

基于文化基因算法的多目标云任务调度研究

郝晓弘, 韩于芳, 晏 燕

(兰州理工大学 计算机通信学院, 甘肃 兰州 730050)

摘 要: 提出了基于禁忌搜索算法的文化基因算法, 从而提高整个云环境的执行效率. 最后在 CloudSim 模拟仿真平台上, 将其与基于遗传算法和基于爬山算法的文化基因算法进行了比较, 从而对基于禁忌搜索算法的文化基因算法的可行性和有效性进行实验验证.

关键词: 云计算; 任务调度; 多目标优化; 文化基因算法

中图分类号: TP31

文献标识码: A

文章编号: 1000-7180(2018)01-0061-05

Research on Multi-Objective Cloud Task Scheduling Based on Memetic Algorithm

HAO Xiao-hong, HAN Yu-fang, YAN Yan

(School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: The MA based on the tabu search algorithm is introduced to improve the MA based on the Hill Climbing Algorithm to improve the efficiency of the implementation of the entire cloud environment. Finally, on the CloudSim simulation platform, the feasibility and validity of Tabu search algorithm are verified by comparison with existing genetic algorithm and Memetic algorithm based on Hill Climbing algorithm.

Key words: cloud computing; task scheduling; multi-objective optimization; memetic algorithm

1 引言

随着数字化进程的快速推进以及移动互联网的快速发展, 数据量以 TB 甚至 PB 级的惊人速度增长着, 已经远远地超出了传统计算模式的处理能力. 基于此背景, 作为商业化计算模式的云计算便应运而生, 它将所有的应用程序、数据和有着巨大规模的 IT 资源, 以服务的方式, 通过无所不在的计算机网络, 提供给庞大的云用户群体使用^[1]. 用户可以访问大量的云计算和存储资源, 而不必去关心这些资源处于什么位置以及这些资源是如何配置的, 对于广大的云用户而言, 云端的资源是无限的. 所谓云任务调度就是在复杂的云环境下将用户提交的任务合理地配置在虚拟机上, 满足用户的所关心的执行时间、执行费用以及系统的负载均衡等条件. 因此, 云环境下的任务调度问题实际上是一个 NP-hard 的多目

标组合优化问题.

近年来, 国内外学者对有目标优化问题做了广泛的研究, 然而正如 Wolpert^[2]等人提出的“优化问题没有免费的午餐”理论, 任何算法不能完全保证在所有问题上得到比其他算法更好的结果. 因此, 应该针对不同的有目标优化问题, 根据其自身的特点, 设计不同的算法. 而文化基因算法就是这样一种算法框架, 它在算法设计过程中既包含了对所研究问题的全局搜索, 也结合了问题的局部搜索, 为不同类型的多目标优化问题研究提供了一个新的思路, 从而可以根据不同的问题, 快速而灵活地构建适合该问题的 Memetic 算法. 自 1989 年, PMoscato 首次提出 Memetic Algorithm 这一术语之后, 文化基因算法已经在各类工程优化问题中得到了很好的应用^[3]. 蔡斌^[4]等人将文化基因算法应用于多目标车间调度理论的研究中, 并将其应用于具体的精益管

理系统中,取得了良好的效果;潘铁文^[5]将文化基因算法应用于图像检索中,取得了较好的应用结果;龙志翔^[6]将文化基因算法应用于多目标的网格计算中,取得了不错的实验结果.

通过对以上文化基因算法的分析与研究不难发现,文化基因算法在解决多目标问题方面有一定的优势,受此启发,本文将改进遗传算法,引入爬山算法和禁忌搜索算法,将遗传算法较好的全局搜索能力以及爬山算法和禁忌搜索算法较好的局部寻优能力进行充分的融合,提出了基于爬山算法和禁忌搜索算法的 MA 算法,以用户所关心的执行时间和执行费用作为主要的优化目标,同时结合负载均衡作为其约束条件,从而解决云环境下多目标问题的云任务调度,最后在 CloudSim 仿真平台上,验证本文所采用的基于禁忌搜索算法的 MA 算法的可行性和有效性.

2 云任务调度模型

2.1 云任务模型

云环境下的任务模型^[7]如图 1 所示.

