RESEARCH ARTICLE

Task scheduling in Cloud Using Hybrid Cuckoo Algorithm

Sumandeep Aujla
Research Scholar, CGC Landran, Mohali, Punjab
Sumandeep.aujla00@gmail.com

Amandeep Ummat
Assistant Professor, CGC Landran, Mohali, Punjab
cecm.cse.amanummat@gmail.com

Abstract – Cloud computing is used for delivering and managing the services over the internet. Cloud computing provides data access and storage devices without the knowledge of the physical location of the end user. Cloud computing has three types of services: software as a service (SaaS), platform as a service (PaaS) and infrastructure as a service (IaaS). Task scheduling is the main issue in cloud computing and it is very important part of cloud computing. Various types of algorithms are used for scheduling for example FIFO, Genetic algorithm, round robin algorithm etc. This paper is focused on the optimization solution for the task scheduling using hybrid cuckoo algorithm. This algorithm combines the results of genetic algorithm and cuckoo search algorithm.

Index Terms – Cloud computing, Task scheduling, Genetic algorithm, Cuckoo algorithm, Hybrid cuckoo algorithm.

1. INTRODUCTION

Cloud computing is basically metaphor for the internet [1]. In Cloud Computing the resources are delivered over the network as a service [2]. Figure 1 shows the cloud environment where various cloud providers like Google, Amazon, Yahoo etc enables users access to resources (Like storage, applications, information etc). Using cloud we can access the information from anywhere and at any time. The cloud removes the need of same physical location. The cloud is very helpful for the small companies that can't afford the storage devices to store their information. These companies store their information on the cloud and removing the cost of purchasing storage devices [3]. Internet connection is the basic need to access the cloud. In cloud computing users access the various services from cloud provider. Information is exchanged between the cloud provider and user whenever user logs in [4]. In cloud computing Clients can store their data on the virtual platform which is called cloud instead of their own PCs, and client can put their applications on the cloud and use the servers within the cloud to do processing and data manipulations etc. cloud computing is very popular in these days because of its high scalability, anytime and anywhere access using the web browser. Cloud computing has three architectures:

Cloud software as a service (SaaS): In this software is made available to the user as service. Cloud applications are generally accessible from various devices like mobile, tablet, laptop, PC, workstations, servers etc. The user has no control over the underlying platform and infrastructure. Examples are Dropbox, Gmail, Gtalk etc.

Cloud platform as a service (PaaS): In this software is made available to the user as service. Programming languages and tools are provided by service provider to develop and deployment services. A user has no control over underlying infrastructure but has control over the deployed applications. Examples are Windows Azure, Google App Engine.

Cloud infrastructure as a service (IaaS): In this Software is made available to the user as a service. A user can demand computing infrastructure, storage infrastructure and network infrastructure etc from service provider. The actual owner of the infrastructure is not user, but has control over operating systems, storage, and deployed applications etc. Example: Amazon, Rack space cloud.
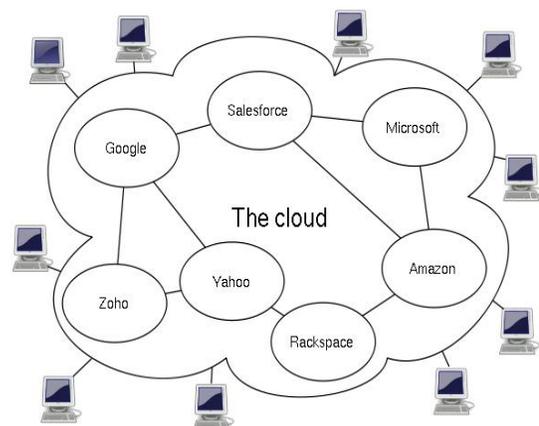

