QoS based Modified Bat Algorithm for Task Scheduling in Cloud

Savita Khurana, Rajesh Kumar Singh

Abstract: ***Cloud computing is prominent technology in the current era. Every company shifted their infrastructure on the cloud to speed up their operation in cost effective way. To speed up the operations, various task scheduling techniques and algorithms were proposed so far but still most of them are not considering both Quality of Service (QoS) and virtual machine optimization factor, which is the upmost important parameter to satisfy the needs of user and effective utilization of resources. This paper propose an approach which considering the multi-objectives QoS parameters-Makespan, cost, user's budget and resource utilization using nature inspired modified bat algorithm. The objective of the proposed approach is to minimize the Makespan, cost by considering user's budget and improve the cloud resource utilization. To fulfill and validation of these objective, proposed approach designed a fitness function.***

Index Terms: ***Cloud Computing; Task Scheduling; Bat Algorithm; QoS; Makespan.***

I. INTRODUCTION

Cloud computing is the most favourite and valuable technology adopted by today's IT industry. It companies migrate their services, infrastructure in cloud to provides services using large scale cloud environment. The cloud computing increases the reliability of services with cost effective way [16]. It is the extension of existing computing i.e. grid and parallel and distributed computing. Most popular and large scale applications i.e. amazon, google use cloud environment to enhance their capabilities, increase reliabilities and provides the services to the users with lesser cost Cloud computing is a internet based computing uses pay-per-go model which means user pay according to usage of resources[19]. . The virtualization is the most important and prominent feature of cloud which is logically division of hardware resources. The machine on which virtual machine is produced is called as host machine and VM is known as guest machine, which is controlled by software called hypervisor. Virtualization provides elasticity, cost effectiveness, infrastructure independency, location independence and scalability. But to effective utilization of VM's is still a challenging task which requires an effective scheduling algorithm to schedule the task on most appropriate resources. Hence, efficient task scheduling algorithm is considered as main challenges in cloud environment as it affect the utilization of resource, achieving even workload at all nodes, increases response time and minimized completion time of tasks[18].

The Waiting time of tasks directly affect the responsiveness of the VMs/nodes which further affect the overall resource utilization and cost of any cloud service provider. So a reliable and efficient scheduler is desirable in cloud environment. In Fig. 1 working of cloud task scheduling is shown. The client submits the tasks through its cloud portal. Resources are not directly allocated to the client, instead before provisioning the resources scheduler firstly checks the resources database based upon quality of service parameters [26]. After verifying the QoS parameters and resource availability, scheduler allocates the resources from resource pool [23]. After completion of tasks, resources are returned back to the resource pool. Scheduler used a number of strategies to schedule the tasks i.e. static as well as dynamic.

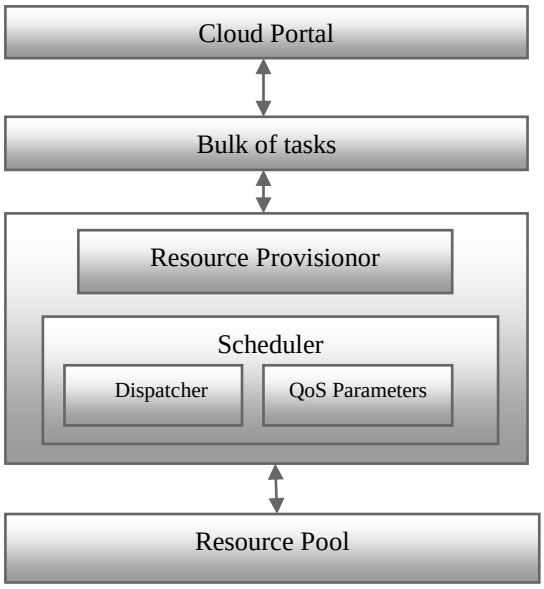

Fig.1: Working of cloud task scheduling

Now a days, researchers are working on optimization algorithms in cloud computing. These algorithms focus only on scheduling of the tasks but not considered the various

Revised Manuscript Received on December 22, 2018.
Savita Khurana, Department of Computer Science & Engineering, I.K.Gujral Punjab Technical University, Jalandhar, Punjab, India.
Rajesh Kumar Singh, Department of computer science & application, SUS Institute of Comupter, Tangori, Mohali, Punjab, India.

