

# 论文题目

---

Scheduling scientific workflow tasks in cloud using swarm intelligence

---

# 论文作者

---

V. Vinothina

---

# 发表期刊信息

---

*2017 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC)*,  
Bangalore, 2017, pp. 1-5.

---

# 技术问题

---

云计算的主要挑战之一是任务调度和云资源的有效利用。一般来说，任务调度问题是NP难问题，而获得最优解是一项具有挑战性的任务。由于云的异构性、动态性和自治性，使得云的调度问题十分复杂。由于云使用的成本，用户必须有效地利用云。科学工作流任务调度是科学应用中的任务调度问题，它是根据一些性能影响因素将每个任务映射到合适的资源上的问题。影响资源利用率的因素主要取决于任务资源调度和分配的算法和实践。

---

# 现实背景

---

最先进的云计算技术之一是分布式计算系统的类型及其特性。它具有巨大的效益和特点，得到了各组织的一致好评。可扩展性和异构性特性使云最适合计算科学的工作流任务，因为工作流包含数千个任务，处理大量数据。基础设施即服务（IaaS）、平台即服务（PaaS）和软件即服务（SaaS）是云的部署模型，允许用户使用按使用付费的策略在云中托管自己的应用程序。工作流应用在科学和工程中很常见，其结构/组成是预先知道的。它可以是服务、应用程序或模块。工作流中每个任务的资源需求取决于其功能能力和数据输入。云的虚拟服务器降低了用户购买、操作和维护物理计算基础设施的成本。此外，虚拟化技术允许在一台物理机上使用多个虚拟机。这使得资源共享更加优化，资源利用更加充分。

---

# 作者思路

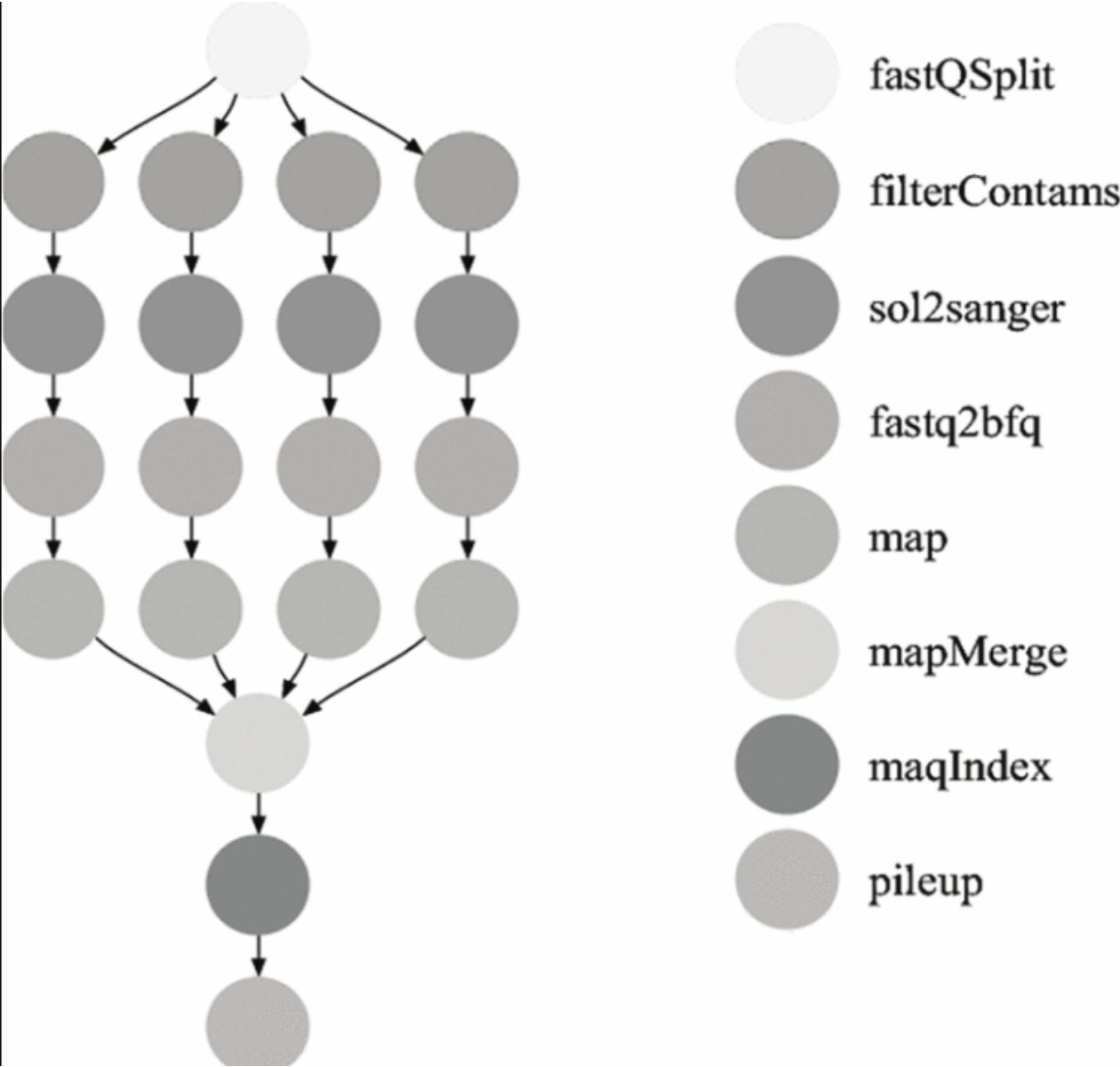
---

在针对云端 workflows 任务的最小生成时间、最小成本、最大资源利用率等不同目标，提出了多种不同的调度算法后，作者在此基础上提出了一种改进的云端 workflow 调度算法。作者之前的工作使用了ACO方法来最小化科学工作流程的完成时间。本文旨在提出一种克服WFSACO算法复杂度的算法。作者给出一些相关的准备工作，之后指出了之前的WFSACO算法的复杂之处，并提出了IWSACO算法的模型和描述。最后对该算法进行实验并进行性能评估。

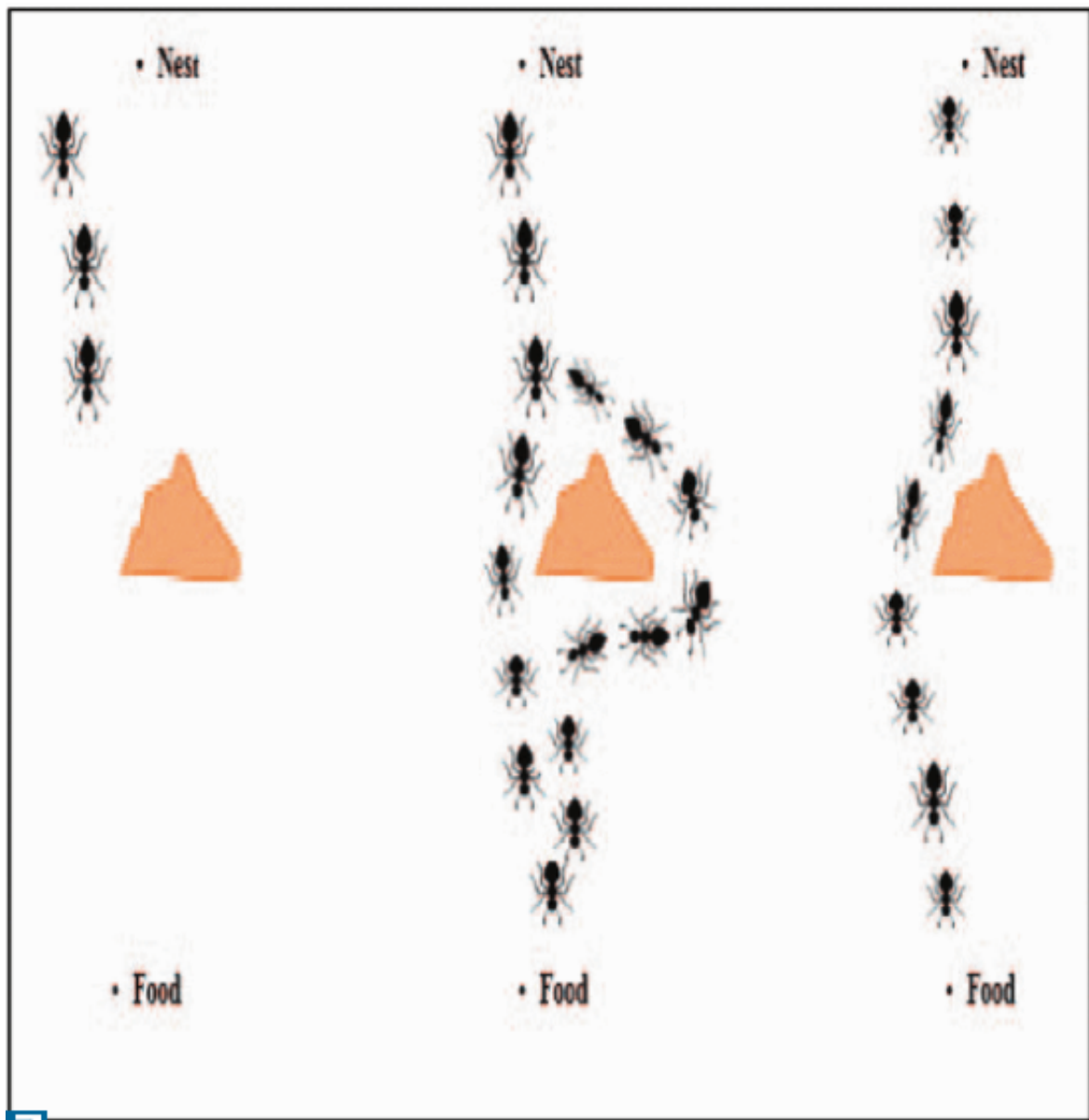
## 解决方案

### 构建模型

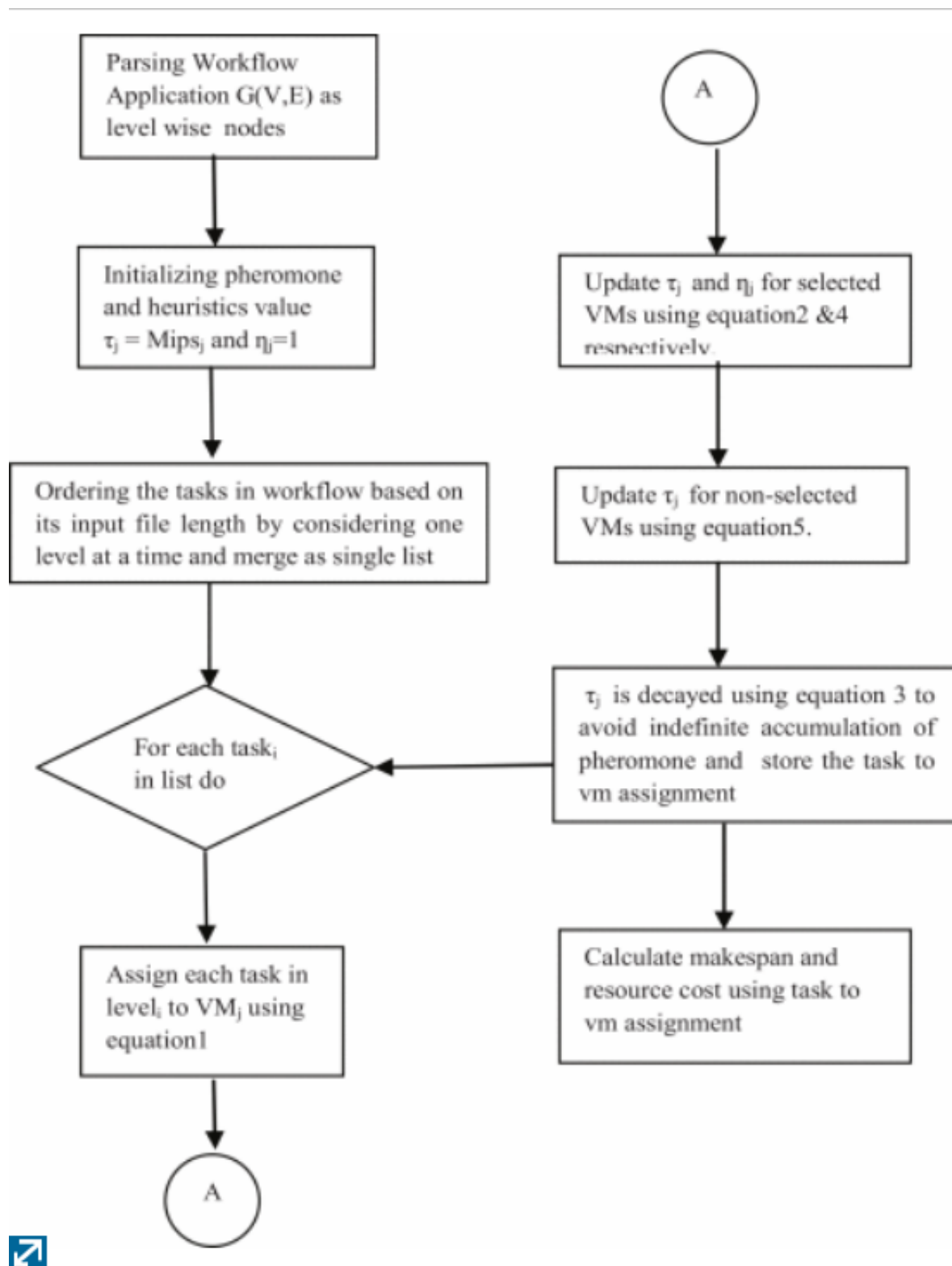
Montage 科学工作流程结构



蚁群算法概率模型



Workflow of iwsaco



## 算法摘要

### lwsaco Algorithm

- 首先，对提交的工作流按有向非循环图进行分析，并根据输入任务的长度按级别排列任务，这样，所有任务都是按顺序排列和编号的。
- Ant的数量经常被设置为1，因为蚂蚁将任务映射到虚拟机，虚拟机将使用tp公式提供最小的生成时间。Ant充当调度代理，并且工作流中所有任务的调度顺序都相同。
- 为了尽量缩短生成时间，Ant尝试使用不同的虚拟机来调度任务。因此，迭代次数等于虚拟

机的数目。虚拟机的数量是根据云主机中可用的计算能力（以mips为单位）创建的。

- 使用转移概率公式，为每个虚拟机选择信息素和启发式值。
- 在任务选择了一个虚拟机之后，信息素和启发式值将在所选虚拟机上更新。为了防止信息素的无限积累，在蚁群算法中采用了信息素尾衰减系数。
- 按照计划，任务映射到虚拟机。可以通过在每个虚拟机中查找空闲时间段来估计资源使用情况。
- 由于所有的任务都安排了一次，与WFSACO算法相比，IWSACO算法将减少大部分的时间，从而提高效率。

## 效果评价

## 创新贡献

- 本文提出了在云端使用ACO进行科学工作流程任务调度的一些基本步骤。蚁群算法是受蚂蚁社会行为启发而建立的概率模型。ACO已成功地应用于NP硬组合问题
- 作者之前的工作是使用了ACO方法来最小化科学工作流程的完成时间。在本文中提出的IWSACO算法成功克服WFSACO算法的复杂度。

结果表明，IWSACO算法比WFSACO算法具有更好的生成时间。显然，iwsaco比wfsaco花费更少的调度时间，因为对工作流程中的所有任务都进行一次排序和调度。而且，在移到下一个级别之前，虚拟机不需要等待其他虚拟机完成其任务。

Workflow Application-Number of Tasks	Makespan In WFSACO (ms)	Makespan in IWSACO (ms)
Montage-25	58.5	56.8
Montage-50	122.42	105.3
Epigenomics-24	2455.3	2100.2
Epigenomics-47	3805.2	3102.7
SIPHT-30	168.4	162.5
SIPHT-60	249.7	199.2

## 个人感想

## 本文的优势

本文提出了一种改进的基于蚁群算法的工作流调度算法，它既能最大限度地缩短工作流调度的时间，又能降低工作流调度的复杂度。

## 本文的劣势

该方法与现有算法的实验结果进行了比较，应该增加相关的调度时间、资源成本等更多指标的比较。在此基础上，还可以继续提出的算法的重新设计。