Web 服务描述模型自身提供的功能属性和非功能特性,搜索和选择满足用户请求的 Web 服务^[3-7];文献[8,9]通过把服务组合流程中的各个 QoS 约束线性加权转化为一个单目标函数,但没有能够解决权重向量的敏感问题,限制了算法的实用性。

遗传算法由于其特有的优势,已被应用到服务选择问题求解中^[10-12],但现有遗传算法的搜索方向多以全局 QoS 最大为目标,用户 QoS 需求仅仅被当做是算法终止条件的一部分,容易导致搜索偏离目标,影响算法的效率。文中从用户 QoS 需求出发,给出了 QoS 组合权重计算模型,并在此基础上提出了基于用户 QoS 需求的服务选择算法,具体如下:

(1) 将 QoS 权重向量分解为专家选择 QoS 权重

收稿日期: 2012-05-15; 修回日期: 2012-08-20

基金项目: 国家自然科学基金资助项目(60903099); 山东省优秀中青年科学家奖励基金(BS2009DX012); 山东省高等学校科技计划项目(J09LG14)

作者简介: 崔纪鹏(1985-),男,山东临沂人,硕士生,CCF 会员,研究方向为服务计算; 马炳先,博士,副教授,硕士生导师,研究方向为 Petri 网、服务计算。

和用户 QoS 需求权重:专家选择 QoS 权重反映特定服务领域各项 QoS 属性的相对重要程度;用户 QoS 需求权重反映用户的偏好。

(2) 定义了依赖 QoS 组合权重的用户满意度函数和适应度函数,使服务选择算法更具指向性和灵活性。

1 候选服务集

Web 服务描述模型^[9]: $W = \{F, Q\}$, 其中 F 为服务的功能属性集合, Q 为服务质量的集合。假设有一个 Web 服务构成的空间 S , 它定义为一个抽象服务类的集合。任何一个抽象服务类 $S_j \in S$ 都可以用来描述一组功能特性 F 相同的 Web 服务, 可做如下定义^[13]:

定义 1 抽象组合服务, 记做 $CS_{abstract} = \{S_1, S_2, \dots, S_n\}$, 表示构成一个组合服务的所有抽象服务类; 具体组合服务, 作为抽象组合服务的一个是实例, 可以通过将抽象组合服务中的每一个抽象服务类 S_j 替换成该类的一个具体 Web 服务得到。

抽象组合服务由一系列的抽象类构成, 这些抽象类所有可能的实例集就是候选服务集, 它描述三方面信息: 服务由哪些抽象子服务类组合而成、抽象子服务类的组合顺序、每个抽象子服务类可用的具体服务。

2 QoS 组合权重计算模型

现有文献中^[14-16]都是假定 QoS 权重向量已给定, 比如有 5 个 QoS 评估指标, 权重向量可设为 $w_1 = (0.2, 0.1, 0.4, 0.1, 0.2)$, 但相关文献并没有给出权重是如何获取的。

用户偏好能够体现用户的主观意愿^[15-17], 它从主观上反映了用户对服务质量的要求, 但单纯由用户自己来指定权重向量会存在一些问题。首先, 用户的偏好是主观的想法, 适合用语言文字描述, 使用时需要量化。其次, 不同的行业、领域的 Web 服务, 可能会涉及到一系列与行业相关的专业化指标。

文中根据运筹学关于权重分解的相关理论, 给出一种 QoS 组合权重计算方法, 将组合权重分解为专家选择 QoS 权重和用户 QoS 需求权重。

2.1 专家选择 QoS 权重的获取

包含领域专家经验的权重指标可以弥补用户偏好的不合理性。专家选择 QoS 权重具有相对的客观性、合理性和稳定性, 文中通过层次分析法获得专家选择 QoS 权重。

2.1.1 层次分析法

层次分析法^[18]是一种解决多目标复杂问题的定性与定量相结合的决策分析方法, 用决策者的经验判断、衡量各目标之间的相对重要程度, 并合理地给出权

值。其特征是合理地将定性与定量决策结合起来, 把决策过程层次化、数量化。

层次分析法在确定权重时, 为了获得定量的结果, 用判断矩阵表示上层某个因素下, 本层所有的因素的相对重要性的比较。判断矩阵的元素用 Santy 的 1~9 标度法给出, 如表 1 所示。

表 1 判断矩阵元素 a_{ij} 的标度方法

标度	含义
1	表示两个因素相比, 具有同样的重要性
3	表示两个因素相比, 一个因素比另一个稍微重要
5	表示两个因素相比, 一个因素比另一个明显重要
7	表示两个因素相比, 一个因素比另一个强烈重要
9	表示两个因素相比, 一个因素比另一个极端重要
2, 4, 6, 8	上述两相邻判断的中值
倒数	因素 i 与因素 j 的判断为 a_{ij} , 则 j 与 i 的判断 $a_{ji} = 1/a_{ij}$