Quality of Service parameters like budget, time, reliability and energy etc. Various optimization algorithms are available like Ant Colony optimization, Genetic algorithm, PSO and BAT algorithm etc have their own significance and limitations. In this paper, BAT algorithm is considered for optimization. It is a meta-heuristic algorithm for searching of prey by bats. Similarly in task scheduling, resources are analogous to prey and various tasks which need to be executed are considered as bats. The resource which provide less time of execution and cost is considered best for a task. In this way tasks are scheduled in a controlled manner using BAT algorithm.

II. RELATED WORK

Yang (2010) proposed Bat Algorithm (BA), an optimization-based meta-heuristic algorithm and relied on searching for the prey of the bats. It considered the echolocation behaviour of bats. The significance of Bat algorithm is that it can provide very quick convergence at a very initial stage by switch from discovery to utilization [1]. The results of BAT are better than PSO and GA.

J.Huang [2] proposed Genetic Algorithm based workflow task scheduling model in the Cloud Computing environment. In this algorithm genetic factor is calculated to meet the task characteristics. It improves the performance of system for workflow task scheduling. The experimental setup showed an improvement in terms of resource utilization and also, reduction in the task completion time.

Xin-She Yang developed a bat-inspired algorithm to target multi-objective optimization problems, that is, solving problems based on global optima and non-linearity. The algorithm focused on validation of subset of test functions, and then capturing the design optimization benchmark in structural engineering. Results showed the supremacy of proposed algorithm over traditional bat algorithm [3].

Chen et al. [4] illustrated an Ant Colony based workflow scheduling algorithm along with the addition of various new parameters and features to attain modified and enhanced results. It considered two parameters cost and deadline. It targeted the cost reduction without losing the focus on deadline. Two types of pheromone were taken into consideration, one was the minimization of makespan time and the other was cost reduction. Heuristic based information was added to guide the ants in fetching their ways. Every ant counted upon one pheromone-type and one heuristics-type, while considering the mentioned parameters' probabilities, in every next iteration. Adaptive modification was done for the parameters in the algorithm [21].

Wen et al. [5] illustrated an ACO algorithm which was merged with various algorithms, namely, Particle Swarm Optimization (PSO) approach to attain high performance. The proposed technique improved the rate of convergence as well as diluted the resource consumption. The approach was success in terms of attaining global results instead of getting trapped in local optima.

Tao et. al., proposed an approach for workflow scheduling based upon Pareto-solution relying on Genetic Algorithm. The two parameters which were considered for optimization were energy conservation and makespan time [6]. Multi-parent crossover and library case were the key highlights of the approach.A library case is the term comprising vectors for task-type, Pareto solutions and dependency matrix for considered tasks. The experimental results proved that the approach worked efficiently in terms of solution diversification, convergence rate and reliability factor.

Jacob [7] modified BAT algorithm for resource scheduling in the cloud computing environment to reduce the makespan time. The designed algorithm provided less incurred cost and précised results in comparison to Genetic Algorithm.

Kumar et al. [8] designed an approach by collaborating the Gravitational Scheduling Approach and the BAT Algorithm to schedule the upcoming tasks in the cloud environment. In this approach a trust model is designed which has been taken deadline as a constraint. The Resources are selected on the behalf of their trust factor for execution of the tasks. CloudSim was used to simulate the results. To prove the efficiency, proposed approach was compared with Random Resource Selection using GSA and was found to be effective in terms of minimizing the makespan time and failure ratio.

Raghavan et al. [9] enhanced the Bat algorithm for scheduling the workflow in cloud environment in order to overpower the processing cost of the overall workflow scenario. The approach was then compared to Best Resource Selection Technique and it was concluded that it efficiently helped in reduction of processing cost incurred for the workflow.