Figure 1 Cloud Computing Environment

RESEARCH ARTICLE

1.1 Task Scheduling

Task scheduling is very important in cloud computing in order to achieve the optimal solution. Task scheduling is the assignment of start and end time of the different tasks that are subject to certain constraints. The constraints can be resource constraint or time constraint. Task scheduling is very important part of cloud computing. Task scheduling helps to maximize the resource utilization and minimizing the execution time by distributing the load on different processors. It has two types: static scheduling and dynamic scheduling. Several algorithms are proposed for scheduling mechanism. Scheduling mechanism is very important to improve the resource utilization.
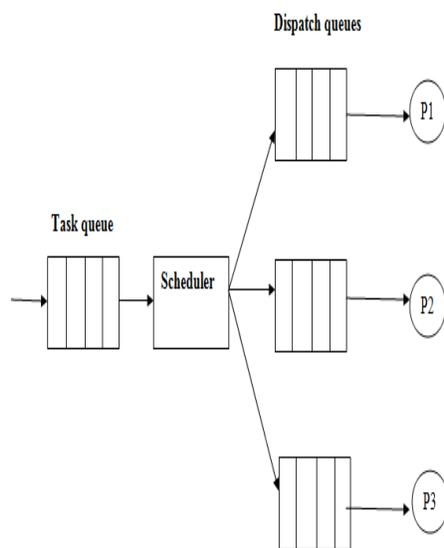

Figure 2 Task scheduling

The basic process of task scheduling is shown in Figure 2. In cloud computing there is a queue of different tasks each task as different priority. Scheduler checks the priority of each task and allocates the tasks to different processors according to their priority.

2. RELATED WORK

Currently many scheduling algorithms are used for the task scheduling problem in cloud computing. Lipsa Tripathy [5] proposed a scheduling mechanism to schedule the various jobs in task scheduling. The jobs are scheduled in cloud by using this mechanism to overcome the shortcomings of earlier protocols. It improves the resource utilization server performance and throughput. Eleonora Maria Mocanu [6] uses the genetic algorithm to find the optimal solution of task scheduling problem. The work was done in the hadoop platform. The Genetic Scheduler is derived from the Job queue scheduler class. The goal of this paper is to improve

Hadoop's functionality by implementing a scheduler based on a genetic algorithm, which solves the stated problem. Nima Jafari Navimipour [7] proposed a new evolutionary algorithm which named cuckoo search algorithm (CSA) to schedule the tasks in Cloud computing. Every cuckoo gives one egg at a time, and dumps it in a nest that is randomly chosen, and then the host can discover an alien egg by a Pa [0, 1] probability. The speed and coverage of the algorithm become very high if the value of Pa is low. Sung Ho Jang, Tae Young Kim [8] proposed a task scheduling model based on the genetic algorithm, which focus on improving the profit for cloud provider. The GA scheduling function creates a set of task schedules, and evaluates the population by using the fitness function that considers the user satisfaction and virtual machine availability. The function iterates reproduce the populations to output the best task schedule. The task scheduling using genetic algorithm shows the effectiveness and efficiency in results as compared to another task scheduling models like ABC based task scheduling model and round robin task scheduling model. Xiaonian Wu [9] proposed an algorithm that is based on quality of service driven. The algorithm finds out the priority of different tasks on the basis of different attributes of tasks and applies the sorting on tasks onto a service which can further complete the task. Mrs.S.Selvarani [10] has proposed a improved cost based algorithm. This algorithm employs the efficient allocation to the resources that are available in cloud. It improves the computation/communication ratio. The results produce using improved cost algorithm is better than ABC algorithm. Jung, Lim [11] presented the scheduling scheme for workload that reduces the waiting time of tasks. Naghibzadeh [12] proposed QoS-based workflow scheduling which minimize the cost of workflow. Choudhary and Peddoju [13] presented an algorithm for scheduling which addresses the major challenges of task scheduling in the Cloud environment. Shaminder kaur[14] have proposed a new algorithm Modified Genetic Algorithm for task scheduling. This algorithm modifies the initial population with SCFP (Shortest Cloudlet to Fastest Processor), LCFP (longest Cloudlet to Fastest Processor)

3. PORPOSED MODELLING

Nowadays scheduling algorithm is required that can reduce the executing time, consumes less energy and improves the resource utilization of the resources. The work presented in this paper will focus on the optimization based scheduling. Hybrid cuckoo algorithm is used in this paper for the task scheduling. Hybrid cuckoo algorithm is the combination of genetic algorithm and cuckoo algorithm. And the results that are produced using hybrid cuckoo algorithm are better than that of Genetic algorithm and cuckoo algorithm.

