Research on Scheduling Strategy based on Trust in Hybrid Cloud

Xiao-Lan Xie[1,2] , Shun-Bai Huang *[1] and Xiu-juan Guo[1]

[1] College of Information Science and Engineering, Guilin University Of
Technology, Guilin ,Guangxi Zhuang Autonomous Region, China, 541004
[2] Guilin University of Technology， GuangXi key Laboratory of Embedded
Technology and Intelligent Information Processing， China, 541004
891524620@qq.com

Abstract

Hybrid cloud can help companies to maximize the potentiality of IT infrastructure, so the hybrid cloud is increasingly popular in the market. Hybrid cloud scheduling has become one of the important issues which cloud providers face in the current. How to effectively and safely make hybrid cloud scheduling is one of the main research topics in the field of hybrid cloud. This paper introduces an effective hybrid cloud resource management model, and proposes a hybrid cloud resource scheduling strategy based on the particle swarm leapfrog algorithm and trust mechanism in order to solve the problems. The experimental results show that the hybrid cloud scheduling strategy is effective in terms of security and throughput.

Keywords: Hybrid cloud; scheduling strategy; Trust mechanism; Shuffled frog leaping algorithm

1. Introduction

As an efficient computing model, cloud computing has been developed rapidly in recent years, and the hybrid cloud has become the preferred way to complete the computing tasks such as research institutions and companies. Public cloud commonly refers to usable cloud provided by the third party, it can use through the Internet. Private cloud is built for a single customer, it provides the most effective control of data, security, and quality of service. Hybrid cloud refers to the combination of public cloud services and private cloud services, this service way helps to provide on-demand, the expansion of the external supply of resources [1]. Cloud scheduling is a distributed dynamic resource allocation and automatic deployment for the submitted work flow by users in the cloud platform and cloud technology [2]. A good scheduling algorithm is the key part of cloud computing services. Nowadays most scheduling algorithm mainly solves the scheduling problem in a single cloud environment, but there is little research on scheduling strategy in hybrid cloud environment. Literature [3] uses integer programming model and task scheduling algorithm, and ensures the quality of service (QoS) of IaaS cloud provider as well as maximum profit. Literature [4] not only considers the total cost of the task execution of the hybrid cloud deadline constraint, but also analyzes the data security requirements and the task dependency problem. In literature [5], hybrid cloud platform architecture is proposed by using the concept of interconnection grid, and a resource allocation strategy based on perceived failure is proposed. These models or algorithms improve the efficiency of hybrid cloud scheduling to some extent, but most of them do not take into account the security issues of cloud computing. This paper introduces trust mechanism to hybrid cloud environment scheduling, to ensure the trust benefits of hybrid cloud scheduling. At

the same, it ensures the efficient scheduling of hybrid cloud by the combination with particle swarm shuffled frog leaping algorithm which has fast convergence and not easy to fall into local optimum.

2. A Hybrid Cloud Model Description

In the hybrid cloud, various types of resources provide services in the form of virtual machine (VM), the cloud task will be deployed to each virtual machine through the scheduling algorithm. Let RNprv and RN_{prb}^{i} represent the number of private cloud (Cprv) and public cloud (C_{prb}^{i}) resources. The hybrid cloud can be shown as the following:

$$C_h : C_{prv} \cup \{C_{pub}^{i}\} \tag{1}$$

$$RN_h = RN_{prv} + \sum_i RN_{pub}^{i} \tag{2}$$

Hybrid cloud formation and scheduling infrastructure architecture is shown in Figure 1, public cloud and private cloud respectively have their own monitoring server, firstly the user task is sent to the hybrid cloud, hybrid cloud scheduler collect the security information for private cloud VM and use, if each private cloud VM have exceeded the safety threshold, then apply for public cloud resources through public cloud addressing, and manage the public cloud resources through the unified resource scheduler. We design an effective task scheduling scheme according to the trust mechanism and the particle swarm-leapfrog algorithm. In the end, the resource allocation task is deployed to the hybrid cloud VM.
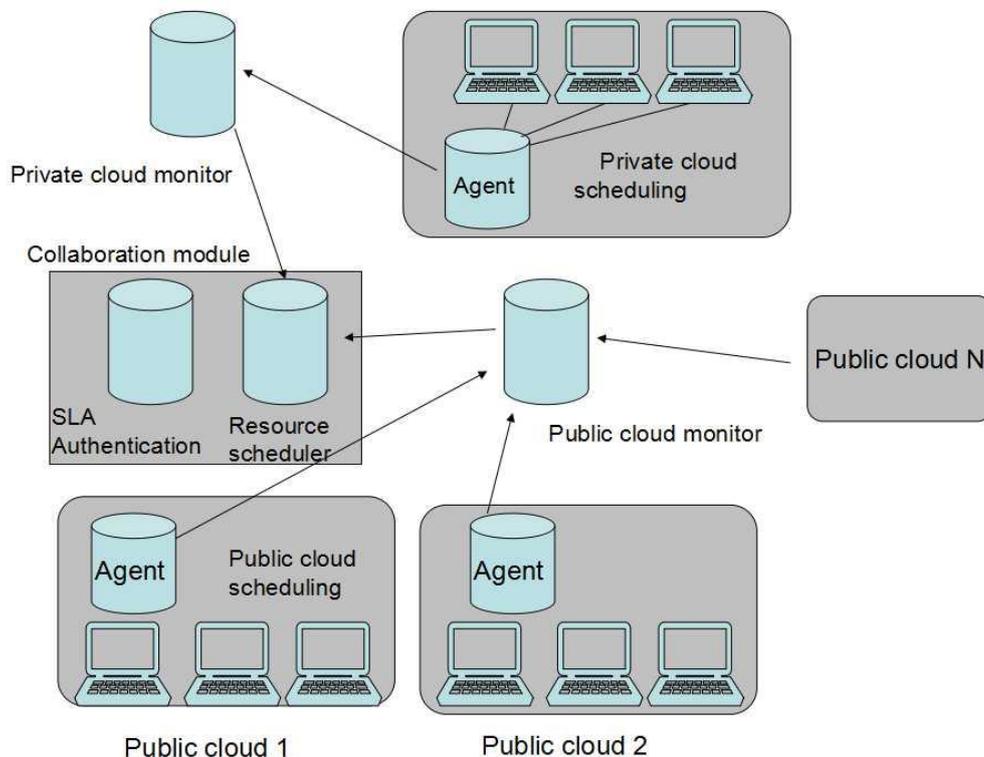
Figure 1. Hybrid Cloud Formation and Scheduling Infrastructure Architecture

3. The Scheduling Algorithm

3.1 Trust Relation Model

In this paper, a trust model based on the hybrid cloud environment is defined on the basis of distinguishing the trust region. The elements of the trust model: TM = (I,J,K , t), I represents the credit provider, J represents the credit receiver, K represents presenter, K = {K1 ,K2 , ...Kn}, among them, I, J, K are the entities in the cloud computing environment. t is the time constraint of trust relation.

Definition 1 Trust: the credit provider assesses the trust degree of a particular entity of the credit receiver in a specific time range. Trust is the evaluation for the credit receiver by the credit provider which is subjective and depends on its own experience and other relevant knowledge

Definition 2 Trust degree: it is an assessment about Credibility of credit receiver. And it can be continuous or discrete, expressed as TV (I, J), it means the trust degree of node I and J.

Definition 3 Direct trust: it expresses that the credit provider and the credit receiver complete direct interaction, then two sides have a direct trust relationship. Denoted as DTV (I, J), which represents the direct trust of I to J.

Definition 4 Recommendation trust: it indicates that the credit provider and the credit receiver establish the trust relationship through other entities' recommendations before their direct interaction. Denoted as RTV (B, J), which represents the recommendation trust of B to J.

Definition 5 Risk: it responds to the unreliable degree of the credit receiver in recent times, it can be achieved by the direct interaction between the credit provider and the credit receiver.

3.1.1 The calculation of trust: (1) Calculate the direct trust degree DTV (I, J, t) in the moment of t between I and J. The formula is as formula (1):

$$DTV(I,J,t) = \gamma DTV(I,J,t_1)A(I,J,t-t_1) - (1-\gamma)R(J,t-t_1) \tag{3}$$

In the formula, γ is weight factor which value is between [0,1], and is used to adjust the time range focused by trust. DTV (I, J, t_1) is the direct trust value of I, J in t1 time interaction, A (I, J, t- t_1) is the time decay function, its calculation method is $A(I,J,t-t_1) = \dfrac{1}{1+(t-t_1)/V}$, V is the decay rate. R (J, t-t_1) is the degree of risk of J at t time.

(2) Calculate the recommendation trust degree RTV (B, J, t) at a certain time t from entity K to J, the formula is as formula (2):

$$RTV(B,J,t) = \frac{\rho(n) \times \sum\limits_{k=1}^{n} RTV(B_k,J,t_1) \times \gamma(t-t_1)}{n} \tag{4}$$

In the formula, γ is the same as formula (1), n is the number of cloud trust domain. In order to resist the defamation attack of some malicious entities, using function $\rho(n)$ to adjust the influence of the number of recommended entities on the recommendation trust, this function reflects that if the number of entities that evaluate the quality of service of the node is more, the evaluation is more close to the average of entity evaluation. We let $\rho(n) = e^{-\frac{1}{n}}$ here.

(3) In conclusion, calculate the comprehensive trust degree TV(I , J ,t) at a certain time t from entity I to J, the formula is as formula :

$$TV(I,J,t) = \alpha DTV(I,J,t) + \beta RTV(B,J,t) \tag{5}$$

Determine the value of α, β according to the specific environment of the cloud, their value is between [0,1] and meet the requirement $\alpha + \beta = 1$.

3.2 The Particle Swarm Leapfrog Fusion Algorithm

3.2.1. Particle Swarm Optimization (PSO) Algorithm and Leapfrog Algorithm: PSO algorithm proposed by Kennedy and Eberhart [6] in 1995, it is affected by the research results of birds behavior, then they carried out the modeling and simulation. The standard PSO algorithm calculate the fitness of each particle i in the iterative process and update themselves by tracking the individual extremum P_i and global extreme value P_g [11]. In the cloud job scheduling model: assuming there are Np particles, particle search space $D = m \times n$, where the number of operations is m, the number of virtual machine resources is n, the position of the **i**th particle X_i is $X_i = (X_{i11}, \ldots X_{imn})$ and speed Vi is $Vi = (V_{i11}, \ldots V_{imn})$. $X_{iab} \in \{0,1\}$ denotes that the **i**th particle assign the **a**th task to the **b**th resource, (a=1...m, b=1...n). The optimal position of the particle i is $P_{iab} = (P_{i11} \ldots P_{imn})$, the global optimal position is $P_{gab} = (P_{g11} \ldots P_{gmn})$, then the velocity update formula is:

$$V_{iab}^{new} = \omega V_{iab} + c_1 r_1 (P_{iab} - X_{iab}) + c_2 r_2 (P_{gab} - X_{iab}) \tag{6}$$

ω is the inertia weight; R1, R2 is a random number between (0,1); C2, C1 is a learning factor.

The shuffled frog leaping algorithm exchange the information obtained by local search with global information [7], it firstly elected the best individual in the local and then exchanged local information and global information by using the shuffled strategy. Job scheduling model in the cloud is: the search space $D = m \times n$, the number of jobs is m, the number of virtual machine resource is n, the number of frogs is L, the **i**th frog allocates a task to a resource can be showed as $X_{iab} = (X_{i11} \ldots X_{imn})$. The best solution in each frog population is X_{fab}, the worst solution is X_{wab}. The optimal solution for all populations is X_{gab}, and the updating formula is as follows:

$$d = wd` + rand() \times (X_{fab} - X_{wab}) \tag{7}$$

$$\tilde{X}_{wab} = d + X_{wab} \tag{8}$$

And d is step Dw, $rand() \in (0,1)$, $d \in [d_{max}, d_{min}]$, d` is the previously moving step, w is the attribute weight coefficient. If $\tilde{X}_{wab}$ is better than X_{wab}, then update the X_W. If there is no improvement about adaptive value, then use X_{gab} replacement X_{wab}. Repeat the updating time until the number of iterations is exhausted.

3.2.2. The Fusion Strategy of Particle Swarm Leapfrog Fusion Algorithm (SFLA-PSO): Firstly produce the initial population in the same size, the number is P. Optimize each population on the basis of PSO algorithm, put the best particles of each group together to form a new group according to the Min-Min algorithm and use the SFLA algorithm to evolve them. It can effectively guide the particle swarm optimization algorithm to avoid local optimization. The fitness function of the algorithm is set as follows:

$$f = \frac{TV(I,J,t)}{MakeSpan} \tag{9}$$

$TV(I,J,t)$ is the only trust degree. The greater the trust degree is, the time span is smaller, the greater the fitness function value is. The updating formulas of the particle swarm leapfrog fusion algorithm are as follows:

$$V_{iab}^{new} = \omega V_{iab} + c_1 r_1 (P_{iab} - X_{iab}) + c_2 r_2 (P_{igab} - X_{iab}) + c_3 r_3 (X_{gab} - X_{iab}) \tag{10}$$

ω is the inertia weight; R1, R2 is a random number between (0,1); C1, C2 is learning factor.

The algorithm flow is as follows:

a.   Set algorithm parameters;

b.   The particle swarm is divided into L (L is the number of group), the number of particles in each group is same;

c.   Calculate the fitness of each particle according to the fitness function and find out the best particle in each group P_{gab};

d.   Update the algorithm according to the formula (10);

e.   Find out the optimal P_{gab} for each group, group P_{gab} in each group by using leapfrog thought;

f.   Solve X_{wab} by using the update formula (7)(8)of leapfrog algorithm;

g.   Output X_{wab} when completing the maximum number of iterations, or return to d.

3.3. The Description of Hybrid cloud Scheduling Algorithm

Due to private cloud local application cost is zero in the hybrid cloud environment, algorithm follows the principle of private cloud priority and user task will first be assigned to a private cloud environment constructed by user, if private cloud resources can't meet the needs of users' trust demand, then apply resources to the public cloud, scheduling algorithm assign resources which meet trust requirements to tasks. The specific scheduling process of algorithm is shown in Figure 2.
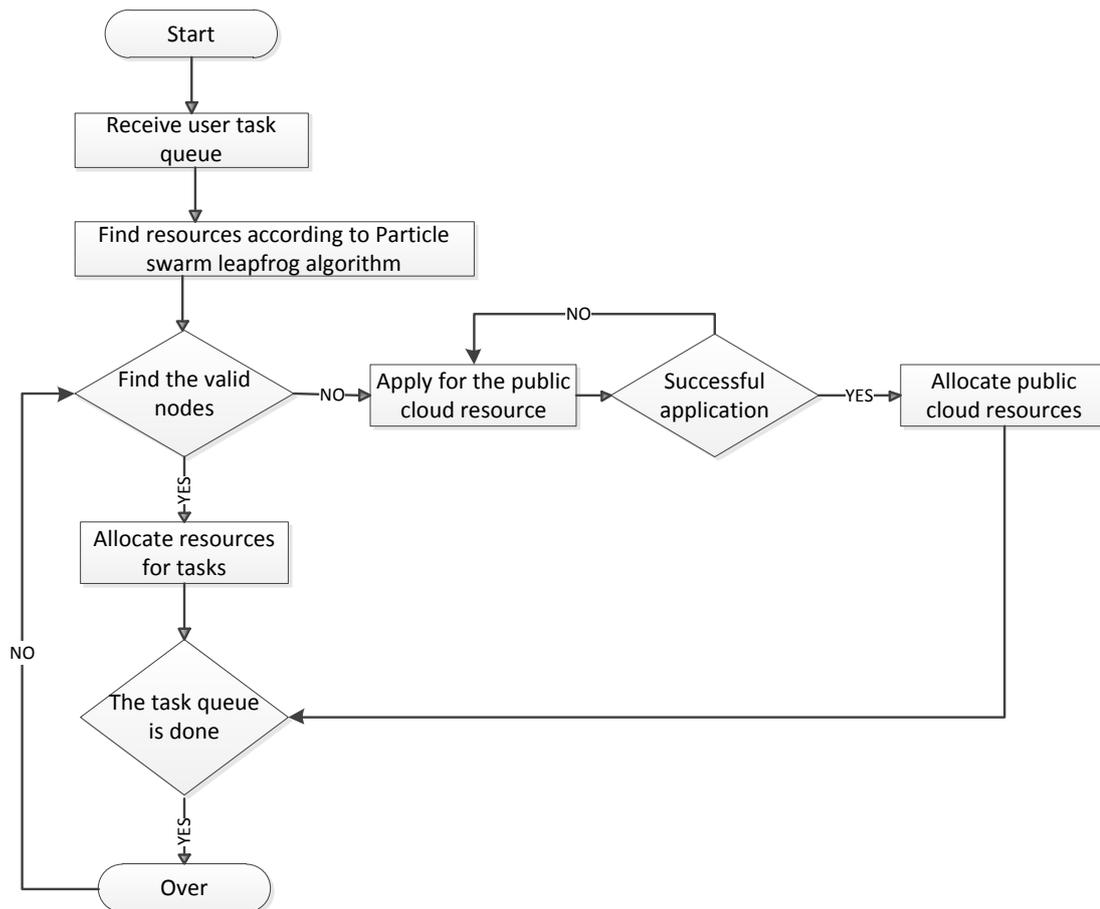

Figure 2. The Hybrid Cloud Scheduling Algorithm

We introduce the set $validP_k = \{p_{k1}, p_{k2},...p_{kn}\}$ to express the collection of valid nodes found by particle k within the required time. According to the mapping from particle swarm leapfrog algorithm to hybrid cloud resource management, the algorithm is

well applied to hybrid cloud virtual machine resource management. It can carry out the unified management and scheduling of the virtual machine resource and the virtual machine resource in the private cloud by the scheduling of the hybrid cloud collaboration module.

4. The Simulation Results and Analysis

This paper uses CloudSim[8] as the cloud simulation platform to carry out the simulation experiment. The hardware of the experiment is as follows: 2.94Ghz dual-core, 500G hard disk, 8G memory; the number of tasks is set at 20~120. The virtual machine configuration of the public cloud provider in the simulation experiment is shown in Table 1, and the number of private cloud hosts is 6. The number of simulation tasks is set to 100-600. In PSO-SFLS algorithm, we set $c_1=c_2=c_3=2$, $\omega =1.2$, The global iteration number of the algorithm is 500. In order to show the superiority of the algorithm, the simulation experiments contrast task execution time and task completion rate between the particle swarm shuffled frog leaping algorithm based on trust mechanism under hybrid cloud environment and TDMin-Min algorithm under a single cloud environment. Experimental results are shown in Figure 3 and Figure 4:

Table 1. Aliyun ECS

Type	the number of processor cores	Memory(GB)
c1.a	1	1
c2.b	1	2
s1.1	4	4
s2.2	4	8
s3.3	8	8
s4.4	8	16
s5.5	10	16
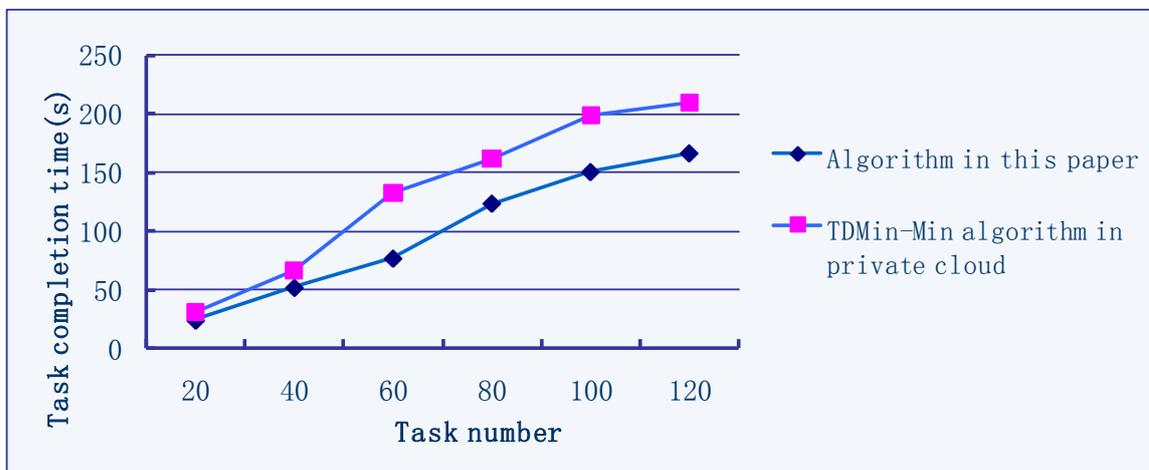
Figure 3. Task Completion Time Contrast
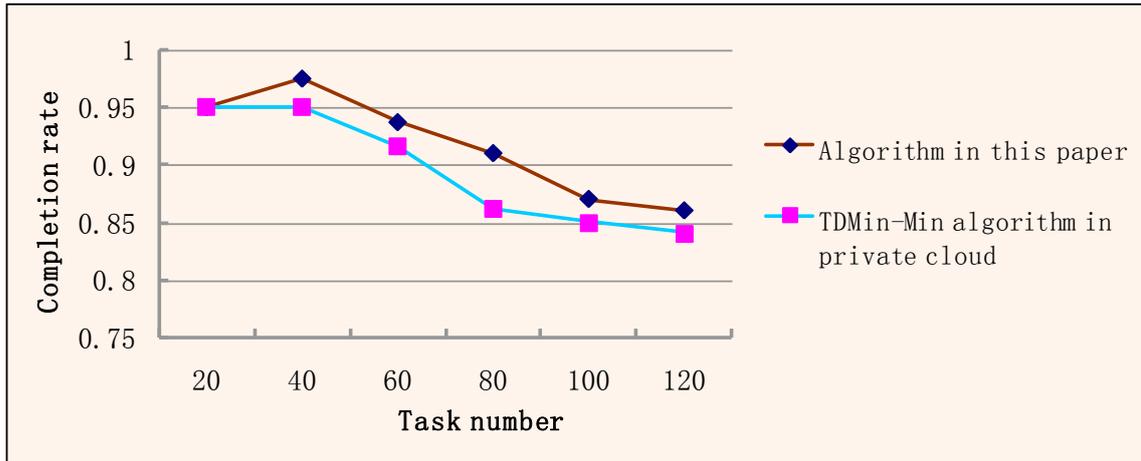
Figure 4. Task Completion Rate Contrast

As shown in Figure 3, due to blind resource allocation of a single cloud TDMin-Min scheduling algorithm, it always allocates the first appropriate resources blindly to the task, this may assign tasks to the worst resources, so it may cause the longest consuming time when increasing the amount of task. As we can be seen from the figure, when performing 300 tasks, a single cloud TDMin-Min scheduling algorithm consuming time increased significantly. This is because computing resources in the hybrid cloud is far more than resources in a single cloud. When the appropriate private cloud resources in the hybrid cloud is not sufficient, it can apply to the public cloud computing resources, so that it saves the time to wait for the right resources for the task. From figure 4, the particle swarm shuffled frog leaping algorithm based on trust mechanism in the hybrid cloud environment greatly increased the selectivity of trusted resources due to the introduction of public cloud. Therefore, the task completion rate is always better than a single cloud TDMin-Min algorithm.