George et al. combined Multi-Objective Bat Algorithm [10] and Particle Swarm Optimization in context of maximizing the profit value in the cloud environment. PSO yield locally optimized results which were further enhanced by BAT Algorithm to attain globally optimized solutions. Simulated results proved the efficiency of results so obtained.

Yang et. al., proposed iterative ordinal optimizations algorithm[11], which was designed to reduce the execution time. Minimum execution time directly affect the cost, so overall expenses has been decreased.. Proposed approach was compared with Monte Carlo and blind pick method, which clearly shows the supremacy of proposed algorithm. This algorithm is also working on scientific workflows in different clouds to provide better results.

Wu, Zhangjun, et al. [12] proposed market oriented scheduling cloud workflow. In this approach two perspectives of workflow scheduling i.e. service level which also used Random scheduling and task level scheduling used ant colony optimization(ACO), particle swarm optimization(PSO) and

Genetic algorithm(GA) for mapping the tasks to virtual machines(VMs). All the three algorithms are compared with respect to Quality of service parameters i.e. Makespan, Cost and no of CPU cycles. The implementation of the algorithm is on SwinDev-C [13] which is specifically designed for cloud.

Zuo et al. [14] presented Improved Ant colony optimization (IACO) algorithm using the cost effective model. In this paper, for achieving the optimal solution, two constraint functions are used for reducing the budget of user and improving the performance of resources. The proposed algorithm provides better results as compare to min-min algorithm.

Shengjun Xue, Wenling shi and Xiaolong Xu [15] proposed a Service-Cost Optimization based Particle Swarm Optimization algorithm for task scheduling. The task completion time as well as its execution time is reduced while considering the user's defined cost. This algorithm enhanced the performance of cloud scheduler by providing better results.

III. SIMPLIFIED BAT ALGORITHM

Simplified Bat algorithm is implemented by Xin-She Yang [1]. It is meta-heuristic algorithm used for optimization of workflow scheduling. Various types of meta-heuristic algorithms are available for scheduling of tasks. But the major problem to find best solution or optimized solution by considering the user's constraints likes budget, time, makespan etc. Bat algorithm uses the principle of randomization for finding the optimal solution[17]. Several types of standard functions are available such as Rastigin's function, Ackley's function, Butterfly method for testing of algorithm [22][24]. An algorithm which has capability to reduce these standardized methods with accuracy is considered as good algorithm. In the same way, simplified bat algorithm is one of the best optimized techniques which provide correct result. The mapping of tasks to appropriate resources in workflow scheduling is the major task of workflow scheduling. Mapping of tasks is an NP-hard problem.

Simplified Bat algorithm was designed to include properties of meta-heuristic algorithms in such as a way that it could present better than other algorithms by considering the pros and cons of all the other algorithms. Various types of scheduling algorithms are available for comparison like simulated annealing, genetic algorithm and particle swarm optimization [20] etc. The simplified bat algorithm accomplished positive results as compare to other meta-heuristic algorithms. Xin-She Yang designed simplified Bat algorithm in 2010 [3]. Bat algorithms used the echolocation properties of microbats. The bats generally search for prey or food. They emit high frequency pulses to find the prey and listen echo which are reflected back from the prey or any obstacle. Bats also have property to recognize among prey/objects. There are various parameters such as wavelength, frequency, loudness and velocity[25]. They have ability to regulate the wavelength and frequency of the emit pulse rate based upon distance from the prey. The following are the assumptions of bat algorithms [1]:-

1. All bats use the feature of echolocation to sense the distance from the prey.

2. Initialize velocity V_i, position X_i with a frequency F_{min}, wavelength λ and loudness A0 randomly to hunt for prey. They emitted pulses by adjusting their wavelength and adjust the pulse rate emission $r \in [0,1]$ according to the target prey.

3. As the bats are nearer to prey, loudness A_0 is decreasing and pulse emission rate is increasing. When A_0 becomes 0; it means bat catches its prey.