RESEARCH ARTICLE

The simulation tool which is used in this work is visual studio 2010. The whole web based work is done in visual studio. To create the cloud environment Microsoft window azure is used.

This paper focus on the following parameters:

- Time
- Resource utilization
- Energy consumption

Hybrid cuckoo algorithm is used in this to achieve the following objectives:

- Proper execution of the tasks in cloud.
- Proper utilization of the resources.
- Less chances of the failure of the system.
- Minimization of the load on different processors.
- To reduce the waiting time.
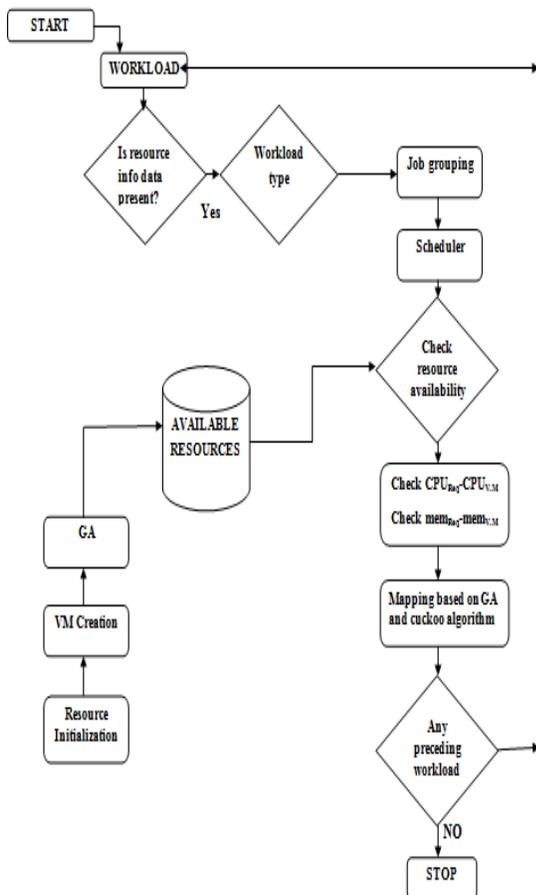
3.1. Basic design of the system

Figure 3 Basic design of the system

The basic design of the system is shown in Figure 3. When jobs are allocated, it checks whether the information about the availability of resources present or not. If the information is present then the system checks the workload type of the jobs, whether it is CPU intensive or memory intensive and what their requirements are. Then the jobs are grouped according to their workload type. Scheduler checks at which time which job has to be executed with which resources and schedule the jobs according to their priority. On the other end resources initialization with its general parameters means temperature value, input and virtual machine creation according to temperature means the capacity of the machine. By using genetic algorithm, cuckoo and hybrid cuckoo algorithm sorting is been done and saved in resource database. Scheduler maps these available resources to jobs according to their requirements. Then checks any pending workload is present there for execution or not if the workload is present then executes it and stops otherwise again the whole loop is started.

3.2 Hybrid Cuckoo Algorithm

Step 1: Initialization.

Step 2: Find out the number of task T that need to be scheduled.

Step 3: Schedule the task T using a Hybrid cuckoo algorithm as following:

1) Firstly tasks are scheduled using the genetic algorithm, and the output result is produced using this algorithm.

2) Then the output which is produced using GA is taken as an input for the cuckoo algorithm. And tasks are scheduled using cuckoo search algorithm.

3) In the final step cuckoo search algorithm is again applied to the results that are produced in 2). And final output result is given.

Step 4: Find out the execution timing, resource utilization.