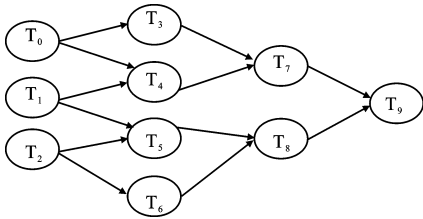


图 1 云任务模型

图 1 所示的 DAG 图就是用户所提交的任务经过云计算的框架分布式并行处理后的云任务形态, T_i 代表了用户所提交的任务, $T_0 \sim T_6$ 用来表示云任务流中的第一个节点, $T_7 \sim T_8$ 用来表示该任务流中第二个节点, T_9 用来表示该任务流中的第三个节点.

2.2 云任务调度模型

云环境下的任务调度模型如图 2 所示.

图 2 中,数据中心采用虚拟化技术,将所需要的计算资源的物理节点集转化为相应的虚拟机集群,随后将相关的属性信息转交给数据中心代理器进行处理.通过统一的用户接口,不同的用户将其需要处理的任务交给数据中心代理器进行相应的任务处理和调度.

所谓云环境下的任务调度,就是根据广大云用户提交的任务列表,对其进行归类,最后将其配置到相应的虚拟机上进行相应的任务调度处理,当将所

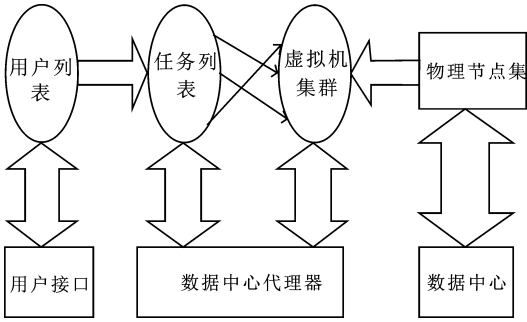


图 2 云任务调度模型

有任务分配完之后,虚拟机就开始进行相应的任务调度,则衡量一个任务调度方案优劣的标准就是其调度过程中所需要的各种开销(如:执行时间、执行费用、能量消耗等)等.通过对含有对于有 M 个虚拟机以及 N 个调度任务的云环境进行粗略的计算,大体可以得到 M^N 种之多的分配调度策略,然后对调度任务的属性、虚拟机的属性以及云环境下相应的组合优化目标进行分析,从而在 M^N 种调度方案中寻找满足条件的最优解,一旦任务被分配完成,该调度方案就会被数据中心代理器交给数据中心进行相应的任务调度处理.

2.3 基于多目标优化的云任务调度

云环境中各参数的数学描述如下所示.

(1) 任务列表:

$T = \{t_1, t_2, t_3, \dots, t_n\}$, 每个任务的数据总量为 Num_i .

(2) 虚拟机列表:

$\text{VM} = \{m_1, m_2, m_3, \dots, m_n\}$, 每个虚拟机的处理能力定义为 Mips , 每个虚拟机的价格为 price_v .

(3) 执行时间

总的执行时间为

$$\text{totaltime} = \max(\text{timeVM}_i) \tag{1}$$

单个虚拟机的执行时间为

$$\text{timeVM}_i = \frac{\sum \text{size}}{\text{Mips}} \tag{2}$$

由于云环境下的虚拟机是采用并行执行的方式,因而执行调度任务所耗费的时间就应该是每个进行任务处理的虚拟机中耗时最长的虚拟机所对应的执行时间, timeVM_i 为第 i 个虚拟机的执行时间, Mips 为第 i 个虚拟机的处理能为该虚拟机上的总的任务数据量.

(4) 执行费用

$$\text{totalcost} = \sum_{i=1}^m \text{price} \times \text{size} \tag{3}$$

式中,Price为单个虚拟机的执行费用,size为该虚拟机处理的任务数量.

(5) 负载均衡值

$$\text{load} = \sqrt{\sum_{i=1}^n \left(\text{Num}_i - \frac{n}{m} \right)^2} \quad (4)$$

由于云环境下的任务调度是采用并发执行的方式,因此,当虚拟机执行较多的任务之时,相应地,该虚拟机的负载就越大.反之,当任务分配越均匀时,其对应地整个云系统的负载就越均衡.

本文的负载均衡值参照虚拟机负载的标准差,如果该虚拟机负载的标准差越小,那么其对应地云系统的负载就越均衡,反之亦然.

3 基于文化基因算法的多目标云任务调度

云环境下的任务调度,本质上而言,就是多目标的组合优化问题,其主要目标是将用户提交的任务合理的分配到虚拟机上,既兼顾用户的服务质量(Quality of Service, QoS)请求(执行时间、执行费用、能源消耗等),又满足云服务提供商的利益(负载均衡),最终达到资源的高效利用.本文融合遗传算法较好的全局寻优能力以及禁忌搜索算法较好的局部寻优能力,改进文化基因算法,将其应用到多目标的云任务调度中.