Algorithm1. Pseudo Code of simplified BA

1) Objective function:F(x), where x=(x1,x2,….xd)
2) Initialize bat variables Xi and velocity Vi, where i=1.....n.
3) Initialize the parameters - frequency Fi, pulse rate Ri and Loudness of Bat Ai.
4) while (t<total no. of iterations)
5) Spawn new solutions by regulating frequency, and updating velocities and location/solutions.
6) F (rand>Ri)
7) Choose a best solution out of the available solutions.
8) Spawn a local optimization around the selected best solution.
9) end if.
10) if (rand< Ai && f(xi)< f(x*))
11) Accept new solutions
12) Increase ri, reduce Ai
13) end if
14) Discover the current best x*
15) end while

IV. PROPOSED APPROACH

Task scheduling Model in cloud computing.

Consider a cloud user submitted a set of *n* tasks such that *A*={*T1, T2, T3,* ⋅ ⋅ *, Tn}*. The users submit the workflow task to the cloud service provider for execution. Cloud provider has a variety of virtual machines (VMs) having varying capability of execution, storage, communication etc.

Cloud provider provides these VMs to the users according to their workflow by discussing a service level agreement (SLA).

Table I: Notations used

T1,T2,…..,Tn	Number of tasks
M	Number of VMs
ST_i	Size of task ST_i in Million of Instructions
BW(m,n)	Bandwidth between Virtual machines
ET_VM_i	Execution time of VM_i
SR_i	Success Rate of VM_i
C(i,k)	Cost of between tasks
FF	Fitness Function of modified BAT

The main elements of the supposed cloud model are following:-

1. Cloud user: initially cloud user submits the workflow task for execution to the cloud broker**.** These tasks are forwarded to the cloud broker via cloud service provider (CSP).

2. Cloud broker: cloud broker considered user requests. The main task of broker is to preserve the eminence of the Virtual machines. It allocates the resources to the users as per their quality of service parameters.

3. Cloud service provider (CSP): it is the interface between the users of a cloud and the broker. Service provider has a pool of all active VMs. The task assignment to the active VM is done by the specific task scheduling algorithm.

Fitness function is the main core of any meta-heuristic algorithm like BAT algorithm. The purpose of this paper is to schedule the tasks in such away that minimize the makespan, increase success rate within budget of the user, maximize the resource utilization. If n number of tasks are assigned to VMs then the execution time of virtual machine, cost and success rate of the tasks are calculated as follows:

Notations and Terminologies

The following terminologies are used to find the fitness function(FF) for modified BAT algorithm. The notations used are given in the Table I.

1. **Processor speed** (*PS*): It is the computation time of a VM and is calculated in million instructions per second (MIPS) and represented by PS_j.

2. **Task Execution time** (ET_{ij}): Assume T_1, T_2, T_3, …., T_n n Tasks, where the ST_i is the size of every task T_i , measured in million instructions. Task execution time of T_i is calculated mathematically as follows.

$$ET_{ij} = \frac{STi}{PSj} \qquad (1)$$

3. **Average Execution time** (Avg_ET): It is the mean of the execution time of the available VMs. In other words,

$$Avg\_ET = \frac{\sum_{j=1}^{m} ETij}{m} \qquad (2)$$

4. **Cost (C)**: Consider a task T_i be scheduled on a VM_m and a successor task T_k be scheduled on a VM_n. The amount of data communication between the tasks T_i and T_k is represented as data(i,k). The bandwidth between two virtual machines VM_m and VM_n be represented as BW(m, n). The data communication cost between the two tasks T_i and T_k is calculated as

$$C(i,k) = \begin{cases} -\frac{data(i,k)}{BW(m,n)}, & if\ m \neq n \\ 0, & if\ m = n \end{cases} \qquad (3)$$

5. Success Rate (SR): it is denoted as SR_m. It is the number of tasks successfully executed over a virtual machine divided by the total number of tasks submitted over it.

$$SR_{m=}\frac{\text{No.of tasks successfully executed on mth VM}}{\text{Total no.of tasks submitted}} \qquad (4)$$

6. Cloud efficiency (*Avg.*C_{eff}): It is average of success rate of the all virtual machines over a cloud. Therefore,

$$Avg.C_{eff} = \frac{\sum_{i=1}^{m} SRi}{m} \qquad (5)$$

Here, m is the number of virtual machines over a cloud.