Step 5: Termination check – When the entire task T has been assigned to the scheduler, the algorithm terminates. Else go to step 2 for scheduling the tasks. In the former steps, Step 3 is the main process of the algorithm.

Hybrid cuckoo algorithm combines the advantages of genetic algorithm and cuckoo search algorithm. Cuckoo search is an optimization algorithm (finding a value x such that f(x) is as small (or as large) as possible) based on cuckoo species which laying their eggs in the nests of other host birds (of other species) [15]. A GA is a stochastic technique based on the principles of natural evolution and genetics. GA combines the exploitation of past results with the exploration of new areas

of the search space. In order to overcome the disadvantages of both cuckoo algorithm and GA the new algorithm hybrid cuckoo algorithm is implemented. The steps that are used in hybrid cuckoo algorithm are shown above.

4. RESULTS AND DISCUSSIONS

In this section, results are presented to evaluate the execution time and resource utilization of the various jobs using C# in visual studio. To convert the web based work to local cloud window azure is used which is platform as a service generated by Microsoft company. Figure 4 shows the snap short of the work that is done in visual studio. We can sort the resources using GA, cuckoo algorithm and hybrid cuckoo algorithm and compare the results of all these algorithms. View machine shows the capacity of the machine. Figure 5 shows the time graph for all executed jobs. Figure 6 shows the energy consumption of the jobs. Figure 7 and 8 shows the resource utilization for all jobs.
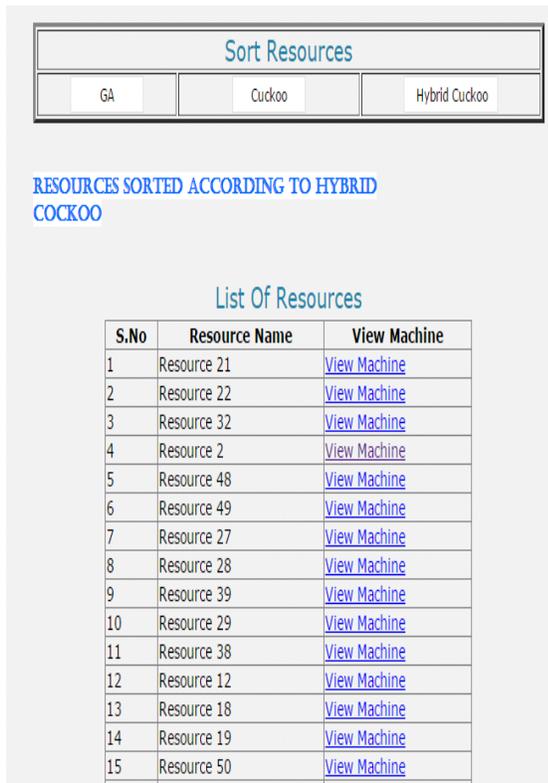

Figure 4 Sorting of resources

4.1 Formulas Used

The Restrigin Function is used to test the performance in this analysis. It is defined as following:

$$f(x) = A_n + \sum_{i=1}^{n}[x_i^2 - A\cos(2\pi x_i)]$$

Where A = 10 and $x_i \in [-5.12, 5.12]$. It has a global minimum at x = 0 where f(x) = 0.

$$Time = \frac{No.\, of\ instructions\ (job\ (i))}{MIPS\ rating\ (Machine\ (i))}$$

$$Energy = \left(\frac{Total\ temperature}{Avg\ Temperature}\right) * \frac{1}{Total\ no.\, of\ resources\ used}$$

Because here,

$$Energy \propto Temperature$$

And

$$Energy \propto \frac{1}{Resources}$$

4.2 Analysis

Following parameters are used for the analysis and after the analysis of these parameters we can say that proposed work is better than the previous modeling.

