

## 论文题目

---

Scheduling Using Improved Genetic Algorithm in Cloud Computing for Independent Tasks

---

## 论文作者

---

Pardeep Kumar and Amandeep Verma

---

## 发表期刊信息

---

ICACCI 12 Proceedings of the International Conference on Advances in Computing, Communications and Informatics. Pages 137-142. 2012

---

## 技术问题

---

云计算中几乎所有的东西，如硬件、软件 and 平台都是可以作为服务提供的，因此云计算受到了广泛欢迎。云计算中的服务按每次使用付费的方式向用户收费，提供商根据客户端请求提供服务。云计算中一个重要的问题是用户请求的调度，即如何将资源分配给这些请求，以便根据用户定义的时间在最短的时间内完成请求的任务。良好的调度技术有助于有效利用资源。本文讨论了min-min、max-min和GA三种调度技术，给出了min-min和max-min的性能指标。通过实例验证了标准遗传算法和改进遗传算法的性能并提出了一种将min-min和max-min相结合的遗传算法。

---

## 现实背景

---

云计算的流行

云计算已经成为高性能计算的一个被广泛接受的范例，因为所有类型的IT设施都作为服务提供给用户——包括IaaS、PaaS和SaaS。云服务是通过互联网提供的。希望访问云服务的设备应该具有Internet访问功能。设备需要非常小的内存，非常轻的操作系统和浏览器。云计算提供了很多好处：它可以节省成本，因为不需要初始安装很多资源；它提供了可扩展性和灵活性，用户可以根据需要增加或减少服务的数量；维护成本非常低，因为所有资源都由云提供商管理。

云 workflow 中的调度（论文中并没有讲这个背景，是其他文献中查阅的）

云 workflow 中，如何在各种限制情况下合理地将任务分配给资源（任务调度）是一个备受关注的课题。云中任务有很多的限制因素，如完工期、成本、吞吐量、资源效用、可信度、周转时间、成功率、延迟、负载均衡、资源可用、预算、截止日期、等待时间和安全性等等。workflow 中的任务又有执行的先后顺序。因此云中的调度方案不能生搬硬套其他的调度方案如FIFO、轮询等等。

---

## 作者思路

先讲述了以前在相关领域所做的相关工作，比如调度技术：min-min、max-min和GA。然后提出了改进的遗传算法并给出了新算法的仿真结果，以及改进遗传算法与标准遗传算法的性能比较。

---

## 解决方案

### 巨人之肩

作者提出 Enhanced GA 之前，先回顾了启发式调度三巨头，min-min、max-min和GA。

- **Min-min**算法以一组所有未分配的任务开始。首先，找出所有任务的最小完成时间。然后在这些最小时间中选择最小值，即任意资源上所有任务的最小时间。然后根据这个最小时间，在相应的机器上调度任务。然后所有其他任务的执行时间更新，机器上通过添加分配的任务的执行时间，机器上的其他任务的执行时间和分配任务从列表中删除的任务分配给机器。然后再次执行相同的过程，直到在资源上分配了所有任务。
- **Max-min**算法在找出最小执行时间后选择最大值，即任意资源上所有任务的最大时间。然后根据最大时间，在相应的机器上调度任务。然后所有其他任务的执行时间更新，机器上通过添加分配的任务的执行时间，机器上的其他任务的执行时间和分配任务从列表中删除的任务分配给机器。然后再次执行相同的过程，直到在资源上分配了所有任务。
- **GA**算法是通过种群进化的生物学概念进行调度的方案，具体步骤如下。个人感觉如何合理设置适应度函数和自然选择的合适概率是决定调度算法好坏的关键。
  - **种群初始化**：初始种群是遗传算法中用来寻找最优解的所有个体的集合。群体中的每一个

解都被称为个体。每个个体都被表示为一个染色体使其适合于基因操作。从最初的种群中选出个体，并对其进行一些操作，形成下一代。

- **适应度函数**：适应度函数是根据给定的优化目标，利用适应度函数来衡量群体中个体的质量。适应度函数可以根据不同的情况而有所不同。在某些情况下，适应度函数可以基于完工时间，而在某些情况下，适应度函数可以基于预算约束。
- **自然选择**：我们使用比例选择算子来确定群体中不同个体遗传给下一代的概率。比例选择算子是指所选择的概率与遗传给下一代群体的概率与个体健康程度成正比。
- **交叉（杂交）**：我们使用单点交叉算子。单点交叉意味着在单个代码中只有一个交叉，在那个点上，一对染色体的一部分交换了。

## 提升方案

其实这个论文没提出什么特别有贡献的方案，也就是GA算法初始化之前先用min-min、max-min得到初始化种群。所以本质上来说也不算是个hybridization，还是启发式的。