3.1 适应度函数

因为云环境下的执行时间和执行成本成本的目标函数值不在一个数量级水平上,因而先规范化处理这两个目标函数值,再通过加权的方法对适应度函数进行设计^[8-9].

假设种群规模为 R ,那么每一代的个体数量为 R .

执行时间函数:

$$f_T(i) = \frac{1/\text{totaltime}(i)}{\sum_{j=1}^r \text{totaltime}(j)} \quad (5)$$

式中,totaltime(i)代表虚拟机 i 在执行调度任务时所耗费的时间,该 totaltime(i) 的值越小,则其对应的函数值就越大.

执行费用函数:

$$f_c(i) = \frac{1/\text{totalcost}(i)}{\sum_{j=1}^r \text{totalcost}(j)} \quad (6)$$

式中,totalcost(i)代表虚拟机 i 执行调度任务时所耗费的成本,该 totalcost(i) 的值越小,则其对应的函数值就越大.

规范化处理执行时间和执行费用之后,其对应

的目标函数的取值分布在区间 $[0, 1]$ 上.

适应度函数:

$$F(i) = \omega f_T(i) + (1 - \omega) f_c(i) \quad (7)$$

式中, ω 位于 $[0, 1]$, ω 表示用户对执行时间的依赖程度.

3.2 本文算法描述

3.2.1 染色体编码

在云环境下的任务调度之中,调度任务的分配方式以染色体的编码方式来呈现,而染色体的编码方式却众说纷纭,因而,本算法采用文献[10]中的资源-任务的间接编码方式最为基于文化基因算法的多目标云任务调度的染色体编码方式.

在染色体的编码方式中,用户提交的任务数量等同于需要执行编码的基因数量,因而,云环境中的虚拟机数量得大于大于基因的数量.假设第 i 个基因的值为 j ,那么就代表第 j 个虚拟机在执行第 i 任务,如果 DAG 任务图中的任务数目是 6,执行调度的虚拟机数目为 3,则染色体编码为: $x = \{1, 1, 2, 2, 3, 3\}$.

上式中关于 x 的染色体集合表示其前两个任务被分配到了第 1 个虚拟机上,第 3、4 个任务被分配到了第 2 个虚拟机上,以这样的方式进行推理,直到所有的任务被分配完成.

当对该染色体执行解码操作时,则能够清楚地知道每个虚拟机上执行的任务情况,即: $\text{Num}_1 = \{1, 2\}; \text{Num}_2 = \{3, 4\}; \text{Num}_3 = \{5, 6\}$.

由此可以看出 Num_i 表示第 i 个虚拟机上所执行的任务集合.通过对每个染色体执行解码操作,就可以获取每个虚拟机上调度任务的相关信息.

3.2.2 遗传操作

(1) 选择操作

所谓轮盘赌法,又称为比例选择法,其基本思想就是个体选中的概率与其适应度函数的大小成正比.由于对目标函数进行了规范化处理,所以群体中个体的适应值相差不是很大,不会造成算法过早地陷入局部最优.

假设种群大小为 G ,个体 i 的适应度表达为 $F(i)$,则个体 i 被选中的遗传概率为

$$p_i = F(i) / \sum_{i=1}^r F(i) \quad (8)$$

(2) 交叉操作

为了避免早熟现象的发生,加快算法的收敛速度,本文采用自适应的交叉算法,在算法开始阶段,首先按较大的交叉概率替换父代中某些个体位置,

从而防止算法过早的收敛,当遗传算法执行到后期时,适当地降低交叉概率,加快算法的收敛速度,从而更快地产生优良个体.

(3)变异操作

本文采用多点变异的方法来执行遗传算法的搜索,通过随机的方式决定需要需要变异的基因位置,最后以选择操作的方式将基因保留到下一代.

3.2.3 禁忌搜索算法

禁忌搜索算法 (Tabu Search Algorithm, TS)^[11] 是 Fred Glover 于 1986 年提出的一种解决优化问题的元启发式算法,其主要思想是按照人类思维的过程,将搜索到的局部最优解存储到禁忌算法的禁忌表中,对该求得的解添加标记,从而避免迂回的搜索,进而跳出局部最优这一局限,提高算法的性能. 其对应的算法流程图如图 3 所示.

