IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-6, Issue-4 E-ISSN: 2347-2693

Task scheduling algorithm optimization based On hybrid HBO and ACO in cloud computing

T. Chopra[1*], R. Singh[1]

[1]Dept of Computer Science and Engineering, IKGPTU Main Campus, Kapurthala, INDIA
[2]Dept of Computer Science and Engineering, IKGPTU Main Campus, Kapurthala, INDIA

**Corresponding Author: tanuchopra.166@gmail.com*

Available online at: www.ijcseonline.org

Received: 20/Mar/2018, Revised: 28/Mar/2018, Accepted: 19/Apr/2018, Published: 30/Apr/2018

Abstract— Cloud computing is a shared pool of compute, storage and network resources that are elastic in nature and can dynamically scale to meet changing demands of an IT organization. Every IT organization has to invest in hardware, software resources to make the business run effectively. Cloud offers guaranteed and reliable access to these resources on pay-as-you-use manner. Increasing demand of cloud services leads to certain challenges. One of the major issue in cloud computing is related to task scheduling in which the goal of service provider is to use the available resources in an efficient manner. A number of meta-heuristic algorithms have been proposed to solve this problem. In this paper, cloud task scheduling policy based on Ant colony Optimization (ACO) and Honey Bee Optimization (HBO) algorithms are combined to improve resource utilization. Minimizing Makespan, Flowtime and reducing cost is the major goal of proposed algorithm.

Keywords— ACO, HBO, Hybridisation, Makespan, Flowtime, Task scheduling.

I. INTRODUCTION

Computing is a term which is heavily used in modern era, which has fascinated people of all tiers of an IT organization, naive users etc. Due to increase of dynamic requirements of Individuals or communities (IT or Non IT), they want to move to cloud to cater to dynamic needs of services as well as potential customers. Realization before choosing the type of service from providers is so extreme importance. Adoption of cloud services is increasing at a rapid rate. At present, more than hundreds of cloud services are available in market. In a nutshell, cloud is elastic and shared pool of resources which are dynamically provisioned based on a service level agreement presented through negotiation between consumers and service providers [1].

Manual assignment of tasks to available resources which is called task scheduling is a difficult task [2]. Hence, a good task scheduling algorithm should adapt its strategy to changing environment and types of tasks. Earlier, there were no concepts like virtualization' hypervisors, shared pooling [3]. Research and therefore, Particle Swarm optimization algorithms [4], such as Ant Colony Optimization and Honey Bee Optimization are suitable for cloud [5][6]. ACO algorithm is based on the behaviour of real ant colonies in nature who wander randomly searching for optimal path between food source and their colony and returning back laying down pheromone trails. While HBO mimics the actions of honey bee colonies to seek best solution to the optimization problem over large distances and multiple directions [7].

In this paper, hybrid of ACO and HBO algorithm is used to overcome disadvantages of each other and to find optimal resource allocation for tasks minimizing overall makespan and flowtime and estimating cost factor [8]. This scheduling strategy was implemented using cloudsim toolkit [9]. Experimental results of proposed algorithm compared with ACO algorithm showed that proposed algorithm outperformed ACO algorithm. Paper is organized as follows: Section I contains the introduction of task scheduling optimization, Section II includes information about task scheduling, Section III contains related work. Meta-heuristic algorithms are described in Section IV. Section V covers the proposed solution and hybrid algorithm. Execution and result evaluation are seen in section VI. In the end, Section VII concludes this paper along with future scope.

II. TASK SCHEDULING

Cloud computing is shared pool of resources provided to user on pay-as-per-use basis without burden of managing the infrastructure. Gaining maximum profit and efficient use of resources is the major objective of service provider [10]. Thus task scheduling becomes a challenging matter in cloud computing. Scheduling refers to a set of policies which define the order of execution of processes [11]. The scheduling in CPU is done to keep it as busy as possible. From all of the available resources of a computer system that needs scheduling before use, the CPU is one of the most critical of them. Multiprogramming is one of the basic and important scheduling technique. In parallel computing, multiple processors have to be scheduled, and it needs to manage the resources for all the processors.

The scheduler distributes jobs across various clusters to allocate the jobs that cannot be assigned to a single cluster and this allocation strategy is referred to as co-allocation(Lee, Leu, & Chen, 2015) [12]. It can reduce waiting time of jobs in queue and hence maximizing the job throughput. So, this helps to improve system utilization and reduce average waiting time in a queue as the jobs that would otherwise wait in the queue for resources can start its execution earlier. However, when the co-allocated jobs compete for inter-cluster network bandwidth then mapping of jobs across the cluster boundaries may results in poor overall performance. Moreover, the heterogeneity of processing and communication resources increases the complexity of the scheduling problem. In recent decade, multi-cluster grid has become the standard execution condition for some large scale (logical) applications with differing attributes [13].

III. RELATED WORK

Hsu and Malyshkin et al [14] the techniques associated with the parallel computing are considered in this case. The parallel system involves large number of job scheduling algorithm. The advantages and disadvantages of each technique are considered in this case. The optimality principal is applied in order to judge best algorithm out of the available resources.

Wang, Xing, Li, Member & Yang et al [15] **a** modified ant colony optimization approach is presented in this paper for the network coding resource minimization problem. It is featured with several attractive mechanisms specially devised for solving the concerned problem**.** During the solution construction phase, the optimization of ant colony can well exploit the local and global information of routing related problems. The results of simulation are represented with the five integrated extended mechanisms. The proposed algorithm outperforms the best solutions obtained and the computational time and provides a number of existing algorithms.

Karaboga, Akay et al [16] Artificial Bee Colony algorithm is one of the most popular swarm intelligence based optimization techniques inspired by the foraging behaviour of real honey bees. Like many, population based evolutionary computation approaches; Artificial Bee Colony algorithm is basically suitable for distributed architectures. Furthermore, to understand the biggest issue for parallel implementations is to decide which food source should be chosen to distribute between sub-colonies and communication. In this study, a new schema for increasing the quality of the distributed source by combining best solutions is proposed. The proposed model of this paper was adopted for ring migration topology. The effectiveness is compared with the ring based topology in which best food sources and the original sequential counterpart in each sub populations are distributed. Comparative results show that the proposed model increased the quality of solutions and early convergence speed while protecting the speedup gain.

Khanmirzaei, Teshnehlab, & Sharifi et al. [17] proposes a Mamdani recurrent neurofuzzy system model (MRNFS), utilizing adjusted Honey Bee Optimization (HBO). In the essential variant of HBO, the calculation plays out a sort of neighbourhood pursuit joined with arbitrary hunt; thus it has the ability of accomplishing worldwide ideal. To enhance the neighbourhood look capacity of HBO and help the calculation to bounce out from the nearby ideal, an adjustment is performed by applying three sorts of hybrids to the first class people. To confirm the execution of the proposed technique, this strategy is connected to some distinguishing proof and expectation benchmarks and its execution contrasted and the essential HBO, Gradient plummet (GD), Differential Evolution (DE) and Particle swarm optimization (PSO), in preparing the MRNFS model.

Fatemeh Rastkhadiv and Kamran Zamanifar et al. [18] *presents a paper that discusses a new scheduler based on load balancing technique using Artificial Bee Colony Optimization. The scheduler is designed to manage the load across the virtual machines and minimizing makespan while maximizing overall throughput. The experimental results implemented using Cloudsim to optimize task scheduling and load balancing is compared with both ACO and FCFS.

IV. META-HEURISTIC ALGORITHMS

The Meta-heuristic algorithms generally make up an important solution of global optimization algorithms. Totally heuristic means to find and discover by error and trial. "Meta-heuristic" can be applied to strategies with a higher level that have modified and guided heuristic methods in a way that can achieve solutions and innovations beyond what is normally accessible in local optimum search. While dealing with optimization problems, there are two types of local and global optima. The local optima are the best solution found in an area of solution spaces but not necessarily the best for the whole problem space. In contrast, the global optima are the best solution to the whole problem space. In most real-life problems finding the global optima is extremely difficult and therefore satisfactory and goodenough solutions are often accepted. To achieve these targets Meta-heuristic algorithms are employed.

IV.I Ant colony Optimization

In the natural world, ants of some species (initially) wander randomly to search food, and while returning back to their colony they lay down pheromone trails. If other ants find such a path, then rather than wandering randomly they are likely to follow the pheromone trail, returning and reinforcing it if they eventually find food. With time, when the pheromone trail starts to evaporate, it reduces its attractive strength. If the ants take more time to travel down the path and back again, the more time the pheromones have to evaporate. Shortest the path more frequently it will be marked as compared to others. This results in increase in density of pheromone on shorter paths than longer ones. By utilizing Pheromone dissipation the meeting to a locally ideal arrangement can be effectively accomplished. In case there was no evaporation at all, the paths chosen by the first ants would become excessively attractive to the following ones. In that case, the exploration of the solution space would be constrained. The influence of pheromone evaporation in real ant systems is unclear, but it is very important in artificial systems.

The overall result is when one ant finds a path to get food then other ants also follow the same food source and positive feedback eventually leads to all the ants following a single path. The ant colony algorithm gets idea of this behaviour from ants that walking around the graph representing the problem to solve

IV.II Honey Bee Optimization

One of the advantage of honey bee optimization over ant colony is, in order to harvest nectar or pollen from multiple food sources (flower patches) a colony of honey bees can extend itself over long distances (over 14 km) and in multiple directions simultaneously. The bee colony constantly searches the surrounding and looking for new flower patches. These scout bees move around the hive randomly, and evaluate the food sources encountered in their path. When the bee starts hive, the scouts deposit the food harvested. The "dance floor", is the area where individuals found a highly profitable food source in the hive and perform a ritual known as the waggle dance. Furthermore, during waggle dance a scout bee communicates the location of its discovery to idle onlookers, which join in the utilization of the flower patch. The dance performed by the scout bee has some characteristics. The length of the dance is proportional to the scout's rating of the food source, the best rated flower patches are recruited with more foragers to harvest. After completing their dancing, the scout returns to the environment so that they can discover more food and collect them. As long as they are evaluated as profitable, when they return to the hive, rich food sources will be advertised by the scouts. Onlooker bees may perform dance as well, to increase the number of recruited bees for highly rewarding flower patches. In autocatalytic process, the colony of bees has an ability to quickly switch the focus of the foraging effort on the most profitable flower patches.

V. PROPOSED SOLUTION

The existing system utilizes the application of ACO or honey bee individually but collaborative approach is not followed. The ant colony optimization is optimal algorithm for determining resource and energy efficiency but problem exists as distance increases. The honey bee algorithm is optimal in terms of more distance but some optimal path may be skipped. So ant colonies and honey bee algorithm must be merged together to achieve the result. The proposed technique combines the features of ant colony and honey bee algorithm in order to achieve optimality. The ant colony algorithm can be utilized if distance increases and features of honey bee can be utilized if distance increases

HYBRID ALGORITHM

1. Initialization
Pheromone trial, random initial solution, number of forager bees, maximum iterations.
2.Iterative loops
   Colony of ants determine the starting loops,
   schedule is contructed for each ant.
3.Repeat
   Apply transition rule for next processing job to execute,
   complete schedule is contructed for every ant.
4. Till complete schedule is constructed do following:
   Honey bee search is processed for searching.
   Pheromone trial is updated.
   Global updation rule is processed.
5.Termination.
   If searching and pheromone updation is completed,
   Then terminate the process
   Else goto step 4

Search Function

1. Iterative loop
   Current best solution is identified.
   Prioritize the population
2. Evaluation
   Evaluate and sort the current solution based on fitness value.
   Search for new nodes and select sites for neighbourhood
   Send bees for chosen sites
3. Condition
   If (
   Value is replaced by new solution.)
   Else
   Repeat step 2

4. Statement
   Select fittest bee
   Current best solution is identified.
   Current best solution is carried to next generation.

FLOW CHART

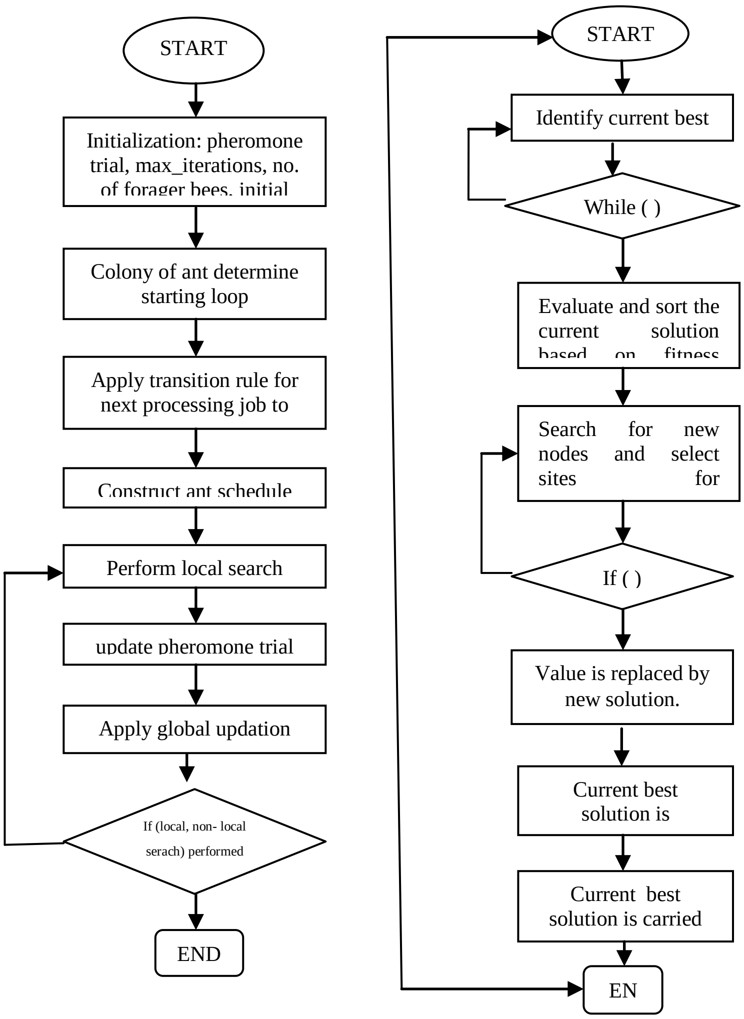

Figure 1: Proposed work

VI. EXECUTION AND RESULT EVALUATION.

The main tools of experiment are Netbeans and CloudSim. It is clear from the configuration file that we need to configure the number, ID, processing power (MIPS) for virtual machines and in case of cloudlets we need to configure ID, length in cloud simulation system. Various parameters that directly or indirectly affect the computation of algorithm are {α, β, ρ, t_{max}, number of ants}. In the experiment values of various parameters are changed and result is computed for different numbers of tasks. In proposed algorithm we also computed cost parameter to estimate the budget for the computation of tasks. Makespan and flowtime value for various parameters is evaluated and performance of ACO and hybrid algorithm is compared.

1. Makespan is the amount of time required to complete the whole job. It should be minimized in order to achieve the energy efficiency such that makespan = Submission time – Response time.

$$\text{Makespan} = \max_{n \in tasks} \{\text{Finish\_time}_i\} \quad [19]$$

2. Flowtime is the sum of response times of all jobs which must be minimized.

$$\text{Flowtime} = \Sigma_{n \in tasks} (\text{Finish\_time}_i) \quad [19]$$

3. Cost is the amount the consumer has to pay to service provider for consuming services.

$$\text{Cost} = \Sigma_{n \in resources} \{\text{Cost}_n * \text{Time}_n\} \quad [19]$$

Where $Cost_n$ is cost of utilising resource per unit time and $Time_n$ depicts the time for the resource n is used.

Parameter Setting of Cloudsim

The experiments are performed with 10 datacenters having 50 virtual machines for 100-1000 tasks with each task having length from 1000-4000 million instructions as shown in table below.

Table 1: Parameter setting of cloudsim.

Entity type	Parameter	Value
Task(Cloudlets)	Length of task	1000-4000
	Total number of tasks	100-1000
Virtual Machine	Total number of VMs	50
	MIPS	500-2000
	Cloudlet scheduler	Time_Shared
	Number of Processing elements requirement	1-2
	Bandwidth	500
	VM memory	1024
Datacenter	Number of datacenter	10
	Number of host	2
	VM Scheduler	Time_Shared

Result Analysis

In result graphical representation shows the strength and weaknesses of experiment. We can only change one value at a time in each experiment, where first we can fix the values of β and ρ and varies the parameter α (alpha) ϵ {0.5, 1.0, 2.0, 2.50}. Similarly we can vary the value of β ϵ {0.5, 1.0, 2.0, 2.50} and ρ ϵ {0.1, 0.2, 0.3, 0.4, 0.5} while keeping all the other parameters fixed. The ACO and hybrid algorithm's performance is computed for different values of parameter. Makespan and flowtime value is obtained for both ACO and hybrid algorithm for 200 and 500 tasks each. Another parameter that is cost parameter is also computed for hybrid algorithm to estimate the cost incurred to run the jobs on virtual machines. Comparison between existing and proposed approach is performed and shown in graphical manner.

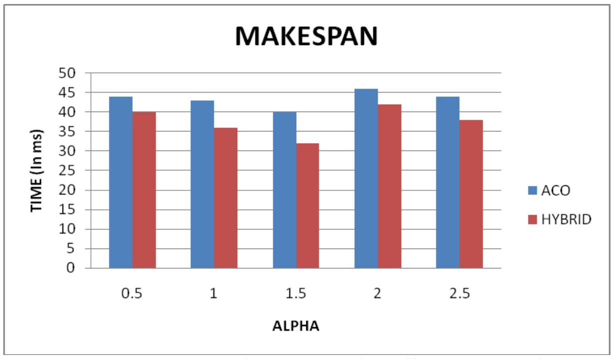

Figure 2: Makespan for 200 tasks for different values of alpha.

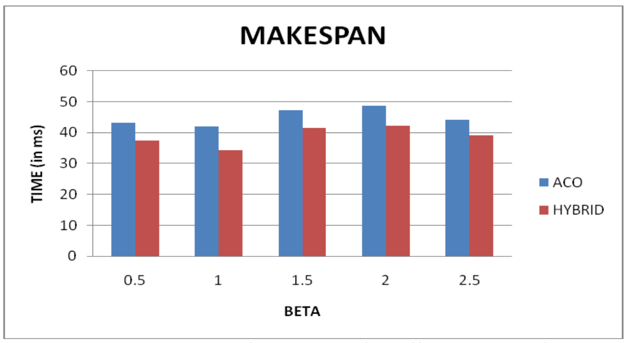

Figure 3: Makespan for 200 tasks for different values of beta.

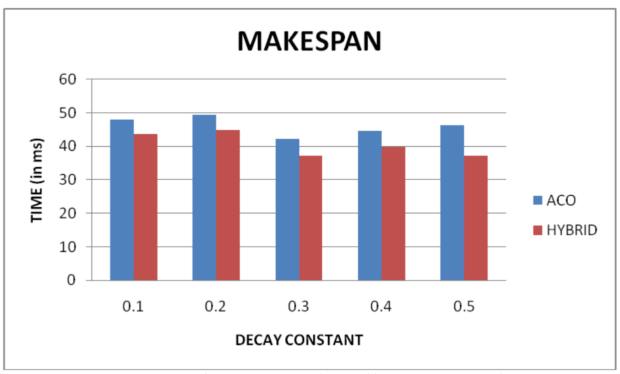

Figure 4: Makespan for 200 tasks for different values of decay constant.

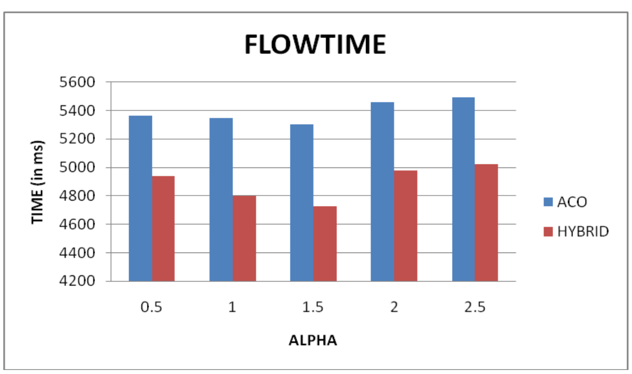

Figure 5: Flowtime for 200 tasks for different values of alpha

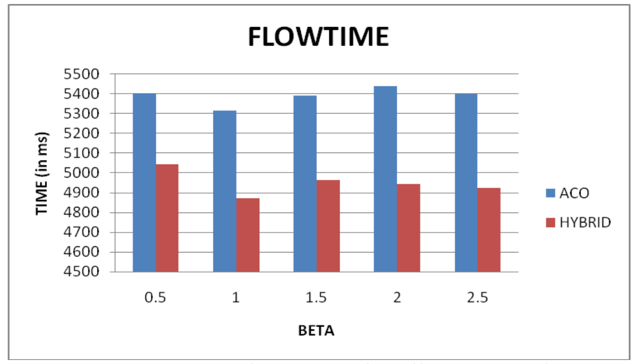

Figure 6: Flowtime for 200 tasks for different values of beta.

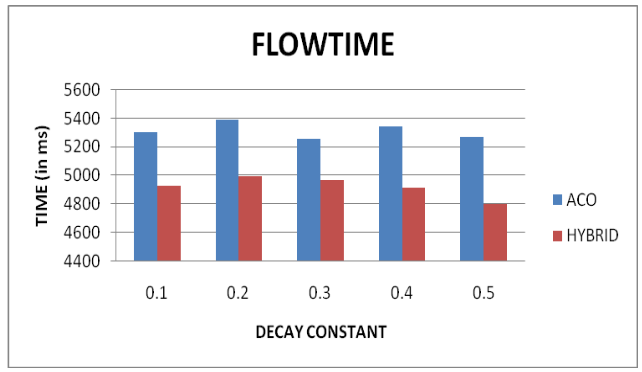

Figure 7: Flowtime for 200 tasks for different values of decay constant.

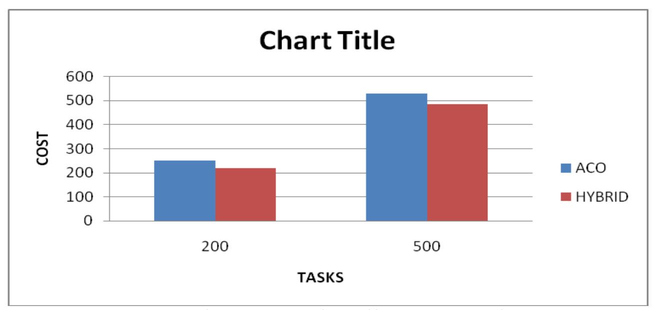

Figure 8: Flowtime for 200 tasks for different values of decay constant.

Table 2: Selected parameters of proposed algorithm.

Parameter	Alpha	Beta	Decay constant
Value	1.5	1	0.5

VII. CONCLUSION AND FUTURE SCOPE

The design of task scheduling algorithm should be such that it supports cloud infrastructure and a large number of applications. The basic mechanism for working of ACO is random search mechanism which is used for allocation of incoming jobs. ACO does little to improve the resource utilization. The combined implementation of ACO and HBO algorithms helped in receiving more optimal results. The main work of the proposed algorithm is to minimize Makespan, Flowtime and efficient resource utilization. Finally we judge that the hybrid work outperformed as compared to ACO algorithm proving worth of study.

In future ant colony and honey bee algorithm can be merged for further reducing makespan and flowtime assosiated with jobs. Furthermore we can vary the number of virtual macines to get better results. The algorithm can be enhanced to solve load balancing problem while task scheduling. Moreover this algorithm can be improved for multi-user environment.

REFERENCES

[1] M.Mezmaz, N.Melab, "A Parallel bi-objective hybrid etaheuristic for energy-aware scheduling for cloud computing systems "J.ParallelDistrib. Comput. 71(2011)1497-1508.

[2] Kaur, H., and Gautam, E.V. (2014). International Journal of Computer Sciences A Survey of Various Cloud Simulators. 3–6.

[3] J. E. Smith and R. Nair. *Virtual Machines: Versatile platforms for systems and processes.* Morgan Kauffmann, 2005.

[4] Chan, T.S. Felix, M. Kumar, Tiwari,"Swarm Intelligence: Focus on Ant and Particle Swarm Optimization", *Vienna, Austria: I-Tech Education and Publishing*, 2007.

[5] Xiangqian Song, Lin Gao, Jieping Wang, "Job scheduling based on ant colony optimization in cloud computing", In Proceedings of 2011 International Conference on Computer Science and Service System, pp.3309 -3312, 2011.

[6] L.D. Dhinesh Babu, P. Venkata Krishna,"Honey bee behavior inspired load balancing of tasks in cloud computing environments" ,*Applied Soft Computing*, Vol. 13, 2013; pp. 2292-2303.

[7] D. Karaboga, B. Basturk,"Artificial Bee Colony (ABC) Optimization Algorithm for Solving Constrained Optimization Problems",*Springer-Verlag Berlin Heidelberg,* 2007,pp. 789–798.

[8] RubingDuan,RaduProdan, "Performance and Cost Optimization for Multiple Large-scale Grid Workflow Applications"2007 ACM 978-1- 59593-764-3/07/0011.

[9] Calheiros, R.N., Ranjan, R., De Rose, C.A.F., and Buyya, R. (2009). CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services.

[10] Vijindra and Sudhir Shenai. A, "Survey of Scheduling Issues in Cloud Computing", 2012, Elsevier Ltd.

[11] Isam Azawi Mohialdeen, "A Comparative Study of Scheduling Algorithms in Cloud Computing Environment" 2013 Science Publications.

[12] Lee, M.C., Leu, F.Y., and Chen, Y.P. (2015). ReMBF: A Reliable Multicast Brute-Force Co-allocation Scheme for Multi-user Data Grids. In Proceedings - International Computer Software and Applications Conference, (IEEE Computer Society), pp. 774–783.

[13] Nandagopal, M. and Uthariaraj, R. V.," Hierarchical status information exchange scheduling and load balancing for computational grid environments," IJCSNS International Journal of Computer Science and Network Security, 10(2):177-185, 2010.

[14] C.H. Hsu and V. Malyshkin, "*Methods and tools of parallel programming multicomputers-2010*", Springer Publication, **Russia**, pp. 10.1007/978-3-642-14822-4

[15] Wang, Z., Xing, H., Li, T., Yang, Y., Qu, R., and Pan, Y. (2016). A Modified Ant Colony Optimization Algorithm for Network Coding Resource Minimization. IEEE Transactions on Evolutionary Computation *20*, 325–342.

[16] D. Karaboga, B. Akay, "A survey: algorithms simulating bee swarm intelligence", *Artif Intell Rev*, vol. 31, no. 1, pp. 68-85, 2009.

[17] Khanmirzaei, Z., Teshnehlab, M., and Sharifi, A. (2010). Modified honey bee optimization for recurrent neuro-fuzzy system model. In 2010 The 2nd International Conference on Computer and Automation Engineering, ICCAE 2010, pp. 780–785.

[18] Rastkhadiv, F., and Zamanifar, K. (2016). Task Scheduling Based On Load Balancing Using Artificial Bee Colony In Cloud Computing Environment. International Journal of Advanced Biotechnology and Research *7*, 1058–1069.

[19] M. Kalra, S. Singh, " *A review of metaheuristic scheduling technique in cloud computing*", Cairo University, Egyptian Informatics Journal, Vol. 16, issue 3, pp 275-295, 2015.

Authors Profile.

Tanu Chopra pursed Bachelor of Engineering and Technology from DAV Institute of Engineering and technology, Jalandhar in 2014 and currently she is pursuing Master of Computer science in Big data from IKG Punjab Technical University, Kapurthala, INDIA.

*Rajdeep Singh p*ursed Bachelor of Technology from MIMIT, Malout, Punjab and Master of Technology from PEC, Chandigarh, India. He has worked in industry for 4 years and has 8 years of experience in teaching. Currently he is working as Assistant Professor at IKGPTU, Kapurthala in the Department of Computer Science and Engineering. He has more than 10 publications in research journal and conferences.