判断矩阵要满足一致性要求。对于矩阵 A , 如果满足 $a_{ij}a_{jk} = a_{ik}$, 则称矩阵 A 为一致性矩阵, 这是为了保证各元素重要程度之间的协调性。实际应用中, 需要对判断矩阵进行一致性检验。

引入一致性指标 CI 来衡量判断矩阵的不一致程度: $CI = \frac{\lambda - n}{n - 1}$, CI 越接近于 0, 矩阵的一致性程度越好。可以用一致性比率 CR 衡量 CI 的大小, CR 定义为 CI 和随机一致性指标 RI 的比值: $CR = CI/RI$ 。其中 RI 可以通过查表的方式获得。

文中选用 QoS 的 5 个指标, 构造的判断矩阵为 5×5 矩阵, 对应的 $RI = 1.12$; 如果经过计算, 一致性比率 CR 的值小于 0.1, 则认为通过一致性检验, 否则可以按文献[19]提供的方法对判断矩阵进行调整。

2.1.2 专家选择 QoS 权重的获取

运用层次分析法获取专家选择 QoS 权重的步骤如下:

(1) 建立层次模型, 如图 1 所示。

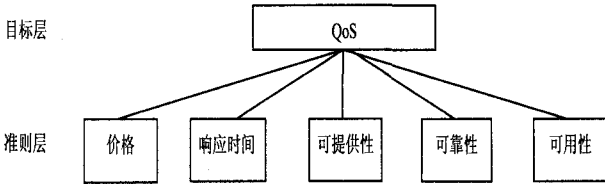


图 1 层次模型

该层次模型中, QoS 为目标层, 表示解决问题的目的; 准则层从价格、响应时间、可提供性、可靠性和可用性 5 个因素对目标层进行分解。通过该模型可以获得

在目标层下,准则层各因素之间的相对重要程度,即各因素的权重。

(2)构造专家评价矩阵,即判断矩阵,如表2所示。

表2 判断矩阵

指标	价格	响应时间	可提供性	可靠性	可用性
价格	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
响应时间	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}
可提供性	X_{31}	X_{32}	X_{33}	X_{34}	X_{35}
可靠性	X_{41}	X_{42}	X_{43}	X_{44}	X_{45}
可用性	X_{51}	X_{52}	X_{53}	X_{54}	X_{55}

其中 $X_{ij} = 1/X_{ji}$, $X_{ii} = 1$ ($i = 1, 2, 3, 4, 5; j = 1, 2, 3, 4, 5$)

(3)对步骤(2)构造的判断矩阵进行一致性检验:如果不能通过一致性检验,则需要对判断矩阵进行调整^[19]。

(4)判断矩阵通过了一致性检验之后,就可以根据判断矩阵得到专家选择 QoS 权重向量,即判断矩阵最大特征值对应的特征向量。

2.2 用户 QoS 需求权重的获取

用户往往习惯于用自然语言的形式描述自己的偏好,比如用{非常重要,重要,一般,不重要,可以忽略}一组表示重要程度的概念来描述。这些概念本身是模糊概念,属于模糊集合的范畴,文中采用三角模糊数^[20]来描述。获取用户 QoS 需求权重有以下四个步骤:

(1)用户对各项 QoS 属性的重要程度进行选择,选择范围为{非常重要,重要,一般,不重要,可以忽略};

(2)将用户的选择结果映射为三角模糊数,可以采用如下的三角模糊数集表示上述重要程度: $\{[10, 10, 8], [5, 7, 9], [3, 5, 6], [1, 3, 4], [1, 1, 2]\}$;

(3)对得到的三角模糊数集进行量化,得到实数集,这些实数按顺序构成用户 QoS 需求权重向量,即 $SP = (s_1, s_2, \dots, s_n)$, s_i 表示用户对第 i 项 QoS 属性重要性程度的选择评价结果,由第 i 项 QoS 属性对应的三角模糊数去模糊化得到;

(4)对得到的用户 QoS 需求权重向量进行归一化处理。

2.3 QoS 组合权重计算

组合 QoS 权重即用户 QoS 需求权重和专家选择 QoS 权重的加权平均,这样得到的权值既考虑了用户的偏好和需求,同时又不失专业领域特性,计算过程如下:

专家选择 QoS 权重为: $OP = (w_1, w_2, \dots, w_n)$

w_i 表示第 i 项 QoS 属性的专家评价权重。

用户 QoS 需求权重为: $SP = (s_1, s_2, \dots, s_n)$

s_i 表示第 i 项 QoS 属性的用户评价权重。

权重组合方法目前研究的很多,文中采取简化的方法,认为重要性的组合是乘积的关系,以此得到组合 QoS 权重,并对其归一化处理得到:

$$CP = \left[\frac{s_1 \times w_1}{\sum_{i=1}^n s_i \times w_i}, \frac{s_2 \times w_2}{\sum_{i=1}^n s_i \times w_i}, \dots, \frac{s_n \times w_n}{\sum_{i=1}^n s_i \times w_i} \right]$$

3 基于遗传算法的服务选择方法

3.1 问题描述

借鉴文献[21]的方法对 Web 服务的质量参数进行选择,并建立了包括 5 个服务质量参数的参数体系,它们分别是花费、响应时间、可靠性、可用性和可提供性。假定组合服务由 n 个子服务构成,则组合服务的各项 QoS 指标可以按照下述公式计算得到^[21]:

花费: $\text{cost} = \sum_{i=1}^n \text{cost}_i$ cost_i 为第 i 个服务的花费;

响应时间: $\text{response} = \sum_{i=1}^n \text{response}_i$ response_i 为第 i 个子服务的响应时间;

可用性: $\text{usable} = \prod_{i=1}^n \text{usable}_i$ usable_i 为第 i 个服务的可用性;

可靠性: $\text{rel} = \prod_{i=1}^n \text{rel}_i$ rel_i 为第 i 个子服务的可靠性;

可提供性: $\text{supl} = \prod_{i=1}^n \text{supl}_i$ supl_i 为第 i 个子服务的可提供性。

QoS 指标可分为积极指标和消极指标:其中可靠性、可用性和可提供性为积极指标,数值越大越好;花费和响应时间为消极指标,数值越小越好。假定用户可以指定各项 QoS 需求,则用户的 QoS 需求可以用下面的不等式组来描述^[21]:

$$\begin{cases} \text{cost} < C \\ \text{response} < \text{RES} \\ \text{rel} > \text{REL} \\ \text{usable} > \text{US} \\ \text{supl} > \text{SUP} \end{cases} \quad (1)$$

其中, $C, \text{RES}, \text{REL}, \text{US}, \text{SUP}$ 均为用户指定的常数,分别代表用户允许的最大花费、最大响应时间、最低可靠性、最低可用性和最低可提供性;服务选择将从满足用户 QoS 需求的角度进行。

3.2 基于遗传算法的服务选择方法

(1)染色体编码。

采用二进制分段编码方式。对于具有 n 个子服务类的组合服务,染色体划分成 n 段,其中每一段为相应子服务类的候选服务预留位置。编码长度应该有自适应性:长度太小不能覆盖整个搜索空间,可能会丢掉最优解;太大则会影响搜索的效率,同时造成资源的浪费。假设第 i 个子服务类的具体服务个数为 a_i ,则该子服务对应的编码长度为: $L_i = \lceil \log a_i \rceil$,组合服务的编码长度为: $L = \sum_{i=1}^n L_i$,编码长度是可变的,取决于子服务类的个数以及子服务类中具体服务个数。

(2) 用户满意度和适应度函数。

服务选择算法应该从最大限度满足用户需求的角度设计。为了方便描述,定义用户满意度的概念。用户满意度是用户对各项 QoS 指标满意程度的加权和,应该能够反映服务选择结果和用户实际需求的差距。当且仅当用户的各项 QoS 需求得到满足时,用户满意度为 1;服务选择结果和用户实际需求差距越大,用户满意度越小。

消极指标满意度定义为:

$$f(x) = \begin{cases} 1 & x \leq X \\ e^{-\frac{x-X}{\Delta}} & x > X \end{cases} \quad (2)$$

积极指标满意度定义为:

$$f(x) = \begin{cases} 1 & x \geq X \\ e^{-\frac{X-x}{\Delta}} & x < X \end{cases} \quad (3)$$

最终用户的满意度函数为:

$$F(x) = \sum_{i=1}^5 \beta_i f_i(x) \quad (4)$$

其中, α 为用户偏好权重, β_i 为组合权重;将(1)~(3)式带入到(4)式中,即可得到对应于某一用户 QoS 输入的用户满意度函数。将用户满意度函数作为适应度函数,即适应度函数: $\text{obj}(x) = F(x)$ 。

(3) 终止条件的设定。

用户满意度满足一定条件的时候算法终止;如果满意解不存在,而搜索空间又很大,若发现群体中一定比例的个体已经是同一个个体,则终止算法的迭代。

(4) 文中用轮盘赌方法实现选择操作;采用二点交叉,交叉率设定为 0.5;对染色体中的每一位二进制数值按照变异概率进行变异操作,变异概率设定为 0.015。

4 实 例

假定一个服务的候选服务集由 5 个抽象服务类构成,每一抽象服务类中分别有 22,7,4,4,13 个候选服务,则搜索空间大小为 $22 \times 7 \times 4 \times 4 \times 13 = 32032$ 。可以根据候选服务集计算并给出实际可提供服务的 QoS 范围,以供用户做进一步的选择,如表 3 所示。

表 3 实际可提供服务 QoS 范围

属性名	下限	上限
花费	568.0	1311.0
响应时间	12.0	195.0
可靠性	0.813	0.99
可用性	0.752	1.0
可提供性	0.67	1.0

用户可根据表 3 的提示,在可满足范围内输入自己的 QoS 需求,比如输入 QoS 向量为 {700,20,0.90,0.92,0.94}。最后,用户需要给出用户 QoS 需求权重向量 SP (假定专家选择 QoS 权重向量 OP 已经事先由领域专家给出)。

用户在选择个人偏好时,将重要性程度映射为区间 [1,10] 内的值,数值从 1 到 10 变化表示用户认为某项 QoS 属性的重要程度依次递增。重要性程度的取值由用户指定,未指定的项默认取值为 5。假设用户认为花费非常重要,而响应时间可以忽略不计,那么他的偏好输入向量可以是 $\text{vec} = (9,1,5,5,5)$ 。算法搜索的结果如表 4 所示:

表 4 权重输入为 $\text{vec} = (9,1,5,5,5)$ 时的输出

服务类	可用服务数	输出服务号
1	22	2
2	7	7
3	4	4
4	4	2
5	13	6

在服务选择时,遗传算法最高进化代数设定为 200 代。算法在两种情况下提前终止:

(1) 用户满意度 Satis 满足一定条件时(这里设定为 $\text{Satis} > 0.99$)算法终止;

(2) 连续 20 代的最优适应度不再变化时算法终止。

该实例下算法的最终进化代数为 40 代,原因是最优适应度从第 21 代开始连续 20 代不再变化。输出组合服务各项 QoS 属性: $\text{cost} = 694 < 700$, $\text{response} = 29 > 20$, $\text{rel} = 0.953 > 0.9$, $\text{usable} = 0.972 > 0.92$, $\text{supl} = 0.942 > 0.94$,最优用户满意度为 $\text{Satis} = 0.988$ 。对比用户输入的 QoS 需求向量可知,该组合结果没有能够满足响应时间 $\text{response} < 20$ 的需求,但由于用户认定响应时间这一指标无关紧要,可以认为该组合结果能够满足用户的需求。

图 2 描述了当其他输入权值固定不变,某单一 QoS 属性权重从 1 到 10 变化时,搜索结果中该属性的单项用户满意度变化规律。为了不失一般性,实验选取可提供性和花费为代表,分别考察了单项积极指标

和单项消极指标满意度随输入权重的变化规律。不难看出,当某一项 QoS 属性的输入权重相对较小时,在算法的输出结果中,该项 QoS 满意度会比较低;当该输入的输入权重增加时,对应的单项用户满意度会逐渐提高。这说明算法输出结果是可控的,用户可以根据自身的需求和算法输出结果,以调整输入权值的方式有效地选择到更为符合自己需求的服务。

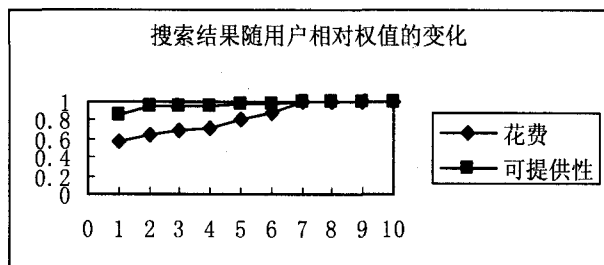


图2 单项用户满意度随其输入权重的变化规律

图3给出了用户满意度随搜索代数数的变化规律。从图中不难看出,相对于普通的遗传算法,文中的方法在收敛速度上有明显的优势,能够以更快的速度找到最能满足用户需求的服务组合。

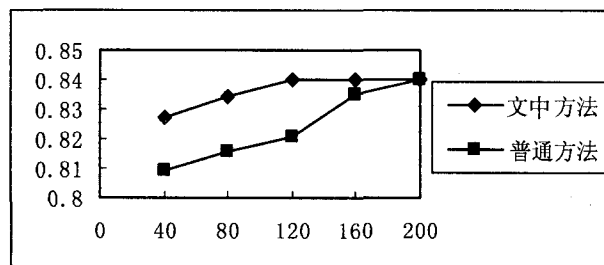


图3 用户满意度随算法搜索代数数的变化

5 结束语

面向用户服务质量需求,文中给出了组合权重计算模型,综合考虑了用户需求的可满足性和合理性,允许用户设置输入权重和 QoS 需求,设计并实现了服务选择的遗传算法,获得满足用户服务质量需求的服务;同时,用户可以根据服务选择算法反馈的搜索结果,有目的地调整输入权重,并通过再选择的方式获得满意度更优的组合方案。

下一步工作主要包括:

1) 进一步优化服务选择算法,寻找更好的用户满意度函数表达式,同时考虑算法参数随问题规模和搜索代数动态变化的可能性;

2) 在实际应用中进一步检验算法的可靠性和可用性,并加以优化改进。

参考文献:

- [1] 王宏琦,刘旭东,王 勇. 基于 QoS 的服务选择系统的研究和实现[J]. 微计算机信息,2006(10X):168-170.
- [2] 岳 琨,王晓玲,周傲英. Web 服务核心支撑技术:研究综

述[J]. 软件学报,2004,15(3):428-442.

- [3] 廖 渊,唐 磊,李明树. 一种基于 QoS 的服务构件组合方法[J]. 计算机学报,2005,28(4):627-634.
- [4] Tomic V. Service offerings for XML web services and their management applications [D]. Ottawa: Carleton University, 2004.
- [5] Benatallah B, Dumas M, Sheng Q Z, et al. Declarative composition and peer to peer provisioning of dynamic web services [C]//ICDE 2002. San Jose CA: IEEE Press, 2002:297-308.
- [6] 陈燕萍,李增智,郭志胜,等. Web 服务组合中基于服务质量的服务选择算法[J]. 西安交通大学学报,2006,40(8):897-900.
- [7] Liu Y T, Ahhl N G, Zeng L. QoS computation and policing in dynamic web services selection [C]//WWW2004. New York: ACM Press, 2004:66-73.
- [8] Zeng L Z, Benatallah B, Dumas M. Quality driven web service composition [C]//WWW2003. Budapest: ACM Press, 2003:411-421.
- [9] 夏 宏,李增智. 粒子群算法求解 Web 服务组合中基于 QoS 的服务选择[J]. Journal of Beijing University of Posts and Telecommunications, 2009,32(4):63-67.
- [10] 张成文,苏 森,陈俊亮. 基于遗传算法的 QoS 感知的 Web 服务选择[J]. 计算机学报,2006,29(7):1029-1037.
- [11] 李淑芝,杨书新,王继升. 基于树形二重结构编码遗传算法的服务选择[J]. 计算机软件,2008,28(12):3017-3020.
- [12] Canfora G, Dipenta M, Esposito R. An approach for QoS-aware service composition based on genetic algorithm [C]//Proc. of the 2005 Conference on Service Oriented Computing. New York: [s. n.], 2004:232-239.
- [13] Alrifai M, Risse T, Dolog P, et al. A Scalable Approach for QoS-based Web Service Selection [C]//Proc of the 1st International Workshop on Quality-of-Service Concerns in Service Oriented Architectures. Berlin: Springer, 2008:190-199.
- [14] Yang Fangchun, Su Sen, Li Zhen. Hybrid QoS-aware semantic web service composition strategies [J]. Information Sciences, 2008,51(11):1822-1840.
- [15] Ran S. A Model for Web Services Discovery With QoS [J]. ACM SIGecom Exchanges, 2003(4):1-10.
- [16] Herssens C, Jureta I J, Faulkner S. Capturing and Using QoS Relationships to Improve Service Selection [C]//CAiSE 2008. [s. l.]: [s. n.], 2008:312-327.
- [17] 邵凌霜,李 田,赵俊峰,等. 一种可扩展的 Web Service QoS 管理框架[J]. 计算机学报,2008(8):1458-1470.
- [18] 岳超源. 决策理论与方法[M]. 北京: 科学出版社, 2003.
- [19] 王传玉. 改进 AHP 中判断矩阵一致性的一种新方法[J]. Journal of Anhui Institute of Mechanical & Electrical Engineering, 2001,16(4):47-50.
- [20] 李登峰. 模糊多目标多人决策与对策[M]. 北京: 国防工业出版社, 2003.
- [21] 王 勇,胡春明,杜宗霞. 服务质量感知的网格工作流调度[J]. Journal of Software, 2006,17(11):2341-2351.