Fuzzy logic based job scheduling algorithm in cloud environment

Pratibha Pandey, Sarvpal Singh*

Department of Computer Science and Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur -273010, India

**Corresponding author's e-mail: Singh_sarvpal@yahoo.co.in*

Received 15 June 2017, www.cmnt.lv

Abstract

Cloud computing is a technology which is growing faster day by day and applied in various fields such as in industry, commerce, and research. Handling resources and the task according to the need of user is the current major issue. In cloud environment when users submit their task, it selects the best virtual machine on which the task can execute. Considering the commercialization and virtualization aspect of the Cloud Environment, this paper proposes an algorithm for Job scheduling which ensures fairness of the resource allocation according to the Quality of service. It mainly focuses on two problems. One is the selection of virtual machine(s) which are eligible to execute the task. Another problem is justification of the task according to the quality of service. Our approach simplifies the complexity of the algorithm and reduces the overhead associated with selecting appropriate and justified virtual machine for a given task. It ensures the fairness of the resource allocation for each classified task and also justifies the overall system allocation. Further, it uses fuzzy logic to adjust the general expectation vector of the task based on the fairness of the allocation of resource.

Keywords:

Cloud Computing,
QoS,
Fairness,
Job scheduling,
Resource allocation

1 Introduction

Cloud computing is a subscription based service like pay-as-you-go model which delivers software, infrastructure and the platform kind of services. These services are called as the Infrastructure as a service (IaaS), Platform as a service (PaaS), and Software as a service (SaaS) in the industry as shown in figure 1. Cloud computing is introduced to reduce the cost of the hardware and software. It also aims to make the next generation data center more powerful so that it will provide dynamic and flexible services to the consumer. Deployment of cloud computation makes the industry stronger and also gives the time to focus on innovation and creativity. This will lead the IT services to the different level and will help in developing the world [3].

Cloud computing is an evolution of the parallel computing, grid computing, and distributed computing. It deals with trading the resources in an efficient way according to the need of the user. Also, it is a large scale of heterogeneous resources that resides in the datacenter. The virtualization ability of the cloud computing hides the heterogeneity of the resources which makes it different from other computing introduced previously. The other feature such as user-oriented approach which gives the services as per need of the user and virtualization technology which is used to pack the resources makes it scalable and flexible.

The working of the cloud computing is to dispatch the tasks to the pool of resources which consists of several computers. It provides various services including storage, power, and several software services according to the need of the task [1, 2]. The business and virtualization technology used by the cloud computing has taken the technology to the new heights. As an example, it leaves the responsibility of resource allocation to the virtualization of virtual machine layer. Further, it pays more attention to the fairness of the resource allocation, likewise, various new features are introduced in cloud computing.

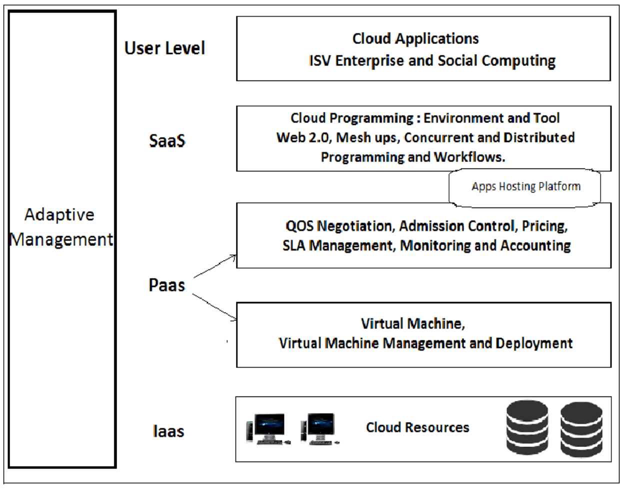

FIGURE 1 Architecture of cloud computing

This paper is organized into various parts. Section II is related work. Section III gives the description of the job scheduling algorithm based on fuzzy logic. Section IV describes the performance and result analysis. Section V provides the future work and conclusion.

2 Related work

Resource allocation of tasks has been core research topic hence there are plenty of papers published in this area. To use the technology of virtualization and allocation of resource efficiently, job scheduling algorithm plays an important role. There are various algorithms, strategies, and techniques proposed in past decade. Berger model [4] is introduced for the first time which discussed the fairness of the allocation of the task. This Algorithm based on Berger Model is taken from the Social theory of distributive justice. It established the dual constraint in which the tasks are classified according to the Quality of service parameters and define the justice evaluation function to judge the fairness in an allocation of the resource.

Bag of Tasks [5] schedules multiple applications which are a set of identical and independent tasks that runs on a heterogeneous master-slave platform. The goal of this algorithm is to minimize the ratio between the actual time taken by the task to execute in a system and the time taken to the task if executed alone. Next, workflow based scheduling algorithm [6] is introduced to find the solution that considers the user Quality of service of the task according to the preference of the user. Work provides the scheduling of the workflows that shows the significant improvement in the utilization of the CPU. This algorithm shows improvement in efficiency along with the fairness in the allocation of the task.

Previously papers have covered the efficiency of the resource allocation schemes but the fairness is missed out in most of the approaches. Those approaches who have considered the fairness of these resource allocation schemes suffer from the performance issues. Hence, this paper has proposed job scheduling algorithm which mainly resolves the problem of fairness while allocating jobs efficiently. Allocation of the jobs according to the Quality of service requirement of the user is one of the major issues in cloud computing, tasks are submitted to the virtual machine according to the availability of the resource. This paper introduces an algorithm which is used to evaluate the justification of the allocation of the task and adjusts the general expectation vector of the task. This paper uses fuzzy logic which helps in adjusting the general expectation vector of the task. It evaluates the fairness of the individual task as well as for classified task. Judging the fairness involves the justice evaluation function (JEF) which calculates the fairness of the allocation of the task [4, 7, 8].

Cloud computing is a pool of resources which consist of several virtual machine in the datacenter. Selecting an appropriate virtual machine is typical because there are thousands of virtual machines in a particular datacenter. To reduce the overhead of selecting it we have proposed a formula to deduce a set of the virtual machine(s) eligible to execute the task. It reduces the complexity of the algorithm and helps in getting better completion time and response time of the task which results in a better performance of the system that improves efficiency.

3 Proposed approach

In this paper, fairness constraint of the allocation of resource is served by the general expectation vector of the task. Task's general expectation vector is defined according to the classification of the task. The classification of the task is performed according to the quality of service *(QoS)* parameter of the task. These QoS parameters are Completion time *(CT)* and Bandwidth *(BW)*.

The general expectation vector of class type task I is: $e_i = [e_{i1}, e_{i2}, e_{i3}]$ and,

$$\sum_{j=1}^{3} e_{ij} = 1, \quad [4]$$

where e_{i1} = Weight of Number of CPU, E_{i2} = Weight of Memory, E_{i3}= Weight of Bandwidth.

Justice evaluation function of the task is calculated by the ratio of the actual reward to expected reward. If the result is more than one that means the task is executed gets more resource than it needed. If the value is less than one, it represents resource allocated is less than needed for the task. The resource is justified when the value of the justification evaluation function is one.

Justice evaluation function J_i for the task i:

$$J_i = \Theta \ln AR/ER, \quad \text{eq(1) [8]}$$

where *AR* is an actual resource allocated to the task and *ER* is an expected resource by the user.

The algorithm starts with the classification of the task according to the need of the user. Later, Actual execution time (AET) of the task is calculated using the file length of the task (FL) and the Million of Instruction per second of the Virtual Machines (MIPS). Now to select the machine(s) eligible for the execution of the task we compare it to the Actual execution time (AET) with the expected execution time (EET) along with this we need to compare the memory required by the task (MRT) and the memory size of the virtual machine (MVM). Similarly, the bandwidth required by the task is also compared with the bandwidth of each virtual machine. The number of tasks requirement of completion time is denoted as *k* and the number of tasks user requirement Bandwidth is *l*. Now justification of each class type is calculated using JCT_i and JBW_i. The Justification bandwidth of the completion time requirement of the task and the completion time of the bandwidth requirement of the task is assumed as 0.9. Now, to get the better mathematical association we have fuzzified JCT_i and JBW_i parameters using Mamdani Fuzzy inference system. It is passed with the rules base which is in the form of *if..else* rule shown in table 1and 2. Table 1 is a rule base for the task whose user requirement is completion time and table 2 is the rule base for the task whose user requirement is bandwidth. Defuzzification is performed using centroid method which is used to give the result in the range [0,1]. Now, Output after the defuzzification denoted with F_out is used to adjust the General expectation vector of classified tasks. The value used for adjusting the vectors are $\alpha_0 = 0.2$, a = 0.1 and b=0.2.

ALGORITHM

BEGIN
STEP 1: Let the set of virtual machine in the environment of cloud computing is
$VM = VM_1 \ldots\ldots VM_m$
STEP 2: Classify the task according to the Quality of service parameters of the user.
STEP 3: Calculate the Actual Execution time (AET) of the task i.
$AET = \frac{FL}{MIPS}$
STEP 4: Produce the set of Virtual Machine which is eligible to execute the task i.
If QoS = CT.
{
If MRT<MVM && AET<=EET
{

Add VM to the VM_i list
}
}
If QoS = BW
{
If (BW_i<BW_{vm})
{
Add VM to the VM_i list
}
}
VM_i={VM_1………VM_t}
Where, t<m

STEP 5: Normalize the virtual machine using the equation (2) to map it with the initial general vector of the task.
STEP 6: Calculate the Euclidean Distance of the virtual Machine.
STEP 7: Repeat from STEP 4 and STEP 5 for each task.
STEP 8: Evaluate the Justification Evaluation Function of task i using eq (1) for type 1 of the task.

$$JCT = \sum_{i=1}^{k} \frac{Actual_{CT}}{Expected_{CT}}$$

Similarly, for the type 2 of task,

$$JBW = \sum_{i=1}^{l} \frac{Actual_{BW}}{Expected_{BW}}$$

STEP 9: The parameter JCT_i and JBW_i are taken as the input. For implementing parameters in fuzzy inference system, we are using following steps.
Define a rule base which is a set of if...else rules.
$R = \{R_1, R_2, R_3, \dots, R_s\}$.
Fuzzify the input parameters using triangular membership function $\mu(v)$.
$\mu(v) \rightarrow [0, 1]$

$$\mu(x) = \begin{cases} 0, & v \le p \\ \frac{v-p}{r-p}, & p \le v \le r \\ \frac{q-v}{r-q}, & r \le v \le q \\ 0, & q \le v \end{cases}$$

Where,

$$x = \{CT, BW\}$$

Apply the inference on rule base based on the values of fuzzified input parameters.
R_i:
$if(L, M, H)\ AND\ if(L, M, H)\ AND\ if(L, M, H) \rightarrow T\_out(L, M, H)$.

$$L \rightarrow Low, M \rightarrow Medium, H \rightarrow High$$

Generate $F\_out(L, M, H) \rightarrow [0,1]$.
Defuzzify the produced F_out by FIS using centroid method,
$F\_out \rightarrow [F_{min}, F_{max}]$
STEP 9: Adjust the General Expectation Vector of the task by using the output of the fuzzy inference system for each classified task.

$$\propto = \propto_0 - a(F\_out - b)$$

END

TABLE 1 Rule base for the task which requires Completion time as QoS parameter

S. no.	Justification Completion time	Justification Bandwidth	Output (F_out)
1	LOW	HIGH	MEDIUM
2	MEDIUM	HIGH	MEDIUM
3	HIGH	HIGH	HIGH

TABLE 2 Rule base for the task which requires Bandwidth as QoS parameter

S. no.	Justification Completion time	Justification Bandwidth	Output
1	HIGH	LOW	MEDIUM
2	HIGH	MEDIUM	MEDIUM
3	HIGH	HIGH	HIGH

4 Performance analysis

The performance of this algorithm is simulated on CloudSim Platform [9]. CloudSim is simulation software which is introduced by University of Melbourne in 2009. There are four levels in this framework that includes GridSim, CloudSim, SimJava, and UserCode. The implementation of the proposed algorithm is performed in CloudSim layer and simulation program in UserCode layer. The proposed algorithm is added in CloudSim by overloading *bindCloudletToVM()* method of the DatacenterBroker class. Cloudlet class is also extended by adding two variables such as expectation completion time and expectation bandwidth and four methods are used to get and set the value of these variable. The fuzzy logic is applied using MATLAB. It has the Fuzzy inference engine which is used to fuzzify and defuzzify the parameters.

The value of the expectation vector of the class 1 type of task is initially taken as e_1 = [0.7, 0.1, 0.2] and the expectation vector of the class 2 type task is e_2 = [0.3, 0.2, 0.5]. e_1 and e_2 are empirical value and are determined in the same experimental environment. To create a cloud computing scenario we have taken data which is listed in table 3 and 4. In table 3, 0 to 3 tasks are class 1 type of task and 4 to 7 are class 2 type of task. Table 3 is the set of the virtual machine which has different preference and performance, data is listed in the table.

Figure 2 and figure 4 is a rule base for the class 1 and 2 type of task in MATLAB, respectively. Similarly, Figure 3 and figure 5 shows the surface view in MATLAB for class 1 and class 2 type of task.

Algorithm 1 is Fuzzy based job scheduling algorithm. Algorithm 2 is job scheduling algorithm based on Berger model. Execution time of task attained by the comparative analysis of result after simulation is demonstrated is figure 6. Overall, after execution, it is noted that efficiency of the algorithm 1 is better than algorithm 2. Task 0-3 prefers high computing power. Also, the completion time of Algorithm 1 is better than algorithm 2.

TABLE 3 Virtual machine parameters

VMid	CPU	Memory	Bandwidth
0	4	2048	1200
1	2	1024	3000
2	2	1024	1000
3	1	512	1200

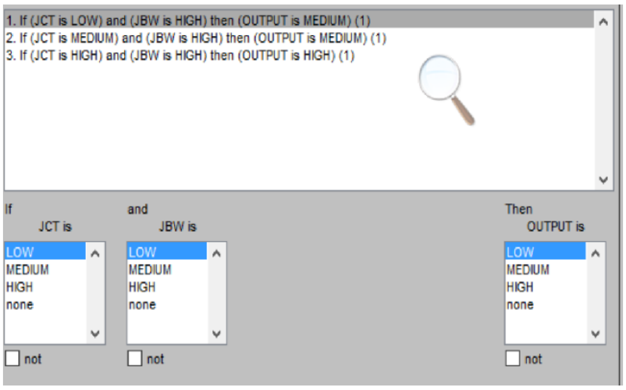

FIGURE 2 Rule base for the class 1 type of task in MATLAB

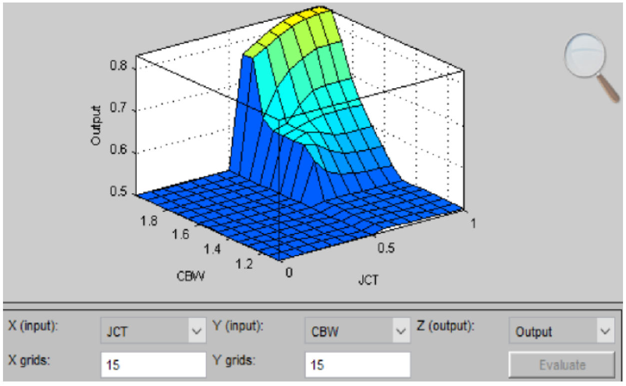

FIGURE 3 Surface View of the parameters in MATLAB for class 1 type of Task

TABLE 4 Task Parameters

CloudletId	Classtype	Length	File_size	Output_size	Expectationtime	ExpectationBW
0	1	4000	2500	500	400	–
1	1	3000	2000	400	200	–
2	1	2000	800	300	150	–
3	1	5000	5000	2000	500	–
4	2	2000	800	300	–	2000
5	2	3000	2000	400	–	3000
6	2	800	300	300	–	1200
7	2	2500	1000	500	–	2000

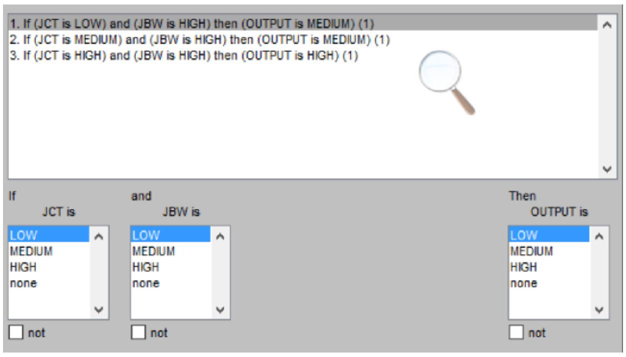

FIGURE 4 Rule base for the class 2 type of task in MATLAB

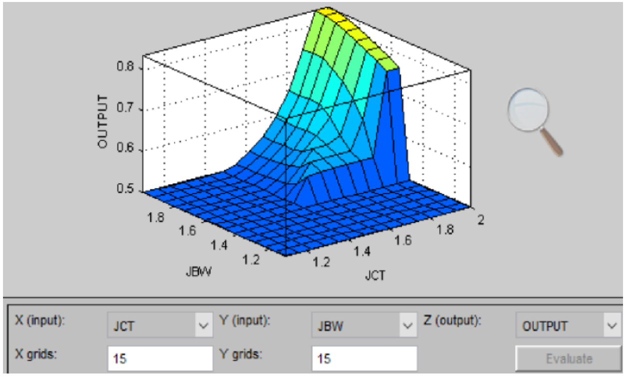

FIGURE 5 Surface View of the parameters in MATLAB for class 2 type Task

Figure 7 presents the comparison of the Justification of user task (J value). J value 0 means the allocation of the resource achieved by the user task are stable with their expectation. J >0 refers that the user achieved resources are higher than its expectation. J<0 refers that the user doesn't meet its expectation. Clearly, Algorithm 1 provides more satisfaction than Algorithm 2.

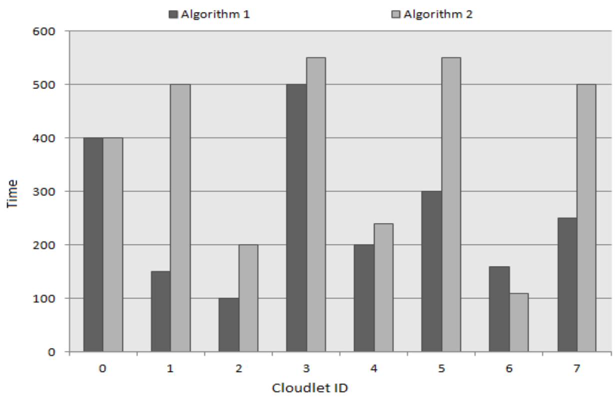

FIGURE 6 Task Execution time comparison

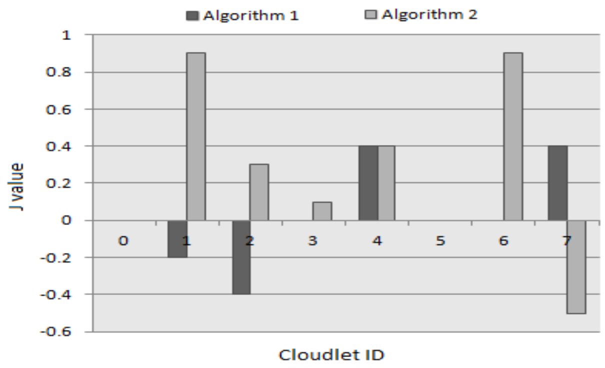

FIGURE 7 User Satisfaction

Class 1 type of task is for the high computing power hence, demands earliest completion time. Figure 7 shows the comparison of the CPU number of the allocated virtual machine for the first type of task. Overall, First algorithm ensures better high performance computation to the task which meets task preference with fairness.

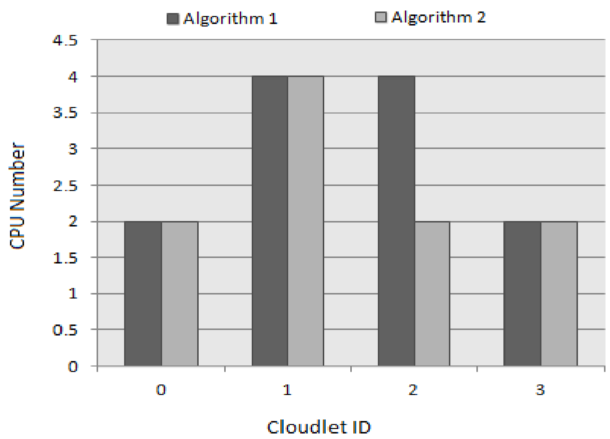

FIGURE 8 Comparison of the first type of task

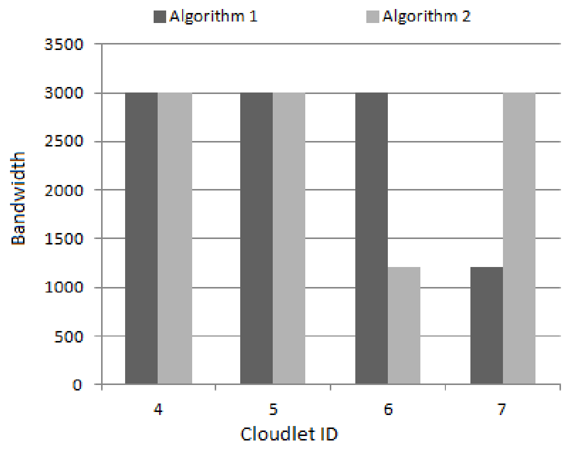

FIGURE 9 Comparison of the Second type of task

Class 2 type of class prefers high bandwidth, which includes task 4-7. Figure 8 shows the comparison of the allocated bandwidth of the virtual machine for the second type of task. Evidently, Algorithm 1 meets better preferences of task with better fairness.

5 Conclusion and future work

In this paper, expectation vector of the task is adjusted using fuzzy logic so that that the fairness in the allocation of the task can be performed. Scheduling algorithm is implemented and verified using extension of CloudSim. The results are compared with the Berger model which shows that the proposed algorithm performs better. It also shows better mathematical association between the parameters to adjust the expectation vector of the classified task. Proposed approach also performed modification in selection of supported virtual machine for a particular task that increased the efficiency of the system.

Further, for accurate results we can use the neural network for quality of service vector of the task to adjust the expectation vector. This will give a better mapping between the task and resources and fairer value of the general vector of the task.