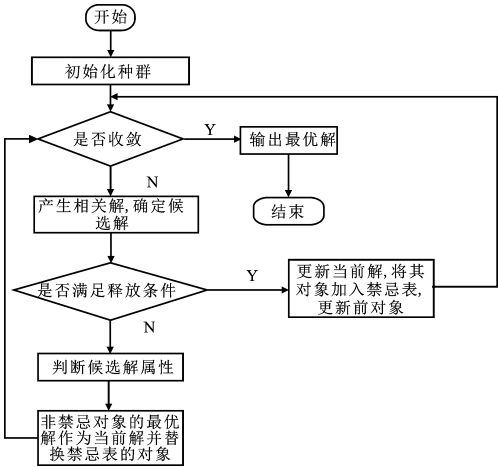


图 3 禁忌搜索算法

3.2.4 基于禁忌搜索算法的文化基因算法

文化基因算法是在模拟文化进化的基础上演变而来的,是一种结合种群全局搜索以及基于个体局部搜索的进化算法. 它和遗传算法在流程上有很大的相似之处,只不过它在交叉变异之后多了一个局部寻优的过程. 文化基因算法充分吸收了遗传算法的全局搜索能力和禁忌搜索算法较强的局部寻优能力,避免了迂回往复的无用搜索,尽快地跳出局部最优,提高了算法的整体性能,其对应的具体算法流程图如图 4 所示.

其对应的算法步骤如下所示.

Step1: 初始化云计算环境中的各参数配置;

Step2: 随机产生可调度任务序列. 初始化整个种群 G , 进化迭代次数 $t=0$;

Step3: 根据式 (7) 对种群 $G(t)$ 进行适应度值评估;

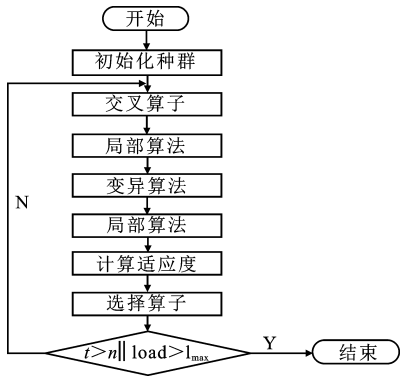


图 4 基于禁忌搜索算法的文化基因算法

Step4: 判断是否满足终止条件, 输出最终解, 结束算法;

Step5: 若不满足终止条件, 对种群 $G(t)$ 进行交叉操作, 并对交叉之后的非支配解进行局部搜索算法(爬山算法和禁忌搜索算法), 产生 $G(t)$;

Step6: 执行变异操作和局部搜索算法(爬山算法和禁忌搜索算法);

Step7: 根据式 (8) 对种群 $G(t)$ 执行轮盘赌法的选择操作, 并将结果保存到 $G(t)$ 之中, 产生新的子种群 $G(t)$, 转到 step3;

4 仿真测试与结果分析

4.1 实验环境及参数配置

为了验证本文所采用的基于禁忌搜索算法的文化基因算法的有效性和可行性, 在 CloudSim 仿真平台上进行了仿真实验. 实验环境: Intel(R) Core (TM) i3-2330M, 2.20 GHz 主频, 8.00 GB 内存.

Cloudsim 仿真程序对 40 台物理机进行模拟, 任务数分别设置为 50, 100, 200, 300, 400. 设置种群规模为 100, 迭代次数 50, 交叉率 0.55, 变异率 0.2.

4.2 实验结果

为了对本文改进的 Memetic 算法的云任务调度策略的有效性和可行性进行评估, 在 CloudSim 仿真平台上对本文改进的 Memetic 算法进行模拟实验, 分别将本文改进的 MA(HC) 和 MA(TS) 算法与传统的 GA 算法进行比较, 从而验证本文改进的 MA(Tabu) 算法的优越性.

4.2.1 有效性证明

为了验证 MA(TS) 算法的优越性, 将本文算法与基于爬山算法的文化基因算法、遗传算法进行实验对比, 均使用同一组输入数据, 三种算法求得的负载均衡率如图 5 所示.

从图 5 可以清晰地看出, 当参与调度的任务数

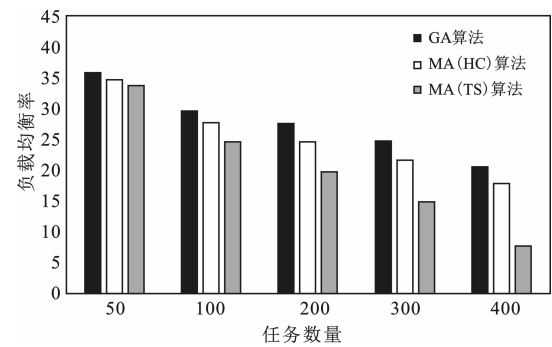


图 5 负载均衡率对比

量较少时,三种算法的负载均衡率的差异不是很大,没有表现出明显的规律性,这是由于当云任务规模较小时,传统的 GA 算法也可以快速地搜索出满意解.但当云任务数量增加时,基于 MA(TS)算法的云任务调度策略的负载均衡水平迅速收敛,并快速超越了基于 GA 算法和 MA(HC)算法的云任务调度算法,这是由于随着参与调度的任务数量的增加,资源分配的随机性降低时,GA 算法的局部求解能力较差,并且随着任务数的增加,该算法易于陷于局部最优而出现退化现象;虽然 MA(HC)算法在 GA 算法的基础上,增加了爬山算法这一局部搜索算法,但是由于爬山算法的的无记忆性,可能在搜索过程中出现局部最大解,到达某高低而失去最佳搜索方向,以及在山脊的两面来回震荡,从而无法执行全面的搜索过程,降低整个算法的搜索效率,从而使其在负载均衡的指标中表现的不是很明显;而 MA(TS)算法充分利用禁忌搜索这一局部启发式算法的优点,将搜索到的局部最优解存储到禁忌算法的禁忌表中,对该求得的解添加标记,从而避免迂回的搜索,进而跳出局部最优这一局限,提高整个算法的性能,从而避免了性能好的节点承担过重的负载,而其他节点负载却过轻,造成资源的浪费,影响云环境下调度任务的高效执行.

4.2.2 执行效率的验证

为了验证基于禁忌搜索算法的文化基因算法的有效性,将本文算法与 MA(HC)和 GA 算法进行实验对比,三组算法均使用同一组输入数据,其对应的执行时间和执行费用的对比分别如图 6 和图 7 所示.

观察图 6 和图 7,可以清晰地看出:当任务数量大于 100 时,基于 MA(TS)算法的云任务调度算法的执行时间和执行费用明显低于基于 GA 算法和 MA(HC)算法的云任务调度算法,这是由于 GA 算法较差的局部求解能力,使其在任务数增多

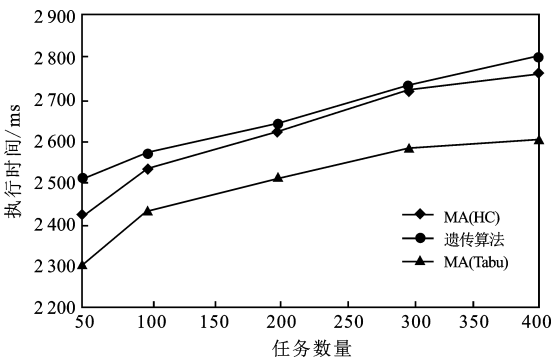


图 6 执行时间对比

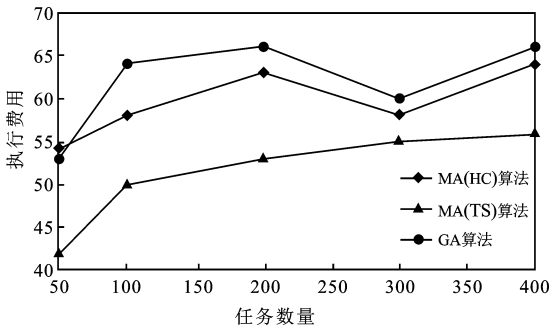


图 7 执行费用对比

的情况下,极易陷入局部最优而出现退化现象;虽然 MA(HC)算法和 MA(TS)都克服了 GA 算法较差的局部求解能力,但是由于爬山算法的的无记忆性,致使其无法执行全面的搜索过程,降低整个算法的搜索效率,从而使其性能明显低于 MA(TS)算法.

综上所述:本文的 MA(TS)算法受任务数量的影响不是很大,执行时间和执行费用维持在一个可控的范围之内,具有良好的稳定性,适合大规模任务的云调度.

5 结束语

本文研究了基于禁忌搜索算法文化基因算法的云任务调度,建立了以执行时间和执行费用为目标函数,以负载均衡为约束条件的多目标云任务调度模型,充分利用了禁忌搜索算法的局部寻优能力和遗传算法的全局搜索能力,防止了算法过早的收敛,增加了种群的多样性.最后在 CloudSim 仿真平台上通过实验证明,该算法在总体性能上优于遗传算法和基于爬山算法的文化基因算法,且适合大规模的云任务调度.

(下转第 71 页)

通过对程序的依赖关系进行分析,进而提取了程序中的循环并对循环进行了分析,提出了一种图像处理类并行性分类的方案,能够帮助编程人员解决如何使用混合架构来实现对图像处理类应用程序的加速,并采用编译器架构对分类进行了初步的实现。

参考文献:

- [1] 代媛. 基于 GPU 并行计算的图像压缩技术研究[D]. 杨凌:西北农林科技大学,2014.
- [2] Silva B D, An B, DHollander E H, et al. Comparing and combining GPU and FPGA accelerators in an image processing context[C]// International Conference on Field Programmable Logic and Applications. IEEE, 2013:1-4.
- [3] 张玮雄, 刘建霞. 数字图像处理技术的发展现状及趋势[J]. 科学之友, 2012,33(6):153.
- [4] 龚志远. 基于图像分析技术的轴承故障诊断[J]. 煤矿机械,2005 ,26(11) :171-173.
- [5] 郭启军,潘炼,钱新恩. 基于图象识别的高抛光透射面

缺陷检测[J]. 计算技术与自动化, 2008, 27 (4) : 95-97.

- [6] 许琳娟,曹文洪,刘春晶,等. 图像处理技术在推移质运动颗粒参数提取中的应用[J]. 长江科学院学报, 2017, 34 (1) :1-5.
- [7] Gonzalez R C, Woods R E. 冈萨雷斯数字图像处理[M]. 北京,电子工业出版社,2003.

作者简介:

王时雨 男,(1984-),博士研究生. 研究方向为微处理器设计、雷达成像、视频处理.

E-mail: onion0709@mail. nwpu. edu. cn.

张盛兵 男,(1968-),博士,教授. 研究方向为嵌入式微处理器的体系结构与微处理器芯片.

安建峰 男,(1977-),博士,副教授. 研究方向为高性能计算机体系结构、嵌入式系统软硬件协同设计.

黄小平 男,(1979-),博士,讲师. 研究方向为计算机体系结构.

白弘骥 男,(1992-),硕士. 研究方向为计算机体系结构.

(上接第 65 页)

参考文献:

- [1] ARMBRUST M, FOX A, GRIFFITH R, et al. A view of cloud computing [J]. Communications of the ACM, 2010, 53(4): 50-58.
- [2] Wolpert D H, Macready W G. No free lunch theorems for optimization[J]. IEEE Transactions on Evolutionary Communication, 1997,1(1): 67-82.
- [3] 刘漫丹. 文化基因算法(Memetic Algorithm)研究进展[J]. 自动化技术与应用, 2007, 26(11):1-4.
- [4] 蔡斌. 基于文化基因算法的车间作业调度理论研究及实践[D]. 重庆:重庆大学, 2012.
- [5] 潘铁文. 基于文化基因算法的图像检索研究[D]. 哈尔滨: 哈尔滨工程大学,2010.
- [6] 龙志翔. 多目标 Memetic 算法在网格作业调度中的应用研究[D]. 武汉:武汉理工大学, 2011.
- [7] 黄婷婷, 梁意文. 云工作流任务调度的模拟退火遗传改进算法[J]. 微电子学与计算机, 2016, 33(1): 41-46.

[8] 黄伟建, 郭芳. 基于烟花算法的云计算多目标任务调度[J]. 计算机应用研究, 2017 , 34 (6) :1718-1720.

[9] 黄超, 胡德敏, 余星. 多目标遗传算法在云计算任务调度中的应用[J]. 信息技术, 2014,38(5):130-134.

[10] 刘之家, 张体荣, 谢雄程. 基于云计算的"用户期待"任务调度算法的研究[J]. 大众科技, 2011,13(4): 75-77.

[11] Chuan-Kang Ting, Cheng-Feng Ko, Chih-Hui Huang. Selecting survivors in genetic algorithm using tabu search strategies [J]. Memetic Comp, 2009 (1): 191-203.

作者简介:

郝晓弘 男,(1960-),硕士,教授,博士生导师. 研究方向为迭代学习控制和智能信息处理.

韩于芳(通讯作者) 女,(1991-),硕士. 研究方向为云任务调度理论及应用. E-mail: cicyyufang@163. com.

晏燕 女,(1980-),博士研究生,副教授. 研究方向为多媒体信息安全、位置隐私保护.