We denote the fitness function by FF. If ET is the maximum execution time of the task between all the Bat individuals and B is the user defined budget then the fitness function is designed as follows:

$$FF = \begin{cases} \frac{ET}{Avg\_ET} * \frac{1}{Avg.\ Ceff}, & if\ \frac{C(T)}{B} \leq 1 \\ \frac{C(T)}{B}, & if\ \frac{C(T)}{B} > 1 \end{cases}$$

Task are executed within the specified QoS constraint then value of fitness function (FF) is less than 1. If FF value is greater than 1, it means that fitness function is unfit to scheduled corresponding task.

V. SIMULATION RESULTS

The proposed algorithm has been implemented on the cloudSim simulator. The proposed algorithm is validated with the help of 8 virtual machines. Each VM consist of 4 GB RAM, core i5 with 2.6GHz processor speed, 32GB Hard Disk with MIPS rate varying from 1000 to 3000. Initial population size is randomly generated and maps the tasks to the machine with minimum makespan. The simulated results are shown in Table II and compare the results with PSO and simplified BAT algorithm which shows the supremacy of the proposed approach.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Table II: Makespan Comparison

Sequence No.	Algorithms	Number of Tasks	Makespan
1	PSO	50	106
	Bat Algorithm		90
	Proposed Modified Bat		87
2	PSO	100	255
	Bat Algorithm		235
	Proposed Modified Bat		230
3	PSO	200	590
	Bat Algorithm		582
	Proposed Modified Bat		560
4	PSO	400	1120
	Bat Algorithm		1092
	Proposed Modified Bat		1015
5	PSO	800	2820
	Bat Algorithm		2710
	Proposed Modified Bat		2640

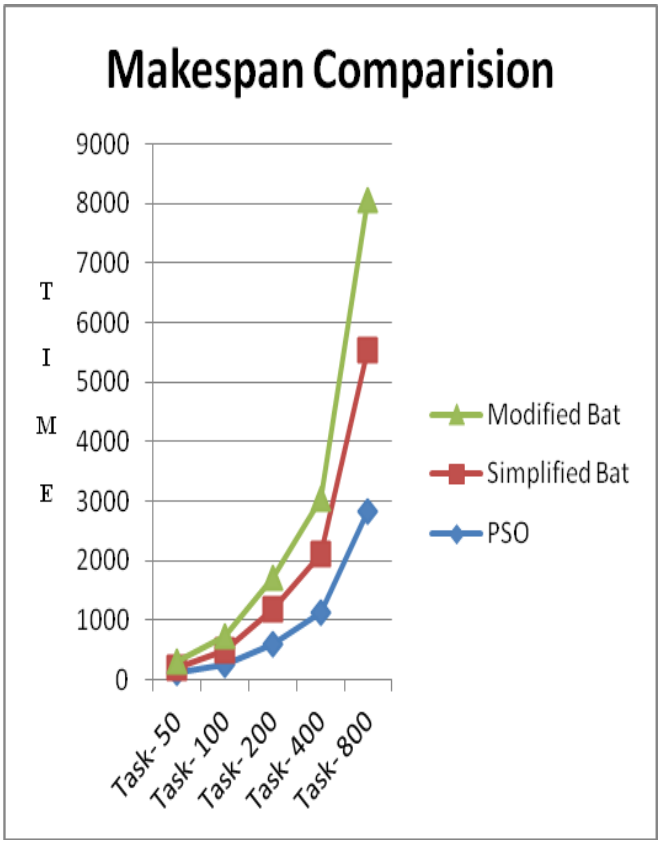

Fig. 2: Makespan comparison for different no. of tasks.

The QoS based modified BAT algorithm's cost is compare with the simplified BAT and PSO.

