

Research for the Task Scheduling Algorithm Optimization based on Hybrid PSO and ACO for Cloud Computing

JieHui JU^{1,2}, WeiZheng BAO³, ZhongYou WANG^{4*}, Ya WANG⁵ and WenJuan LI^{*6}

¹*Zhejiang University of Science & Technology, Hangzhou 310023, China*

²*University of Colorado Boulder, Colorado 80309-0xxx, USA*

³*Surveying and Mapping Management Office, Jinhua Planning Bureau, Jinhua Zhejiang 321000, China*

⁴*Zhejiang Communications Industry Services Co,Ltd.,Hangzhou 310050, China*

⁵*The Second Surveying and Mapping Institute of Zhejiang Province, Hangzhou 310012, China*

⁶*Key Lab of E-Business, Hangzhou Normal University, Hangzhou 310036, China*

jjh1mail@163.com, wzbao1970@126.com, 798026426@qq.com,

hzwyzwy@163.com, iellie@163.com

Abstract

In cloud computing environment, there are a large number of users which lead to huge amount of tasks to be processed by system. In order to make the system complete the service requests efficiently, how to schedule the tasks becomes the focus of cloud computing Research. A task scheduling algorithm based on PSO and ACO for cloud computing is presented in this paper. First, the algorithm uses particle swarm optimization algorithm to get the initial solution quickly, and then according to this scheduling result the initial pheromone distribution of ant colony algorithm is generated. Finally, the ant colony algorithm is used to get the optimal solution of task scheduling. The experiment simulated on CloudSim platform shows that the algorithm has good effect in real-time performance and optimization capability. It is an effective task scheduling algorithm.

Keywords: *Cloud Computing, Task Scheduling, Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO)*

1. Introduction

Cloud computing is commercial implementation which combines traditional computer and network technology such as distributed computing, parallel computing, grid computing, utility computing, virtualization and network storage, etc., It is a kind of computing service mode which is based on the Internet. The convenient access for shared configurable computing resources (network, servers, storage, applications and services, etc.) and on-demand access for resources are realized [1]. These resources can be quickly prepared and used as required by tiny management cost or interactions with the service provider. Cloud services can be divided into three layers: infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS).The basic principle of cloud computing is that through the network, the huge computing program automatically can be divided into many smaller subroutine, then the large system that are composed of a number of servers, sends the results back to the user after searching, calculating and analyzing. Through this technology,

service providers can deal with terabytes of data in a few seconds on the far side which can achieve the same powerful performance of network with super computer service, such as Map/Reduce framework proposed by Google.

2. The Problem Statement for Task Scheduling Algorithm Optimization

In business, in addition to data mining, searching, data storage which can be used offline task model, all the other applications (such as electronic commerce, transaction processing, etc.,) are higher requirements for real-time business, which requires that the cloud computing system can complete the tasks sent by the users in the shortest time as far as possible. Since users are numerous in a cloud environment, system would deal with a very large amount of tasks, so service instance is generated in order to make sure that the system can quickly handle the service request. In the cloud system the problem that tasks are reasonable scheduled efficiently to realize the global optimization system, becomes the key points and difficulties of cloud computing research.

As a new computing mode, Cloud computing, it is the development of grid computing, parallel computing and distributed computing, as well as the new technology of next generation of Internet and application. Resource Scheduling Policy in Cloud computing is an important part of cloud computing technology, it mainly focus on how to allocate compute nodes for the task submitted by users, how to carry on the dynamic extension of the compute nodes in the case of meeting the requirements of service quality from customers and taking the shortest execution time to create the highest degree of load balancing, and its efficiency directly affects the performance of the entire cloud computing environment.

The ant colony algorithm and The particle swarm optimization are two swarm intelligence algorithms in the field of computational Intelligence, the former is based on the simulation of ant colonies to collect food, and the latter simulates the process of the flock seek for food. Ant colony algorithm in the experiments of solving traveling salesman problem, assignment problem, scheduling, etc, turns out to be practical. It highlights its efficiency and superiority in solving complex problems, especially in solving discrete optimization problems. Ant colony algorithm has a great prospect of development in the future. Particle swarm optimization is a kind of efficient parallel search algorithm, whose concept is simple and the algorithm is easy to implement. And this algorithm is good at solving continuous optimization problems.