In short, the algorithm ensures the security of hybrid cloud scheduling. To a certain extent, it also ensures the overall system throughput.

5. Conclusion

In order to improve the safety of hybrid cloud task scheduling, the trust mechanism is introduced into the hybrid cloud task scheduling algorithm which combined with the particle swarm shuffled frog leaping algorithm to improve the security and efficiency of hybrid cloud scheduling. Simulation results show that the scheduling algorithm in this paper is more efficient than the single TDMin-Min algorithm. With the continuous development of cloud computing technologies, cloud structure and processing will become more complex, the next step we will optimize the hybrid cloud scheduling and research more efficient algorithms to meet the scheduling needs.

6. Acknowledgements

7. References

[1]   Yunzhu Gao. Construction and Design of the Resource Scheduling Models Based on Hybrid Cloud Models [D]. XiDian University, 2013.

[2]   Xiaonian WU,Mengqing Deng,A task scheduling algorithm based on Qos-driven in Cloud Computing [J],Procedia computer Science 17,2013: 1162-1169.

[3]   Guoxiang Zhang. Research on PSO algorithm based Task Scheduling in a Hybrid Cloud[D]. Beijing University Of Posts And Telecommunications,2012.

[4]   Jiangyong Wang,Biao Xu, Coupled Task Scheduling Strategy Based on Data Protection in Hybrid Clouds[J],Modern Computer,2014:1007-1423.

[5]   Bahman Javadi, Jemal Abawajy, Rajkumar Buyya. Failure-Aware Resource Provisioning for Hybrid Cloud Infrastructure[J]. Journal of Parallel and Distributed Computing, 2012,72(10):1318-1331.

[6]   Eberhart R., Kennedy J. A new optimizer using particle swarm theory[C]. Proceedings of the 6th International Symposium on Micro Machine and Human Science, 1995, pp.39-43.

[7]   Yao-chen Luan. Development status of Shuffled frog leaping algorithm And its Research [J] DA ZHONG KE JI, 2009,(1): 24-25.

[8]   Yan-Ni Wang, Wen-Hui Wu. The Analysis of Simulation Process of Cloudsim3.0[J], SOFTWARE,2014,35(4),63-64.

[9]   Lu lu.Research on Job Scheduling in Cloud Environment[D].Nanjing University of Science and Technology,2013.

[10] Dong Guofang. Research on Trust Mechanism-based Resource Scheduling in Grid Environment[D]. Shandong University,2007.

Authors

Xiao-Lan Xie, She got her PhD in Xidian University, Shan Xi, China. She is a Professor in School of information science and engineering, Guilin University of Technology. Areas of research include Cloud computing, Grid computing and Intelligent decision system. She is a committee member and deputy secretary general of Cloud computing expert committee of China communication society. She is also a member of China computer society CCF and IEEE CS.

Shun-Bai Huang, He is M. S. candidate. His research interests include Cloud computing, Multi-Agent, data mining, etc

Xiu-Juan Guo, She is M. S. candidate. Her research interests include Cloud computing, data mining, etc

作者姓名	职称/学位	单位	地址	手机	Email
谢晓兰	教授	桂林理工大学	雁山校区	15277336998	237290696@qq.com
黄顺柏	硕士	桂林理工大学	雁山校区	18877323344	891524620@qq.com
郭秀娟	硕士	桂林理工大学	雁山校区	18778380821	920717554@qq.com