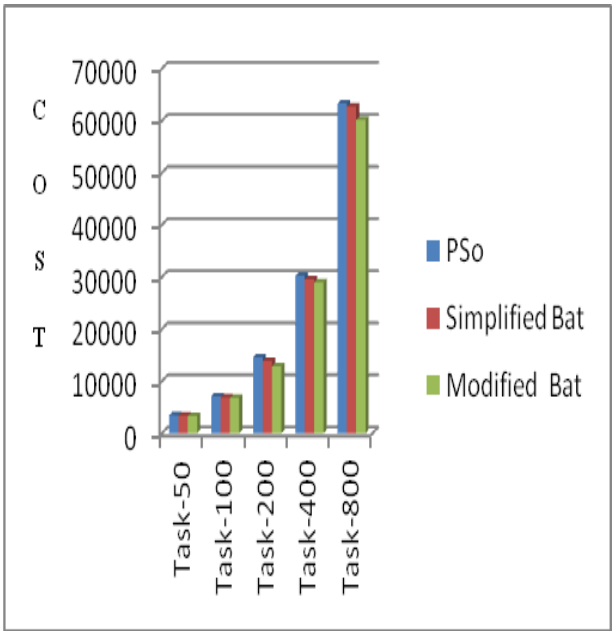

Fig. 3: Comparison of scheduling algorithms based upon cost

Task execution cost is calculated by taking the cost of communication between the virtual machines and computing cost by considering the bandwidth between the two virtual machines. Cost value is calculated against the user's defined budget. The results of proposed work are shown in Fig. 3 and compared with the existing meta-heuristic algorithm PSO and simplified BAT algorithm. It outperforms than simplified BAT algorithm and PSO in respect to Makespan, cost and cloud efficiency.

VI. CONCLUSION

Task scheduling is one of the major factor in computing which effects the cloud resource utilization. Various task scheduling algorithms like simplified BAT, PSO, genetic etc. algorithms are available. In this paper a modified- BAT algorithm has been proposed for task scheduling in cloud computing to improve the performance of simple BAT algorithm. In this paper, a fitness function is designed by considering the conflicting QoS parameters- Makespan, cost, user's budget, Machine's success rate and cloud resources efficiency. Fitness function is used to map the tasks to VMs in such a way that minimize the Makespan, cost and increase the resource utilization rate as its concurrence rate is growing. The results are simulated with the help of cloudSim simulator. The results of proposed algorithm are shown better results as compare to the other meta-heuristic algorithms-PSO, simplified BAT algorithm.

REFERENCES

1. Yang X. S. (2010), "A New Meta heuristic Bat-Inspired Algorithm, in: Nature Inspired Cooperative Strategies for Optimization, Studies in Computational Intelligence", Springer Berlin, Springer, 284, 65-74.