References

[1] Liu Peng *The definition of cloud computing and characteristics* http://www.chinacloud.cn/.2009-2-25

[2] *IBM center for high performance on demand solutions* IBM Cloud Computing White Paper http://www.ibm.com/developerworks/websphere/zones/hipods/

[3] Armbrust M, Fox A, Griffith R, Joseph A, Katz R, Konwinski A, Lee G, Patterson D, Rabkin A, Stoica I, Zaharia M 2010 A view of cloud computing *Communications of the ACM* **53**(4) 50–8 Trans. Roy. Soc. London, vol. A247, 529-51, April 1955

[4] XU Baomin C Z 2011 Job scheduling algorithm based on berger model in cloud environment *Advances in Engineering Software* **42** 419–25

[5] M. L. P. J.-F. R. Y. V. F. Benoit A 2011 Offline and online master-worker scheduling of concurrent bags-of-tasks on heterogeneous platforms *Parallel and Distributed Processing* **190** CCIS, 1–8

[6] Varalakshmi A B a P V P, Aravindh Ramaswamy 2011 An optimal workflow based scheduling and resource allocation in cloud *Advances in Computing and Communications* 411–20

[7] Berger Z M J, Cohen B P 1966 Status characteristics and expectation states *Sociological theories in progress* **1** 29–46

[8] Berger C T Z M J, Cohen B P 1966 Status characteristics and expectation states: a process model *Sociological theories in progress* **1** 47–74

[9] Buyya R, Ranjan R, Calheiros R N 2009 Modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities *Proceedings of the Conference on High Performance Computing and Simulation (HPCS 2009), Leipzig, Germany. IEEE Press: New York, U.S.A., 21–24 June 2009* **21** 1–11

[10] Sutherland I E 1968 A futures market in computer time *Commun ACM* **11**(6) 449–51

[11] Rajkumar Buyya, Rajiv Ranjan, Calheiros R N 2009 Modeling and simulation of scalable cloud computing environments and the CloudSimToolkit: challenges and opportunities *In: Proceedings of the seventh high performance computing and simulation conference (HPCS 2009, ISBN:978-1-4244-49071), Leipzig, Germany. New York, USA: IEEE Press; June 21–24, 2009*

[12] Braun T D, Siegel H J, Beck N, et al. 2001 A Comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributedcomputing system *J Parallel Dist Comput* **61**(6) 810–37

[13] LI Mao-Sheng, YANG Shou-Bao, FU Qian-Fei1, YANG Jin 2006 A Grid Resource Transaction Model Based on Compensation *J Softw* **17**(3) 472–80

[14] Berger J, Zelditch Jr M, Anderson B, Cohen B P 1972 Structural aspects of distributive justice: a status value equation *Sociological*

theories in progress **2** Boston: Houghton 119–46

[15] Broberg J, Venugopal S, Buyya R 2008 Market-oriented grids and utility computing: the state-of-the-art and future directions *J Grid Comput* **6**(3) 255–76

[16] Kumar S, Dutta K, Mookerjee V 2009 Maximizing business value by optimal assignment of jobs to resources in grid computing *Eur J Oper Res* **194**(3) 856–72

[17] Rajkumar Buyya, Chee Shin Yeo, Srikumar Venugopal 2008 Market-oriented cloud computing: vision, hype, and reality for delivering IT services as computing utilities *In: Proceedings of the tenth IEEE international conference on high performance computing and communications (HPCC 2008), Dalian, China. Los Alamitos, CA, USA: IEEE CS Press; September 25–27*

AUTHORS

Pratibha Pandey

Current position: Guest Faculty, Department of Computer Sc. & Engineering M. M.M.University of Technology, Gorakhpur-273 010 Uttar Pradesh, India
University studied: M.tech from MMMUT,India
Scientific interest: Cloud Computing, Networking and Operating System.
Publications: two other papers in the same area.
Experience: 9 months of experience in Teaching.

Sarvpal Singh

Current position: Associate Professor, Department of Computer Sc. & Engineering M.M.M.University of Technology, Gorakhpur-273 010 Uttar Pradesh, India
University studied: Phd awarded from Department of Computer Science, Deen Dayal Upadhyay Gorakhpur University, Gorakhpur, India, 2013.
Scientific interest: Wired/Wireless Networks (Mobile Database System)
Publications: 50
Experience: about 22 years & 6 months