In cloud computing environment, task scheduling is a NP-completed problem. Simulated annealing, genetic algorithm and ant colony optimization, particle swarm intelligence optimization algorithms are very suitable for solving NP problems. Genetic algorithm, particle swarm optimization algorithm and ant colony algorithm have been used to cloud computing task scheduling problems. Literature [1] shows that the effective scheduling for independent tasks is completed by using genetic algorithm in the cloud computing environment. Literature [2] presents a cloud computing task scheduling algorithm based on particle swarm optimization, which optimizes task transfer time and processing time. Literature [3] uses ant colony algorithm to schedule cloud computing task, which make the total completion time and average task completion time minimum. But any algorithm has its advantages and disadvantages, such as the genetic algorithm is fast random global searching ability. But more parameters lead that programming is more complex and the algorithm is easier to fall into local optimum. And the convergence speed and local search ability of particle swarm optimization algorithm in the early is good. The late convergence speed is slow. But compared with genetic algorithm, convergence speed and optimization performance of particle swarm optimization algorithm is superior to the genetic method. What's more the programming is easier to implement and the parameters which need to

adjust are less. Although the ant colony algorithm has good optimization ability, the initial pheromone is scarce and the convergence speed is slow. On the basis of the study, this paper puts forward a kind of particle swarm optimization and ant colony optimization based on cloud computing task scheduling algorithm. This algorithm absorbs the rapid convergence of particle swarm optimization algorithm and optimization ability of ant colony algorithm. The time of the system process scheduling problem and the total task execution time are reduced, which correspondingly improves the efficiency of the cloud computing task scheduling.

3. The Description for the Task Scheduling based on Cloud Computing

At present, Map/Reduce programming model which is put forward by Google is most used in cloud computing environment [4]. A large task will be divided into many smaller subtasks through the two stages of Mapping and Reducing. And then these subtasks are assigned to a number of virtual resource nodes which can realize the parallel execution. Finally the results are returned to users. This paper only considers the subtasks are independent with each other. The goal of cloud computing task scheduling is to make the total task completion time minimum by reasonable allocation of resources for the subtasks.

The problem for the task scheduling based on cloud computing can be described as follows: n subtasks which are independent with each other are assigned to m virtual resource nodes for execution ($m < n$). And the set of tasks are expressed as $T = \{t_1, t_2, \dots, t_n\}$. The set of the virtual resource nodes are expressed as $VM = \{vm_1, vm_2, \dots, vm_m\}$, t_j ($j = 1, 2, \dots, n$) represents the first j subtask. vm_i ($i = 1, 2, \dots, m$) represents the first i virtual resource node. Every subtask can only execute in one virtual resource. The distribution relationship between the tasks' set $T = \{t_1, t_2, \dots, t_n\}$ and the virtual resource nodes $VM = \{vm_1, vm_2, \dots, vm_m\}$ can be represented by the following matrix X .

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} \quad (1)$$

Among the expression (1), x_{ij} represents the distribution relationship between the subtask t_j and the virtual resource vm_i ($x_{ij} \in \{0,1\}$). If the subtask t_j execute in the resource vm_i , $x_{ij} = 1$, otherwise, $x_{ij} = 0$. The ETC_{ij} is the expected execution time of the subtask t_j in the virtual resource vm_i . The distribution relationship X between the corresponding task and the virtual resource can consist of the ETC matrix.

$$ETC = \begin{pmatrix} ETC_{11} & ETC_{12} & \cdots & ETC_{1n} \\ ETC_{21} & ETC_{22} & \cdots & ETC_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ ETC_{m1} & ETC_{m2} & \cdots & ETC_{mn} \end{pmatrix} \quad (2)$$

Among the expression (2), b_j is the beginning time of task t_j . $\max\{CT_{ij}\}$ is the total time of all the tasks. $C_{\max} = \max\{CT_{ij}\}$ is the fitness value of the scheduling. The optimization target is to make the value minimum.

4. The Algorithm Designing for the Cloud Computing Task Scheduling

Particle swarm optimization algorithm proposed by Kennedy and Eberhart in 1995 is inspired by birds foraging behavior, which is a kind of bionic optimization algorithm. Since the method is simple, easy to realize and has less parameters etc., this algorithm has shown good effect in continuous optimization problems and discrete optimization problems. Ant colony algorithm is put forward by Dorigo Macro and others in the 1990s. The algorithm simulates ant colony foraging behavior. Parallel algorithm is introduced into the positive feedback mechanism, which is equipped with strong robustness, excellent distributed computing mechanism and easily combination with other algorithms. It is widely used to solve complex combinatorial optimization problem [5].

Although the particle swarm optimization algorithm and ant colony algorithm respectively achieve good effect in solving the optimization problem, each algorithm has its disadvantages. The convergence speed of particle swarm optimization algorithm is fast in the initial stage, but the local searching ability will decrease in the later due to the lack of population diversity. And although ant colony algorithm has good optimization ability, the early convergence speed is slow due to the lack of pheromone in initial stage. Combined with the advantages of both algorithms, this paper puts forward a new cloud computing task scheduling algorithm based on particle swarm optimization and ant colony optimization. The algorithm firstly uses the fast convergence of the particle swarm optimization algorithm to generate initial solution quickly. According to the scheduling results, the initial pheromone distribution of ant colony algorithm is generated. At last, ant colony algorithm is used to find the optimal solution of task scheduling.

4.1 Particle Swarm Optimization Algorithm: The cloud computing task scheduling is a discrete problem. And the particle swarm optimization algorithm is mainly suitable for solving continuous optimization problems. In the discrete version of particle swarm optimization (PSO) algorithm, each particle position vector is set as 0 or 1. Particle velocity no longer means the speed of particles in continuous space flight. It means the probability of calculating the particle position vector values to 0 or 1.

Based on the characteristic of the cloud computing task scheduling, the position and velocity of every particle should be redefined. In the particle swarm, the position of each particle means the task scheduling scheme. According to the distribution relationship matrix, the position of first i particle can be represented as the expression (3).

$$X_i = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} \quad (3)$$

Among the expression, $x_{ij} \in \{0,1\}$, $\sum_{i=1}^m x_{ij} = 1, i \in \{1, 2, \dots, m\}, j \in \{1, 2, \dots, n\}$.

Similarly, the velocity of the first i particle is defined as the expression (4).

$$V_i = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{m1} & v_{m2} & \cdots & v_{mn} \end{pmatrix} \quad (4)$$

Among the expression, $v_{ij} \in [-v_{max}, v_{max}]$, $i \in \{1, 2, \dots, m\}$, $j \in \{1, 2, \dots, n\}$.

In the discrete particle swarm optimization algorithm, the fuzzy function is introduced to ensure the value of the position vector is 0 or 1.

$$sig(v_{ij}^{k+1}) = \frac{1}{1 + exp(-v_{ij}^{k+1})} \quad (5)$$

4.2. Ant Colony Optimization Algorithm: With reference to the cloud computing task scheduling and resource allocation relationship matrix, the undirected complete graph is built up. With the ant colony algorithm, a group of nodes will be selected from the set which will make the optimization target minimum.

In the initialization stage of the ant colony algorithm, the pheromone on the nodes of the initial solution is strengthened a certain multiple based on particle swarm optimization (PSO) algorithm for scheduling results, which can makes convergence of ant colony algorithm accelerating. The ants can quickly close to the optimal solution in the process of traveling. Through the many times of experiments, the results show that 6 times pheromone is best.

Some ants are randomly placed on the nodes to travel. The transition probability of node x_{ij} for the first k ant in the time of t are as follows:

$$p_{ij}^k(t) = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}(t)]^\beta}{\sum_{xis \notin tabu_k} [\tau_{is}(t)]^\alpha [\eta_{is}(t)]^\beta}, & x_{ij} \notin tabu_k \\ 0, & \text{else} \end{cases} \quad (6)$$

$$\eta_{ij}(t) = \frac{1}{CT_{ij}^k(t)} \quad (7)$$

$\tau_{ij}(t)$ means the remaining pheromone on the node x_{ij} in the time t . $\eta_{ij}(t)$ represents the inspiring information. $tabu_k$ ($k = 1, 2, \dots, m$) is the taboo table of the ant k . Every subtask t_j is corresponding to a group of nodes $\{x_{ij}\}_{i=1}^m$. The ant k selects the node x_{ij} in the time of t , then the $\{x_{ij}\}_{i=1}^m$ will be added into the taboo table $tabu_k$. α and β are respectively the relative important degree of the information amount and heuristic information.

When the ant selects a node, the task was assigned to some virtual resource. And then the pheromone of node will be changed. According to the expression (8) and expression (9), the pheromone can be updated locally.

$$\tau_{ij}(t+1) = (1-p) \cdot \tau_{ij}(t) + \sum_{k=1}^m \Delta \tau_{ij}^k(t) \quad (8)$$

$$\Delta \tau_{ij}^k(t) = \begin{cases} \frac{Q}{CT_{ij}^k(t)} & \\ 0 & \end{cases} \quad (9)$$

Among $\rho \in [0,1]$ is pheromone volatilization coefficient. $1 - \rho$ represents the information remaining coefficient. $\Delta \tau_{ij}^k(t)$ means the information amount remained in the path (i, j) . And Q is constant.

5. The Experiments and Analysis

CloudSim is developed by Dr Rajkumar Buyya in university of Melbourne Australia which is the cloud computing simulation platform. This platform is mainly used to simulate the cloud environment.

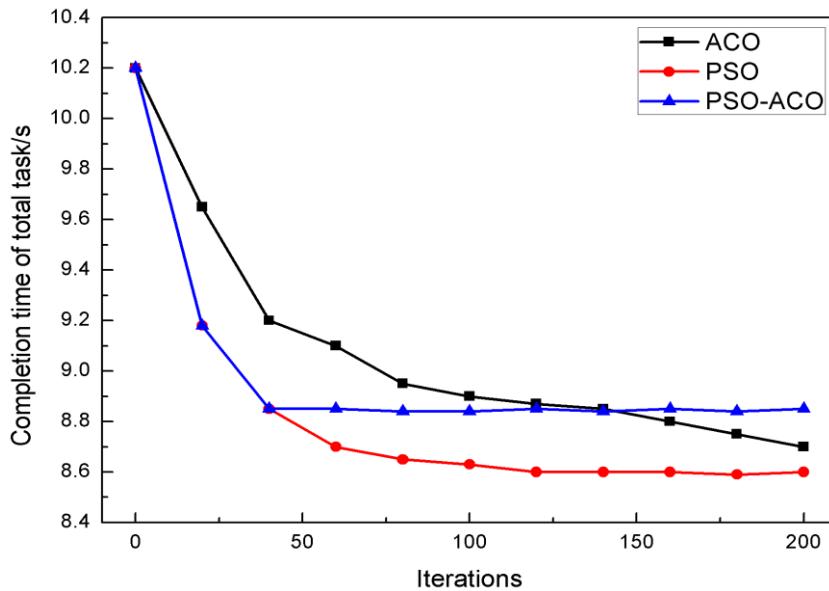


Figure 1. The Execution Time Curve of the Total Task (5 Recourses and 50 Subtasks)

This article uses the CloudSim platform to simulate the task scheduling algorithm in the cloud environment. To check the performance of the scheduling algorithm, this paper respectively do experiments under the condition of five virtual computing resources and 50 subtasks, and five virtual computing resources, 500 subtasks. The comparison experiments are carried out by respectively using particle swarm optimization algorithm (PSO), ant colony algorithm (ACO), and the hybrid PSO-ACO algorithm. The dealing capacity of each virtual computing resource is {400 MIPS, 600 MIPS, 800 MIPS, 1000 MIPS, and 1200 MIPS} in these experiments. The subtask assignment lengths are in the range of [400 MI, 400 MI]. The algorithm parameter Settings are as follows: in the part of particle swarm optimization (PSO) algorithm, the group scale size =100, $c_1 = c_2 = 2$ and the number of iterations is 40. In the part of the ant colony algorithm, population scale size =100, $\alpha = 1$, $\beta = 1$, $\rho = 0.2$ and the number of iterations is 160. The parameters of the particle swarm algorithm, ant colony algorithm are the same with the hybrid algorithm in this paper. The number of iterations is 200. And the

algorithm runs repeatedly 20 times. The experimental results are shown in Figure 1 and Figure 2.

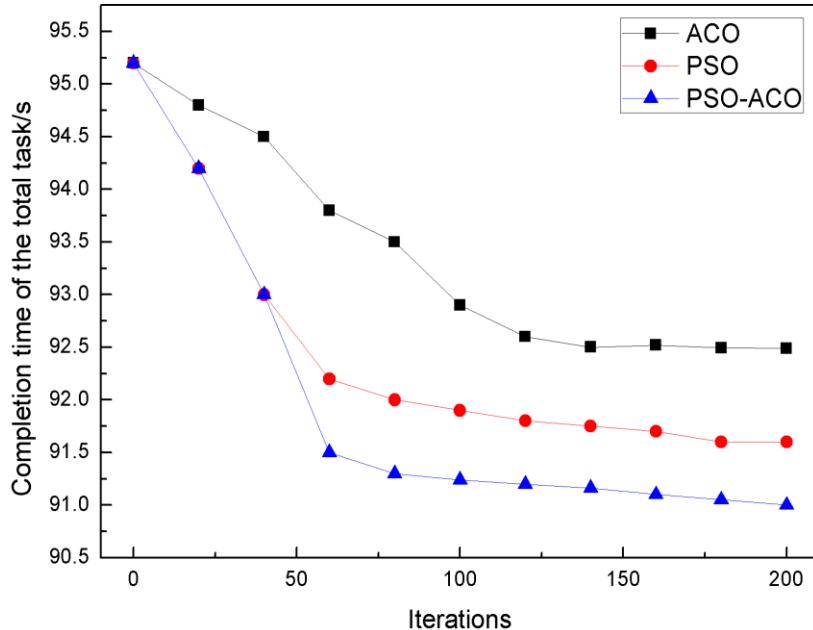


Figure 2. The Execution Time Curve of the Total Task (5 Recourses and 500 Subtasks)

Based on the above experimental results, this paper proposed PSO- ACO task scheduling algorithm combines the fast convergence ability of PSO and the optimization ability of ACO. At the beginning of the iteration, the task scheduling optimization effect of PSO-ACO is better than that of the ACO. Through the simulation experiments under the two kinds of deployment, the results show that the proposed scheduling algorithm obviously shortens the scheduling time and reduces the execution time under the total task. At the same time, the time performance and optimization performance achieve good results.

6. Conclusion

In view of the cloud computing task scheduling problem [6-7], this paper puts forward a cloud computing optimization algorithm based on particle swarm optimization and ant colony optimization algorithm. Through many simulation experiments, it is proved that the algorithm has a good real-time performance and optimization ability and is an effective scheduling algorithm.

Acknowledgements

This work is supported by the Scientific Research Fund of Zhejiang Provincial Education Department (No.Y201223199), the open fund for Key Lab of Visual Media Intelligence Handles Technology of Zhejiang Province (NO.2012017), the Natural Science Fund of Zhejiang Province (NO.LQ12G02016,LQ12F02006), the open fund for Key Lab of E-business Market Application Technology of Guangdong Province (2011GDECOF07).

References

- [1] M. A. Arfeen, K. Pawlikowski and A. Willig, "A Framework for Resource Allocation Strategies in Cloud Computing Environment", Computer Software and Applications Conference Workshops (COMPSACW) (2011) IEEE 35th Annual, (2011), pp. 261-266.
- [2] G. Lizheng, Z. Shuguang, S. Shigen, *et al.*, "Task Scheduling Optimization in Cloud Computing Based on Heuristic Algorithm", Journal of Networks, vol. 7, no. 3, (2012), pp. 547-553.
- [3] J. Kennedy and W. Spears, "Matching Algorithms to Problems: an Experimental Test of the Particle Swarm and Some Genetic Algorithms on the Multimodal Problem Generator", Proc IEEE International Conference on Evolutionary Computation. Piscataway, NJ: IEEE Service Center, (1998), pp. 78-83.
- [4] J. Dean and S. Ghemawat, "Map Reduce: Simplified data processing on large clusters", Proceedings of the 6th Symposium on Operating System Design and Implementation. New York: ACM, (2004), pp. 137-150.
- [5] J. Kennedy and R. C. Eberhart, "Particle swarm optimization", Proc IEEE International Conference on Neural Networks, IV. Piscataway, NJ: IEEE Service Center, (1995), pp. 1942-1948.
- [6] W. Jiyi, Z. Jianlin, W. Tong and S. Qianli, "Study on Redundant Strategies in Peer to Peer Cloud Storage Systems", Applied Mathematics & Information Sciences, vol. 5, no. 2, (2011), pp. 235S-242S.
- [7] W. Ji-yi, F. Jian-qing, P. Ling-di, X. Qi, "Study on the P2P Cloud Storage System", Acta Electronica Sinica, vol. 39, no. 5, (2011), pp. 1100-1107.

Authors



Jiehui Ju, she is a Lecturer at the Zhejiang University of Science and Technology. Currently, she is also a visiting scholar of the University of Colorado Boulder in America (from May 10, 2013 to Jun 5, 2014). Her research interests include Cloud Computing , Data Mining.



ZhongYou Wang, he is the deputy director at Technology R & D Center of Zhejiang Communications Industry Services Co, Ltd. He received Master's Degree in computer science in 2009 at Hangzhou Dianzi University. His research interests include service trust and security, software engineering.



WeiZheng Bao, he was born in 1970. He is a senior engineer and director of surveying and mapping management office, Jinhua Planning Bureau, Zhejiang Province, China. His research interests include cloud computing and digital city.



Ya Wang, she was born in 1973. She is a senior engineer of surveying and mapping, the second surveying and mapping institute of Zhejiang Province. Her research interests include cloud computing, cadastre and virtual reality.



WenJuan LI, she was born in 1978. She is a lecturer at Hangzhou Normal University. She received PhD degree in 2011 from Zhejiang University, Hangzhou China, in computer science. Her research interests include cloud computing, trust and reputation. She has published more than 10 journal articles indexed by SCI, EI.