2. Huang J, "The Workflow Task Scheduling Algorithm Based on the GA Model in the Cloud Computing Environment." Journal of Software, Vol. 9, No 4, PP. 873-880, April 2014.
3. Yang X. S., (2011), "Bat algorithm for multi-objective optimisation", International Journal of Bio-Inspired Computation, Vol. 3, No.5, pp. 267-274.
4. Chen W.-N, Zhang J, Yu Y.,"Workflow scheduling in grids: an ant colony optimization", IEEE Congress Evolutionary Computation, pp. 3308-3315, 2007, DOI: 10.1109/CEC.2007.4424898
5. Wen X, Huang M, Shi J," Study on resources scheduling based on ACO algorithm and PSO algorithm in cloud computing", 11th International Symposium on Distributed Computing and Applications to Business, Engineering & Science, pp. 19–22, 2012, http://dx.doi.org/10.1109/DCABES.2012.63.
6. Tao F, Feng Y, Zhang L, Liao T.W., "CLPS-GA: a case library and Pareto solution-based hybrid genetic algorithm for energy-aware cloud service scheduling", Applied Soft Computing, Vol- 19 (2014), pp. 264-279, 10.1016/j.asoc.2014.01.036.
7. Jacob L., "Bat algorithm for resource scheduling in cloud computing", International Journal for Research in Applied Science and Engineering Technology, 2 (2014), pp. 53-57.
8. Kumar V. S., Aramudhan,"Hybrid optimized list scheduling and trust based resource selection in cloud computing", Journal of Theoretical and Applied Information Technology, 69, pp. 434-442, 2014.
9. Raghavan S, Marimuthu, C, Sarwesh, P, & Chandrasekaran K., "Bat algorithm for scheduling workflow applications in cloud", International Conference Electronic Design, Computer Networks & Automated Verification (EDCAV), IEEE, pp. 139–44, 2015.
10. George S, "Hybrid PSO-MOBA for profit maximization in cloud computing", Published in International Journal of Advanced Computer Science and Applications(IJACSA), Volume 6 Issue 2, 159–63, 2015.
11. Yang X.S., "Bat algorithm for multi-objective optimisation", International Journal of Bio-Inspired Computation, 3(5):267–274, 2011.
12. Wu Z., Ni Z., Gu L., Liu X.,"A revised discrete particle swarm optimization for cloud workflow scheduling", International Conference on Computational Intelligence and Security (CIS), IEEE, 2010.
13. Malawski M, et al. "Cost-and deadline-constrained and provisioning for scientific workflow ensembles in IaaS clouds", Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis. IEEE Computer Society Press, 2012.
14. Zuo L, Shu L, Dong S, Zhu C, "A Multi-Objective Optimization Scheduling Method Based on the Ant Colony Algorithm in Cloud Computing", Article in IEEE access 3:1-1 · December 2015, DOI: 10.1109/ACCESS.2015.2508940
15. Xue S, Shi W. and Xu X," A Heuristic Scheduling Algorithm based on PSO in the Cloud Computing Environment", International Journal of u- and e- Services, Science and Technology, Vol-.09, pp.349-362, 2016.
16. Bawa R.K. and Sharma G, "Reliable resource selection in grid environment". International Journal of Grid Computing & Applications, Volume 3, Number 1, pp. 1-10 2012.
17. Rani T. S, Kannan S, "Task Scheduling on Virtual Machines using BAT Strategy for Efficient Utilization of Resources in Cloud Environment", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 17 (2017) pp. 6663-6669.
18. Hussin M., Fatma A., "An Enhanced Task Scheduling Algorithm on Cloud Computing Environment", International Journal of Grid and Distributed Computing, Vol. 9, No. 7 (2016), pp.91-100, http://dx.doi.org/10.14257/ijgdc.2016.9.7.10.
19. Bawa R. K., and Sharma G., "Modified Min –Min Heuristic for Job Scheduling based on QoS in Grid Environment", IEEE Conference on International Management in the Knowledge Economy, 19-20 December 2013.
20. Abdi S., Motamedi S. A. and Sharifian S., "Task Scheduling using Modified PSO Algorithm in Cloud Computing Environment", International Conference on Machine Learning, Electrical and Mechanical Engineering (ICMLEME'2014) Jan. 8-9, 2014 Dubai (UAE).
21. Talal H, and Bellaiche M., "Evaluation and selection of Cloud security services based on Multi-Criteria Analysis MCA." International Conference on Computing, Networking and Communications (ICNC), IEEE, 2017. DOI: 10.1109/ICCNC.2017.7876216
22. Cai X, Gao X.-Z., Xue Y., "Improved bat algorithm with optimal forage strategy and random disturbance strategy", International Journal of Bio-Inspired Computation, 2017, https://doi.org/10.1504/IJBIC.2016.078666
23. Syed Hamid Hussain et. Al.,, Resource scheduling for infrastructure as a service (iaas) in cloud computing: Challenges and opportunities, Journal of Network and Computer Applications, Volume 68, June 2016, pp. 173-200, https://doi.org/10.1016/j.jnca.2016.04.016
24. Gitanjali J. et. al., "Scheduling Tasks Using Bat's Algorithm in Cloud Sim", International Journal of Pure and Applied Mathematics,Vol- 118 No. 18, 3919-3928, 2018.
25. Sagnika S. et. al., "Workflow Scheduling in Cloud Computing Environment using Bat Algorithm," Proceedings of First International Conference on Smart System, Innovations and Computing", DOI:10.1007/978-981-10-5828-8_15, Jan 2018.
26. Kalra M. and Singh S, "A review of metaheuristic scheduling techniques in cloud computing", Egyptian Informatics Journal (2015) 16, 275-295.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication