

基于 QoS 的卫星信息应用链服务调度方法^①

郭少阳, 邱涤珊, 张 斌

(国防科学技术大学 信息系统工程重点实验室, 长沙 410073)

摘 要: 服务调度问题是卫星信息应用链功能结构设计完成之后必须解决的问题. 在卫星信息应用链服务调度问题的相关概念进行描述之后, 提出了卫星信息应用链的服务质量参数体系, 为链的性能评价和服务选择提供了依据, 并在此基础上建立了卫星信息应用链的服务调度问题的数学模型, 提出了一种基于 GA 的服务调度算法. 实验结果表明, 本文所提出的服务质量参数体系能有效解决应用链性能评估问题, 基于 QoS 的卫星信息应用链服务调度方法可行性较强.

关键词: 卫星信息应用链; 服务调度; 服务质量; 遗传算法

QoS-Based Service Scheduling Method of Satellite Information Application Chain

GUO Shao-Yang, QIU Di-Shan, ZHANG Bin

(Science and Technology on Information Systems Engineering Laboratory, National University of Defence Technology, Changsha 410073, China)

Abstract: The problem of service scheduling is proposed after the design of structure and function of Satellite Information Application Chain(SIAC). After introducing some basic conceptions about the service scheduling problem of SIAC, the QoS parameters of SIAC and the calculus algorithms are discussed, which can be used as criterions for performance evaluation and service selection of SIAC. After that, the mathematical model of service scheduling problem of SIAC is established, and service scheduling algorithms based on GA are presented. The experimental results indicates the proposed QoS parameters are effective for the evaluation of SIAC and the QoS-based method used for SIAC service scheduling is feasible.

Key words: satellite information application chain; service scheduling; QoS; genetic algorithm

卫星信息数据包括可见光、红外、超光谱、SAR 以及电子等多种类型, 各类信息数据的时空属性、特点、作用以及处理流程模式不尽相同. 但是根据卫星信息处理的一般流程, 其处理功能一般包括用户需求提交、侦察任务规划、卫星数据接收、通用数据处理、专业数据处理、情报产品生成以及分发等^[1,2], 卫星信息应用链将这些代表了不同的信息处理功能的节点按一定的顺序联结起来, 每个功能节点内部可能存在大量的候选服务. 在应用链的设计完成之后, 如何评价其性能, 并为每个功能节点指派特定的候选服务, 是应用链投入运行前需要解决的问题.

文献[3-5]采用排队网络方法对加拿大军队地理信息部基础行动数据(FBO)的生产过程进行建模, 假设不同服务节点的服务台的功能属性和非功能属性是相同的, 并使用 lingo 软件求解服务台资源分配方案. 显然, 这一假设在实际应用中有很大的局限性. 文献[6-7]提出通过并行服务的思想来提高 workflow 执行的成功率和时效性, 所考虑的是 QoS 相同的多个服务同时执行的情况, 而对 QoS 不同的情况却未考虑. 在无线网络通信领域^[8-10], 关于 QoS 的研究很多, 802.16 标准提出了完备的 QoS 保证机制, 据此建立可靠的网络通信质量参数体系, 并在此基础上对 PMP 和 Mesh 模式下 QoS 架

^① 收稿时间:2013-04-25;收到修改稿时间:2013-05-20

构、包调度及接纳算法进行深入研究,在实际中具有成熟的应用,对于本文研究的问题具有启示意义。

针对上述分析,本文提出了卫星信息应用链的服务调度问题,并展开了系统研究.第 1 节对该问题进行了分析和描述,第 2 节建立了应用链服务质量模型和服务调度模型,第 3 节提出了应用链的 QoS 估算算法和基于应用链 QoS 的服务调度算法,第 4 节通过案例对上述模型和算法进行了验证,第 5 节总结全文。

1 卫星信息应用链服务调度问题概述

卫星信息应用链(Satellite Information Application Chain, SIAC)是指为满足作战任务需求,以按照一定的方式将卫星应用装备、基础设施、人员等资源有机组织起来所构成的功能网链结构为承载的信息服务网络.一般来说 SIAC 节点的属性包括功能属性和非功能属性.功能属性是节点的固有属性,反映了其所能执行的功能,SIAC 就是依据卫星信息服务的运行流程将不同功能的节点进行联结而得到的;非功能属性反映了 SIAC 的性能水平,主要是指节点的服务质量(QoS),如响应时间、可靠性等.它反映了节点完成任务的效果,在实际运行时受多种因素的影响。

对于一条设计好的SIAC,其各节点的功能属性一般就已确定了,而非功能属性可以在运行中因设置不同而发生变化.由于在各功能节点中往往存在大量功能相同而 QoS 不同的候选服务,选择不同的候选服务将最终导致整个 SIAC 服务质量的不同.那么,各节点的 QoS 与 SIAC 的 QoS 之间的关系如何,以及如何根据用户对 SIAC 的 QoS 要求来从各节点中选择候选服务呢?