- Time
- Energy consumption
- Resource Utilization

(1) Time – As shown in the Figure 5 execution time of the jobs in hybrid cuckoo is less as compared to First in first out (FIFO), Genetic algorithm and Cuckoo algorithm so we can say that hybrid cuckoo algorithm minimizes the execution time of the jobs and gives the optimized result. Values are shown in Table 1 when we execute 50 jobs hybrid cuckoo takes only 48.72 seconds for the execution of the jobs but cuckoo algorithm, GA and FIFO takes 61.46, 101.8, 109.18 seconds respectively. So time taken by hybrid cuckoo algorithm is less compared to other algorithm.
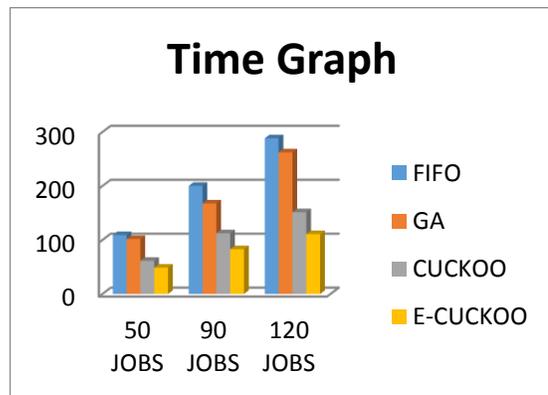

Figure 5 Time graph for all jobs

RESEARCH ARTICLE

	FIFO	GA	CUCKOO	E-CUCKOO
50 JOBS	109.18	101.8	61.46	48.72
90 JOBS	200.23	167.69	112.33	82.92
120 JOBS	288.45	262.28	151.37	110.85

Table 1 Modeling Table of time

(2) Energy consumption - Energy consumption in hybrid cuckoo algorithm is less as compared to GA and cuckoo algorithm. The values are given in Table 2 and it is clear from the table that jobs that are executed using hybrid cuckoo algorithm consume very less energy and improves the performance of the system. Values are shown graphically in Figure 6.
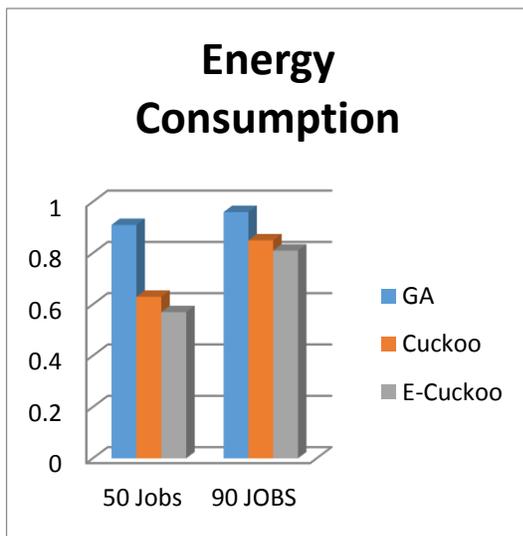

Figure 6 Energy for all jobs

Table 2 Modeling Table of energy

(3) Resource utilization - We consider 3 values [-1, 0, 1]. If the value is -1 then the resources are less utilize, if value is 0 the resources are properly utilize and if the value is 1 then resources are over utilize. It is shown in Table 3 and Table 4 that resources are properly utilized in hybrid cuckoo algorithm for 50 and 90 jobs respectively but in GA and cuckoo algorithm resources are less or over utilize. So we can say that hybrid cuckoo algorithm maximize the resource utilization as shown in Figure 7 and 8.
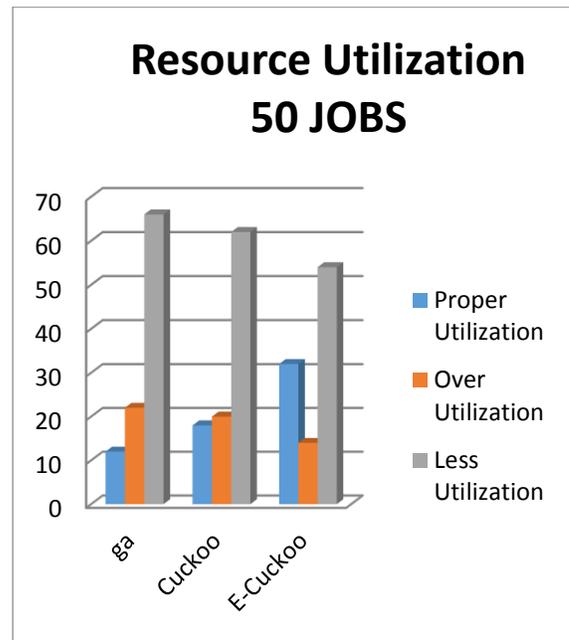

Figure 7 Resource Utilization for 50 jobs

50 JOBS			
	GA	Cuckoo	E-Cuckoo
Proper Utilization	12	18	32
Over Utilization	22	20	14
Less Utilization	66	62	54

Table 3 Modeling table of resources for 50 jobs

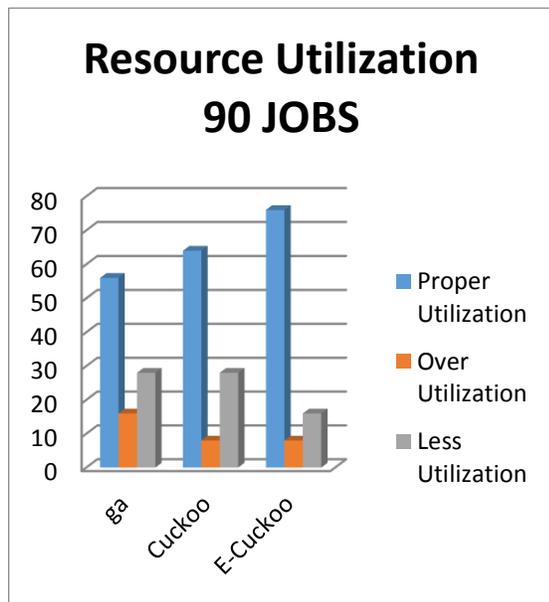|	50 Jobs	90 JOBS
GA	0.91	0.96
Cuckoo	0.63	0.85
E-Cuckoo	0.57	0.81

RESEARCH ARTICLE

Figure 8 Resource Utilization for 90 jobs

90 JOBS			
	GA	**Cuckoo**	**E-Cuckoo**
Proper Utilization	56	64	76
Over Utilization	16	8	8

Table 4 Modeling Table of resources for 90 jobs

5. CONCLUSION

Cloud computing is an emerging computing technology that uses the internet and central remote servers to maintain data and application. Task scheduling in cloud computing is very complex problem. Using the hybrid cuckoo algorithm to find the optimal solution in scheduling is very simple approach. This algorithm shows the better solution for scheduling in visual studio framework. This algorithm is better than previous modelling because it minimizes the execution time and maximizes the resource utilization and consumes less energy when we work with 50, 90 and 120 jobs than that of the GA, FIFO and cuckoo algorithm. The results that are produced using hybrid cuckoo algorithm are compared with the GA and cuckoo algorithm and we conclude that hybrid cuckoo algorithm is better than others.

In future work we can check the results for more than 120 jobs. And we can compare the hybrid cuckoo algorithm with another task scheduling algorithms.

REFERENCES

[1] K. Kc, K. Anyanwu "Scheduling Hadoop Jobs to meet deadlines" International Conference on Management of Data, 2010.

[2] J. Geelan "Twenty-One Experts Define Cloud Computing Virtualization" Electronic Magazine Article, available at the link http://virtualization.sys-con.com/node/612375, 2008.

[3] Alexa Huth and James Cebula "The Basics of Cloud Computing" © 2011 Carnegie Mellon University. Produced for US-CERT, a government organization.

[4] Aarti Singh, Manisha Malhotra "Security Concerns at various Levels of Cloud Computing Paradigm: A Review" Intenation al Journel of Computer Networks and Appplications Volumn 2, Issue 2, March – April 2015.