## 直观表述

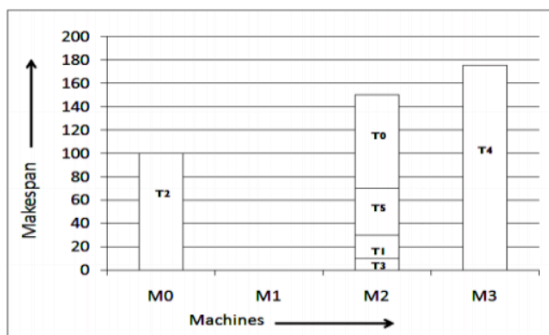


Figure 1. Task Assignment by Min-Min Algorithm

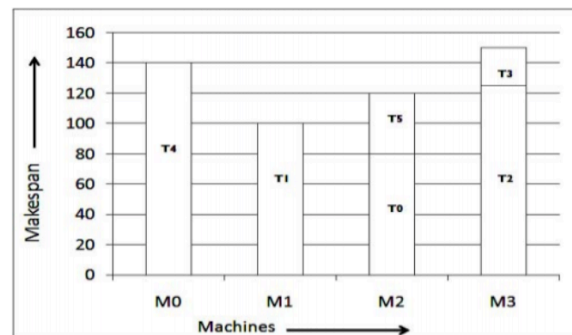


Figure 2. Task Assignment by Max-Min Algorithm

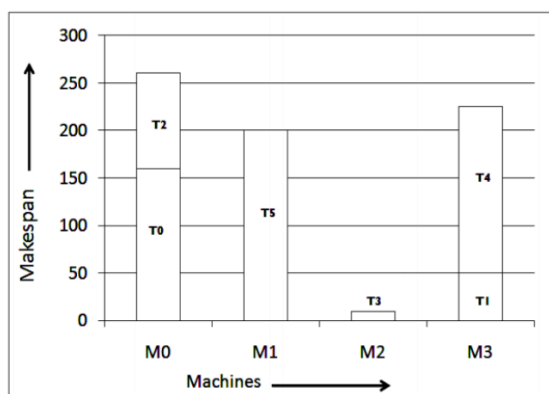


Figure 3. Task assignment by Standard Genetic Algorithm

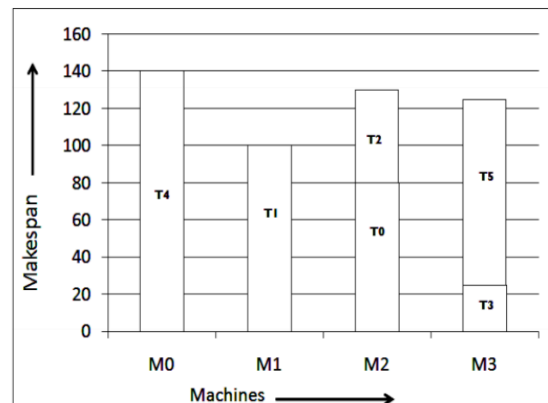
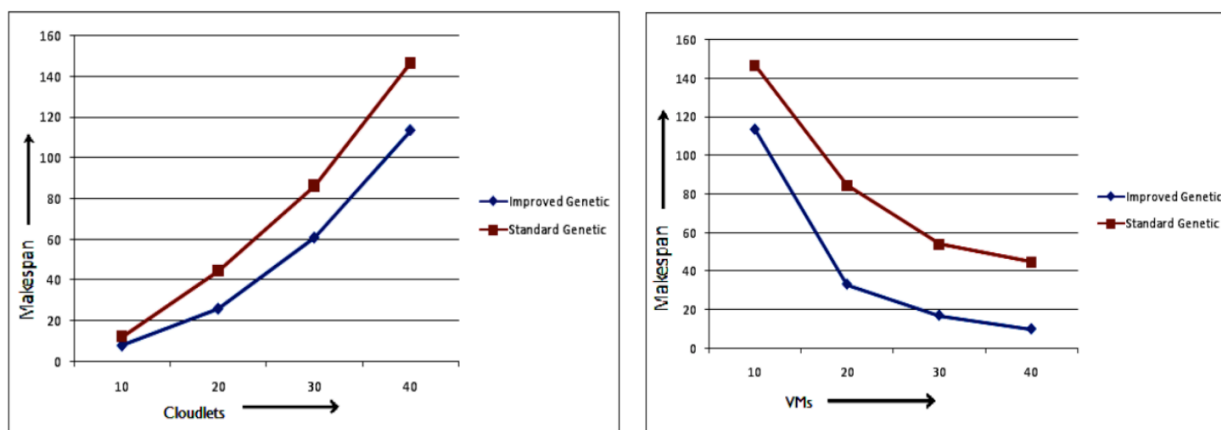


Figure 4. Task assignment by Improved Genetic Algorithm

---

## 效果评价

论文使用CloudSim作为模拟器来验证改进算法和标准遗传算法的性能。实验中，虚拟机是资源，而cloudlet是任务。进行了两个实验：固定虚拟机数量，改变cloudlet的数量；固定了cloudlet的数量，改变虚拟机的数量。



---

## 创新贡献

- 把启发式三巨头融合了；
- GA初始化前用了min-min，max-min；
- makespan 变小了。

---

## 个人感想

### 本文的优势

这个文章前部分介绍启发式三个算法还是很通俗易懂的，画的直观表示图很容易理解，可以算个入门好文。

### 本文的劣势

这个文章算法虽然让完工期短了，但是at the expense of时间复杂度。而且改进很简单，基本上就是两种方法前后顺序运行了一下，没啥实际贡献。个人感觉improvement应该多在适应度函数上做文章，比如融合一下min-min和max-min的思想，或者在自然选择概率上有启发式地向min-min和max-min迁移（比如设计一个概率），这样更有实效。

---