SIAC 调度是指依据应用链中各个成员服务的执行次序,从各节点存在的大量候选服务中选择若干合适的服务进行执行,以满足用户对 SIAC 各性能指标的要求,其核心功能是依据候选服务的服务质量信息进行选择.基于 QoS 的调度把对整个 SIAC 的执行要求规划到各个成员服务,并作为服务的非功能特性方面的要求.在进行调度之前,需要建立 SIAC 的 QoS 要求与其所包含的成员服务 QoS 之间的关系。

2 基于QoS的SIAC服务调度模型

2.1 SIAC 基本模型

从本质上来讲,SIAC 是一种卫星信息服务(Satellite Information Service, SIS),它是以用户特定任务与目标为

导向、以卫星所获取的目标与环境数据为原料、以卫星信息保障资源为依托、以卫星信息服务内容为基础、按照一定的策略为用户提供信息服务创造价值并满足用户需求的专业活动. SIS 按粒度大小可分为:基本服务、链路级服务和网络级服务.其中,基本服务是与具体操作相关的活动,具有原子性,如辐射校正.由若干基本服务依据一定的方式进行组合嵌套可形成链路级服务,链路级服务与特定的功能相关,同理可得网络级服务. SIAC 节点之间的互联主要有以下几种组合方式。

表 1 基本服务组合表

服务组合类型	服务组合图例	含义
顺序(Sequence)	WS1 → WS2	成员服务依次执行
循环(Loop)	WS1 (循环箭头)	服务不断循环直至检测到终止条件
选择(Switch)	XOR { WS1, WS2 }	选择进入某一个分支服务
并行(Parallel)	Start { WS1, WS2 }	同时进入多个分支服务
同步(And-Join)	{ WS1, WS2 } → End	多个服务进行合并后方可进入下一节点
合并(XOR-Join)	{ WS1, WS2 } → XOR → End	多个服务进行比较后进入下一节点

一种 SIAC 网络级服务模型和链路级服务模型如图 1 所示。

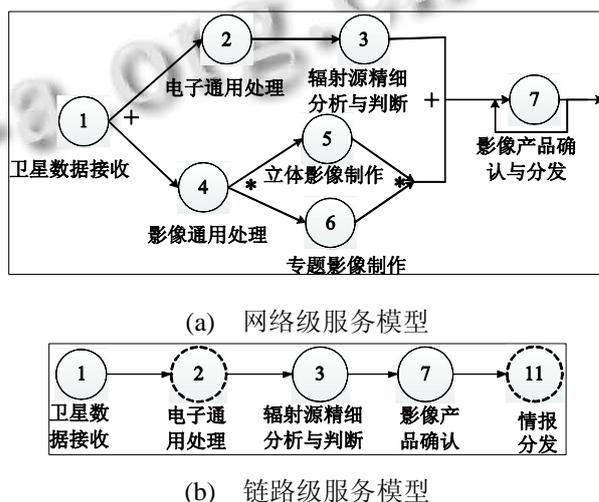


图 1 不同粒度的 SIAC 服务模型

2.2 SIAC 服务质量模型

从应用链设计的角度来看,SIAC 可形式化表示为 $\langle S, E, Q \rangle$ 三元组的形式.其中 $S = \{S_1, S_2, \dots, S_n\}$ 表示 SIAC

中功能节点的集合, 代表服务的功能需求. $E=S \times S$ 是服务之间关系的集合, 如果 $\langle S_i, S_j \rangle \in E$, 则表示服务 S_i 是服务的前提, 即只有在服务 S_i 完成的情况下服务 S_j 才能开始. $Q=\{Q_1, Q_2, \dots, Q_k\}$ 表示应用链中服务或服务组合的 QoS 约束. 本文借鉴文献[7]的军事信息服务三维 QoS 模型, 将卫星信息应用链各成员服务的质量(QoS)模型表示为:

$$Q = \{T, A, Rel, C\}$$

其中:

(1) 执行时间(T): 表示从发出服务调用请求到获得该服务处理结果所花的时间. 服务执行时间包括服务本身处理用户请求所花时间 $T_{process}$ 和数据在网上传输所需时间 T_{trans} , 即服务时间 $T = T_{process} + T_{trans}$.

(2) 可用性(A): 表示服务可以被访问, 即可用的概率. 可以通过平均无故障时间和平均修复时间^[11]来定义, 令 $A = MTF_s / (MTTR_s + MTF_s)$, $MTTF_s$ 表示服务 S 的平均无故障时间, $MTTR_s$ 表示服务 S 的平均修复时间.

(3) 可靠性(Rel): 表示服务请求被成功响应(服务操作得到成功执行, 同时表示操作成功执行的消息成功的返回给服务消费者), 服务成功率通过以往服务被成功执行的次数与服务被调用的总次数的比值来度量, $Rel = N_c(s) / N_{invoke}$, 其中 $N_c(s)$ 为服务被成功调用的次数, N_{invoke} 为服务被调用的总次数.

(4) 执行代价(C): 服务的调用会对网络资源带来诸如带宽、服务宿主节点计算能力等的开销, 如果这部分开销过大会影响使用该服务的其它任务的完成, 可用系统负载反应服务的执行代价, 令 $C = e^{(\Delta x / L)^2}$, 其中 Δx 为服务调用占用的系统负载, L 为系统总的负载能力.

2.3 SIAC 服务调度模型

假定 SIAC 中包含了 n 个成员服务, 就可以得到如下的执行时间(T)、可用性(A)、可靠性(Rel)和执行代价(C)的关系:

$$\begin{cases} t = t(t_1, t_2, \dots, t_n) \\ a = a(a_1, a_2, \dots, a_n) \\ rel = rel(r_1, r_2, \dots, r_n) \\ c = c(c_1, c_2, \dots, c_n) \end{cases} \quad (1)$$

其中, t_i, a_i, rel_i, c_i 分别是第 i 个服务的执行时间、可用性、可靠性和执行代价; t, a, rel, c 分别是加和、连乘、Max 等运算的复合.

假定用户对 SIAC 的总的服务要求为: $t < T, c < C, rel > REL, a > A$ 其中 T, A, Rel, C 为常数. 结合上述工作流与成员服务的服务质量参数之间的关系(式(1)), 可得如下不等式组:

$$\begin{cases} t(t_1, t_2, \dots, t_n) < T \\ a(a_1, a_2, \dots, a_n) > A \\ rel(r_1, r_2, \dots, r_n) > REL \\ c(c_1, c_2, \dots, c_n) < C \end{cases} \quad (2)$$

每一个成员服务在网格中都存在一定数量的候选服务, 假定所有成员服务的候选服务的 QoS 参数都注册在军事信息栅格(MIG)之中, 通过 MIG 提供的查询接口, 可以获得所有成员服务的 QoS 信息. 记第 i 个服务 r 候选服务数量为 m_i , 则 SIAC 调度过程就是从所有成员服务的候选服务中选择合适的服务, 使得能够满足(2)式所列出的执行要求. 这样, SIAC 的调度问题就转变成(2)的组合优化问题.

3 SIAC 服务调度问题求解算法

3.1 SIAC 服务质量估算算法

考虑 SIS 的一般过程, 选择表 1 中的顺序、选择、循环、并行四种服务组合方式, 用 T_i, A_i, Rel_i, C_i 分别表示某组合类型中第 i 个成员服务 a_i 的执行时间、可用性值、可靠性值和执行代价(循环活动以 a 表示). 则可以得到服务组合的服务质量计算公式^[12].

表 2 SIAC 服务质量计算公式

组合方式	顺序组合(sequence)	选择组合(switch)	循环组合(Loop)	并行组合(Parallel)
执行时间(T)	$\sum_{i=1}^N T_i$	$\sum_{i=1}^N P_i \cdot T_i$	$T_a / (1 - p)$	$\max_{1 \leq i \leq N} T_i$
可用性(A)	$\prod_{i=1}^N A_i$	$\sum_{i=1}^N P_i \cdot A_i$	$(1 - p) \cdot A_a / (1 - p \cdot A_a)$	$\prod_{i=1}^N A_i$
可靠性(Rel)	$\prod_{i=1}^N Rel_i$	$\sum_{i=1}^N P_i \cdot Rel_i$	$(1 - p) \cdot Rel_a / (1 - p \cdot Rel_a)$	$\prod_{i=1}^N Rel_i$
执行代价(C)	$\sum_{i=1}^N C_i$	$\sum_{i=1}^N P_i \cdot C_i$	$C_a / (1 - p)$	$\sum_{i=1}^N C_i$

其中, 选择关系和循环关系中涉及了转移概率 p_i 和循环概率 p , 该参数应由应用链设计人员给出. 但是转移概率的值, 往往很难给出. 所以这种方法一般适用于应用链初次运行的时候. 在应用链运行一定时间以后, 转移概率通过应该使用转移频率(可由应用链执

行日志计算得出)来代替. 应用链运行的次数越多, 转移概率就越准确. SIAC 的服务质量可通过以下算法求

表 3 SIAC 服务质量求解算法

SIAC 的服务执行时间求解算法	
1	GetExecutionTime
2	{
3	executionTime = 0; //赋初值。
4	While(SIAC 中的成员服务) //对成员服务进行遍历,获取成员服务之间的连接关系类型。
5	{
6	if (connection = sequence) //顺序关系
7	{
8	GetMember{s ₁ ,s ₂ ,...,s _n }; //获取成员服务
9	executionTime=Σ {getExecutionTime(S ₁),..., getExecutionTime(S _N)};
10	}
11	else if (connection = switch) //选择关系
12	{
13	GetMember{(s ₁ ,p ₁),(s ₂ ,p ₂),(s _n ,p _n)} //获取所有成员服务及转移概率;
14	executionTime=Σ { p ₁ × ..., p _n × getExecutionTime(S _N)};
15	}
16	else if (connection = loop) //循环关系
17	{
18	GetMember{(s ₁ ,p ₁),(s ₂ ,p ₂),(s _n ,p _n)} //获取所有成员服务及循环概率;
19	executionTime= getExecutionTime(S)/(1-p);
20	}
21	else if (connection = parallel) //并行关系
22	{
23	GetMember{s ₁ ,s ₂ ,...,s _n };
24	executionTime=Max {getExecutionTime(S ₁),..., getExecutionTime(S _N)};
25	}
26	}
27	return executionTime .
28	}

该算法的复杂性与成员服务的嵌套层次有关, 假设应用链成员服务的最大嵌套层次为 l , 则该算法的复杂性为 $O(l)$. 可靠性、可用性和执行代价的求解算法与此类似, 这里不再重复.

以上提出了一种应用链服务质量的估算算法, 解决了在已知成员服务的服务质量参数的情况下, 计算相应的应用链服务质量参数的问题, 利用该算法可以建立应用链的服务质量要求与成员服务的服务质量要求之间的关系. 因为该算法最终要递归到对网格中服务的相应服务质量参数的求解上来, 所以执行该算法时, 每需要一个服务的相应的服务质量参数值, 就以未知变量来替代, 这样在算法执行结束后, 就可以建

立起应用链的服务质量参数与成员服务质量参数之间的关系.

3.2 基于 GA 的 SIAC 服务调度算法

当各节点的候选服务较少时, 可采用穷举法来求解 SIAC 调度问题. 然而当问题规模逐渐变大时, 穷举法的搜索空间将急剧扩大, 采用穷举法求解就不切实际了. 这类非线性组合优化问题一般属于 NP 难问题, 遗传算法(Genetic Algorithm, GA)已经被证实寻求满意解的最佳工具之一, 可以考虑采用遗传算法来进行求解^[13].

(1) 染色体编码

本文采用二进制形式对染色体进行编码, 对于具

有 n 个成员服务的工作流, 染色体被划分为 n 段, 其中每一个分段为相应成员服务的候选服务位置. 如 $xxxx|xxx|xxxx$ 是具有 3 个成员服务的编码. 每一个分段为相应成员服务的候选服务的编码, 如 $1101|101|0111$ 代表第 1 个成员服务的第 13 个候选服务、第 2 个成员服务的第 5 个候选服务和第 3 个成员服务的第 7 个候选服务.

(2) 种群初始化

设置种群规模, 并得到一组初始化解作为随后进行进化的基础. 例如设置种群规模为 20, 则最初的 20 条解, 无论其目标值为多少, 均可被接受.

采用随机方式来产生初始种群, 一般情况下种群的数量视搜索的大小 ($\prod_{i=1}^n m_i$) 而确定, 其取值在 10~160 之间浮动.

(3) 适应度函数

将(2)式转变为一个单目标的组合优化问题, 可以得到目标优化函数:

$$\begin{aligned} \min \text{obj} = & t(t_1, t_2, \dots, t_n) / T + c(c_1, c_2, \dots, c_n) / C \\ & - \text{rel}(\text{rel}_1, \text{rel}_2, \dots, \text{rel}_n) \\ & - a(a_1, a_2, \dots, a_n) / A \end{aligned} \quad (3)$$

一旦第 i 个成员服务的候选服务 j 确定以后, 可以通过查表得到 $\text{rel}_i, a_i, t_i, c_i$ 的值.

(4) 终止条件设定

$$\begin{cases} t_1 + 0.4t_2 + 0.4t_3 + 0.6t_4 + 0.6\text{MAX}(t_5, t_6) + 5t_7 < 600 \\ c_1 + 0.4c_2 + 0.4c_3 + 0.6c_4 + 0.6c_5 + 0.6c_6 + 5c_7 < 50 \\ \text{rel}_1 \times (0.4\text{rel}_2 \times \text{rel}_3 + 0.6\text{rel}_4 \times \text{rel}_5 \times \text{rel}_6) \times 0.8\text{rel}_7 / (1 - 0.2\text{rel}_7) > 0.8 \\ a_1 \times (0.4a_2 \times a_3 + 0.6a_4 \times a_5 \times a_6) \times 0.8a_7 / (1 - 0.2a_7) > 0.7 \end{cases} \quad (4)$$

令:

$$\begin{aligned} I_1 &= (t_1 + 0.4t_2 + 0.4t_3 + 0.6t_4 + 0.6\text{MAX}(t_5, t_6) + 5t_7) / 600 \\ I_2 &= (c_1 + 0.4c_2 + 0.4c_3 + 0.6c_4 + 0.6c_5 + 0.6c_6 + 5c_7) / 50 \\ I_3 &= (\text{rel}_1 \times (0.4\text{rel}_2 \times \text{rel}_3 + 0.6\text{rel}_4 \times \text{rel}_5 \times \text{rel}_6) \times 0.8\text{rel}_7 / (1 - 0.2\text{rel}_7)) / 0.8 \\ I_4 &= (a_1 \times (0.4a_2 \times a_3 + 0.6a_4 \times a_5 \times a_6) \times 0.8a_7 / (1 - 0.2a_7)) / 0.7 \end{aligned}$$

则适应度函数可表示为: $\min f = I_1 + I_2 - I_3 - I_4$.

通过仿真, 可以看到适应度函数收敛的情况, 仿真结果如图 2 所示. 可以看出: 在算法运行前期, 最佳适应度函数值与最差适应度函数值之间的差距较大, 平均适应度函数值变化较为明显; 随着算法的迭代, 在当种群进化到 20 代左右时, 适应度函数值的最佳值

由于遗传算法的迭代终止条件尚无定论, 本文采用(2)式作为迭代终止条件. 即当发现合适的染色体编码能够满足(2)式的要求, 则终止迭代过程. 或者当满意解不存在, 而群体中个体的进化又已趋于稳定, 则迭代终止.

(5) 选择

采用适应度比例方法来对种群的个体进行选择.

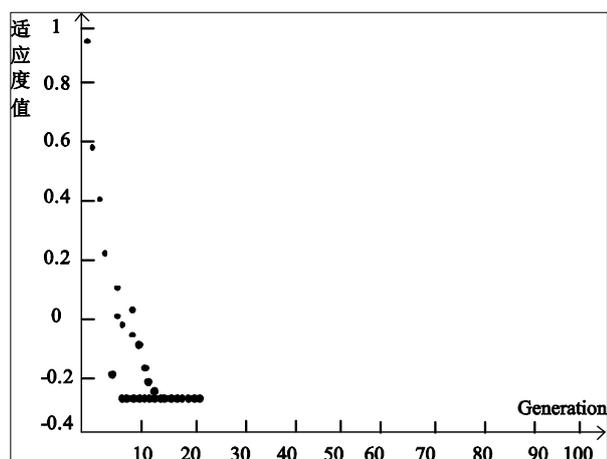
(6) 交叉和变异

本文选用二点交叉算子, 交叉概率为 p_c , 并以事先确定的变异概率 p_m 对染色体基因座的基因值进行变异.

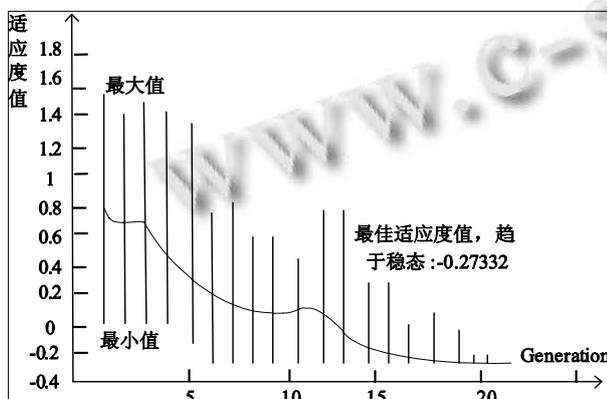
4 案例分析

在图 1(a)中所示的网络级 SIAC 服务模型中, 设置输入参数如下: 假定各节点的候选服务数量依次为 4、16、8、4、4、16、8, 各候选服务的服务质量参数可通过查表得到, 用户对应用链服务质量的总要求为 $t < 600, c < 500, \text{rel} > 0.8, a > 0.7$. 通过应用链运行的历史求得两个分支的转移概率分别为 0.4 和 0.6, 循环活动的执行概率为 0.8, 算法交叉和变异概率设置为: $p_c=0.8, p_m=0.09$, 其余参数值随机生成, 依据 3.1 节中的算法可建立如下不等式组:

与最差值之间的差距迅速变小(趋于 0), 平均值趋于平稳(-0.27), 也就是说种群中个体的进化已趋于稳态, 此时算法终止. 选择的候选服务序列为: 3,9,5,4,4,10,3, 计算可得卫星信息应用链的服务质量分别为:526,440,0.85,0.82, 满足用户的要求.



(a)



(b)

图2 适应度函数的收敛

5 结论

本文对卫星信息服务组合方式进行了分析,建立了卫星信息应用链服务质量参数体系和计算模型,并给出了相应的求解算法.基于服务质量,提出了卫星信息应用链服务调度问题的数学模型,并采用遗传算法进行了实现,最后通过一个案例进行了验证,证明了模型的正确性和服务质量作为卫星信息应用链性能评价标准的可行性,可以为卫星信息应用链服务能力研究提供评价依据.

下一步的工作包括两个方面:一是进一步考察SIAC的服务质量要求,建立更为完善、更加符合作战实际的服务质量参数体系;另一方面是把本文提出的服务质量计算方法和调度算法应用到SIAC的服务质量管理中去,并在实际应用中加以验证和改进.

参考文献

- 1 文江平.卫星军事应用技术.北京:国防工业出版社.2007:470-517.
- 2 王建江,邱涤珊,彭黎.基于排队网络的空间信息数据处理系统效能评估.计算机应用,2012,32(3):870-873.
- 3 Ng KYK, Ghanmi A, Lam MN. Workforce configuration and workflow analysis of an information technology organization: a queuing network approach. IEEE T ENG MANAGE, 2002, 32(1): 724-732.
- 4 Ghanmi A. Modeling and analysis of a Canadian Forces Geomatics division workflow. European Journal of Operational Research, 2006, 170: 1001-1016.
- 5 Ng KYK, Lam MN. Workflow analysis of an information technology organization: the case of the operational mode during a time of crisis. IEEE T ENG MANAGE, 2005, 52(3): 396-403.
- 6 于晓浩,胡丹,罗雪山,刘俊先.军事信息系统服务资源并行优化方法.系统工程理论与实践,2012,32(9):2078-2085.
- 7 于晓浩,刘俊先,罗雪山,张嘉焱.军事信息服务 QoS 模型及服务选择方法.系统工程,2010,28(9):91-97.
- 8 Zhang YM, Sheng Y. Scheduling algorithm with quality of service support in IEEE 802.16 networks. Information and Communications Technologies, 2009, 4: 44-49.
- 9 Sun J, Yao Y, Zhu H. Quality of service scheduling for 802.16 broad wireless access system. IEEE 63rd Vehicular Technology Conference. Melbourne, 2006, 3: 1221-1228.
- 10 Rangel V, Gómez J, Ortiz J. Performance analysis of QoS scheduling in broadband IEEE 802.16 based networks. Proc. of Opnetwork 2006 Technology Conference, Washington D.C, August 2006.
- 11 金海,陈汉文,吕志鹏,宁小敏.CGSP 作业管理器合成服务的 QoS 优化模型及其求解.计算机学报,2005,28(4):578-585.
- 12 王勇,胡春明,杜宗霞.服务质量感知的网格 workflow 调度.软件学报,2006,17(11):2341-2351.
- 13 Baker S, Bender A, Abbass H, Sarker R. A scenario-based evolutionary scheduling approach for assessing future supply chain fleet capabilities. Studies in Computational Intelligence, 2007, 49: 485-511.