[5] Lipsa Tripathy, Rasmi Ranjan Patra "Scheduling in cloud computing" International Journal on Cloud Computing: Services and Architecture (IJCCSA) ,Vol. 4, No. 5, October 2014.

[6] Eleonora Maria Mocanu, Mihai Florea, Mugurel Ionuţ Andreica, Nicolae Ţăpuş "Cloud Computing – Task Scheduling based on Genetic Algorithms" ©2012 IEEE

[7] Nima Jafari Navimipour and Farnaz Sharifi Milani,"Task Scheduling in the Cloud Computing Based on the Cuckoo Search Algorithm" International Journal of Modeling and Optimization, Vol. 5, No. 1, February 2015.

[8] Sung Ho Jang, Tae Young Kim, Jae Kwon Kim and Jong Sik Lee "The Study of Genetic Algorithm-based Task Scheduling for Cloud Computing" International Journal of Control and Automation Vol. 5, No. 4, December, 2012.

[9] Xiaonian Wu, Mengqing Deng , Runlian Zhang , Bing Zeng, Shengyuan Zhou "A task Scheduling algorithm based on QOS- driven" International conference on Information Technology and quantitative Management(ITQM) pp. 1162-1169 , ELSEVIER(2013).

[10] Mrs.S.Selvarani1, Dr.G.Sudha Sadhasivam "Improved cost based algorithm for task scheduling in cloud computing" ©2010 IEEE.

[11] D. Jung et al., "A workflow scheduling technique for task distribution in spot instance-based cloud," Ubiquitous Information Technologies and Applications, Springer Berlin Heidelberg. pp. 409-416, 2014

[12] S. Abrishami and M. Naghibzadeh, "Deadline-constrained workflow scheduling in software as a service cloud," Scientia Iranica, vol. 19, Issue 3, June 2012.

[13] M Choudhary and S. K. Peddoju, "A dynamic optimization algorithm for task scheduling in cloud environment," International Journal of Engineering Research and Applications (IJERA), Vol. 2, Issue 3, pp. 2564-2568, 2012.

[14] Shaminder kaur, Amandeep Verma "An efficient Appproach to Genetic Algorithm for Task Scheduling in Cloud Computibg Environment" International Journal of Information Technology & Computer Science, Sep2012, Vol. 4 Issue 10.

[15] S. Tayal. Tasks Scheduling optimization for the Cloud Computing Systems, IJAEST, vol. 5, issue no. 2. S.a.

[16] J. Dean, S. Ghemawat. MapReduce: simplified data processing on large clusters. 2004.

[17] X.-S. Yang; S. Deb (December 2009). Cuckoo search via Lévy flights. World Congress on Nature & Biologically Inspired Computing (NaBIC 2009). IEEE Publications.

[18] K. Kaur, A. Chhabra, G. Singh. "Heuristics Based Genetic Algorithm for Scheduling Static Tasks in Homogeneous Parallel System" International Journal of Computer Science and Security, vol. 4, issue 2, 2010.

[19] Savita.P, Geetha Ready "A Review Work on Task Scheduling in Cloud Computing using Genetic Algorithm" International Journal of Scientific and Technology Research, Vol 2 Issue 8, August 2013.

RESEARCH ARTICLE

[20] X.-S. Yang and S. Deb, "Engineering optimisation by cuckoo search," International Journal of Mathematical Modelling and Numerical Optimisation, vol. 1, Issue 4, pp. 330-343, 2010.

Authors

Sumandeep Aujla is pursuing her Mtech degree in computer science and engineering from CGC Landran, Mohali, Punjab. She received the Btech degree from BBSBEC, Fathegarh Sahib, Punjab in 2013. Her research area lies in the field of cloud computing. Her current research work is based on the task scheduling in cloud computing.

Amandeep Ummat works as assistant professor in CGC, landran, Mohali, Punjab. He gives the guidance in this paper. He has done his master degree in the field of computer science and engineering. He done his research work in the field of cloud computing.