

Enhanced Genetic Algorithm based Task Scheduling in Cloud Computing

Rajveer Kaur

Research Fellow, Department of Computer Science and Engineering, SGGSWU, Fatehgarh Sahib, India, Punjab (140406)

Supriya Kinger

Asst.Professor, Department of Computer Science and Engineering, SGGSWU, Fatehgarh Sahib, India, Punjab (140406)

ABSTRACT

Cloud computing is basically internet based computing while software, information and shared resources are provided to devices and computers on demand, like electricity grid. With the fusion of network technology and traditional computing technology such as distributed computing parallel computing, grid computing a cloud computing product is formed. Task scheduling is the major concern in the field of cloud computing. As the use of cloud computing increases, the burden on the cloud network also increases. So, it's the duty of the scheduler to make cloud efficient to solve client's tasks. This work focuses on the same to achieve the objective of optimized task scheduling where improved genetic algorithm is proposed. Genetic algorithm is artificial intelligent based soft computing technique to optimize the process. Here in this work, genetic algorithm is enhanced using new fitness function based on mean and grand mean values. This optimization can be implemented on both ends, for job scheduling and resource scheduling. This will schedule the whole process and optimize as much as possible. The results analysis also proves the cloud system's increased efficiency for task scheduling.

Keywords

Cloud Computing, Task Scheduling, Genetic Algorithm, Enhanced Genetic Algorithm.

1. INTRODUCTION

Cloud computing is growing fast and is inexpensive [1]. Cloud computing provides IT services by taking advantage of the web applications, networks and the computing systems. For high resource utilization and to fulfill the user's requirements scheduling must be needs to allocate the resources. This is called as "the cloud task scheduling problem". Cloud computing service is focused to provide high quality and low-cost information services by pay-per-use model in which guarantees are offered by the cloud service providers [2]. Service provider provides the service to user. These services are Infrastructure as a service (IaaS), Platform as a Service (PaaS), Software as a Service[3]. In the cloud computing environment, task scheduling and resource assignment have been managed by providers through virtualized technology. Task scheduling becomes more complex because of the transparent and dynamic flexibility of cloud computing system, and the different needs for recourses by different applications. Task scheduling strategies only focus on equity or efficiency will increase the cost of time, space, and throughput. It improves the quality of service of the entire cloud computing at the same time. At the same time it also has number of challenges. Cloud providers and Cloud users are the two players in the environment of cloud computing for achieving different goals. Providers want to

achieve high resource utilization for having maximum revenue, on the other hand, users try to minimize the expenses, while keeping the requirement regarding the performance. Due to lack of information shared between the resources, it is tuff to allocate recourses in mutual optimal way. More over variability of the environment and ever-increasing heterogeneity gives even bigger challenges for both parties. Cloud computing is recently a booming area and has been emerging as a commercial reality in the information technology domain. However the technology is still not fully developed. There are still some areas that are needs to be focused on [4]. These areas can be broadly classified into two domains:

- Resource Management
- Task Scheduling

In computer science, scheduling is the method by which processes or data flows are given access to system resources. The scheduler is concerned mainly with:

- Throughput - The total number of processes that complete their execution per time unit.
- Turnaround time - Total time between submission of a process and its completion.
- Response time - amount of time it takes from when a request was submitted until the first response is produced.
- Fairness - Equal CPU time to each process (or more generally appropriate times according to each process' priority and workload).
- Waiting Time - The time the process remains in the ready queue.

2. TASK & RESOURCE SCHEDULING

2.1 Task Scheduling

The task scheduling goals of Cloud computing is to provide optimal tasks scheduling for users, and provide the entire cloud system throughput and QoS at the same time. Specific goals are load balance, quality of service (QoS), economic principle, and the optimal operation time and system throughput [5]. As the number of users of Cloud computing Systems increased, the tasks to be scheduled in Cloud increased proportionally. Therefore, there is a need for better algorithms to schedule tasks on these systems. Algorithms required to schedule tasks are service oriented and differ in different environments. Task Scheduling algorithms in cloud computing aim at minimizing the makespan of tasks with minimum resources efficiently. Cloud computing, uses low-power hosts to achieve high usability. The cloud computing refers to a class of systems and applications that employ distributed resources to perform a function in a decentralized

manner. Cloud computing is to utilize the computing resources (service nodes) on the network to facilitate the execution of complicated tasks that require large-scale computation. Thus, the selecting nodes for executing a task in the cloud computing must be considered [6]. A task is an activity that uses set of inputs to produce a set of outputs. In Cloud computing, user applications will run on virtual systems where distributed resources are allocated dynamically. Dynamic load-balancing mechanism has to allocate tasks to the processors dynamically as they arrive. Redistribution of tasks has to take place when some processors are overloaded. Every application is completely different in nature and independent where some require more CPU time to compute complex task, and some others may need more memory to store data. Different scheduling algorithms can be used depending on the type of the task to be scheduled. The scheduling algorithms can utilize better executing efficiency and maintain the load balancing of system. The efficiency of the cloud depends on the algorithms used for task scheduling.

2.1.1 Basic features of job scheduling include

- Automatic submission of jobs.
- Interfaces to monitor the executions.
- Priorities and queues to control the execution order of unrelated jobs.
- Interfaces which help to define workflows and job dependencies.

2.2 Resource Scheduling

Resource Scheduling became complicated task in cloud computing as required resources are limited and the number of users increase day by day so it is most important to manage these resources in such a way that resources are properly utilize and the waiting time for resources is decreases[7]. For proper scheduling of resources many algorithm are available as well as methods in cloud computing. Optimum resource scheduling help both user as well as service provider cloud provider get benefits in term of efficient resource management which in turn provides more resources to allocate user without postponing or declining user request for resources, cloud user get benefit in term of money proper utilization of money is done. Resource scheduling is the basic and key process for clouds in Infrastructure as a Service (IaaS) as the need of the request processing is must in the cloud. Every server has limited resources so jobs/requests needs to be scheduled. Each application in the cloud computing is designed as a business processes including a set of abstract processes. To allocate the resources to the tasks there need to schedule of the resources as well as tasks coming to the resources. There need to be a Service Level Agreements (SLAs) for Quality of Service (QoS).

There is enormous need for the cloud services to the schedule the resources as this scheduling will further followed by the job/task scheduling inside of the resources. There may be many instances of the single resource that they can be run at the same time [8]. There is need of checking of availability and reliability and also the load must be balanced among the resources of the same type. There are network strategies that could provide services like compute, storage and bandwidth management at less cost. The best way is combine together the computability of network strategies with scheduling algorithms. Usually when tasks are scheduled they are done according to user's requirements and requests but while

looking into all the aspects the computation needs to be done. Application scalability is the main aim for the cloud services to achieve. In cloud scalability of resources allows real time provisioning of the resources for services. Cloud has complex execution environment but it has to provide the QoS to its users. There have been a lot of work done on resource scheduling in cloud computing. New algorithms and management techniques for resource scheduling in cloud computing are being advised to make cloud computing a best experience for sources are used best for the fully customizable configuration environment for application providers as well as customers. The resource scheduling is been a tough job in cloud specially as it is the one which decides which process will be allocated to which resource and for how much time. For proper scheduling of resources many algorithm are available as well as methods in cloud computing such as Genetic algorithm.

2.3 Genetic Algorithm

Genetic algorithm is based on biological concept of generation of the population, a rapid growing area of Artificial intelligence [9][10]. GA's are inspired by Darwin's theory about Evolution. According to the Darwin "Survival of the fittest" It also a used as the method of scheduling in which the tasks are assigned resources according schedules in context of scheduling, which tells about which resource is to be assigned to which task. Genetic Algorithm is based on the biological concept of population generation.

2.3.1 Advantages of GA

- For reducing the scheduling time Genetic Algorithm is used.
- It can solve every optimization problem
- It solves problems with multiple solutions
- Genetic algorithm is a method which is very easy to understand.
- GA represents a solution to the problem as a genome (chromosome).

2.3.2 Disadvantages of GA

- The GA was very difficult for the branch and bound techniques.
- The GA did not perform well on problem in which the resources were tightly constrained.
- The genetic algorithm did not perform well on the jobshop problem.

2.3.3 Application of GA

- Manufacturing
- Computation science
- Engineering
- Chemistry
- Physics
- Mathematics
- Economics

3. PROPOSED WORK

The proposed work is to implement Task Scheduling environment by using .NET Environment and also to implement a new Enhanced GA based Task scheduling mechanism to integrate it with task grouping & priority of jobs. Figure (1) shows the work flow in this a local cloud environment is generated using .Net technology. A client-server architecture is been created. In which client assign different types of jobs and send to the server for execution. Once the server receives the jobs it initializes the jobs and also initializes the resources available. Once the initialization is done it creates the tasks then priority is been assigned

according to different factors. Now the scheduler sorts the resources according to their capability. Then the available jobs are set into groups and the execution is done by different algorithm i.e. FCFS, Priority, SJF, GA and enhanced GA. Finally the results are analyzed according to the time taken for the execution of the jobs.

Step 1: Generate Cloud Environment

In this work using .Net technology a local cloud environment is generated for execution of different Tasks. A user has to have login account for accessing this cloud environment so that user can assign task to the server. In this, Firstly select the group from which user belongs to and then user can login after authentication. Only Admin can see the full environment. It means only admin have the full access to see all the assigned jobs and to execute them, all existing resources and users. In this admin can manage the resources. Check the available jobs which are assigned by different users. Admin can assign its own jobs also. It has all the details of all the assigned jobs and executed jobs.

Step 2: Initialize Input Data

In this user can submit their job by adding job type and also have option to add description about their job which will state some properties related to the job. Submitted jobs are added to the queue in the cloud system which will be later executed by the server.

Step 3: Initialization of Resources

In this the available resources are initialized by the cloud system which can be used for the execution of the jobs. It shows the list of all the resources with their status that they are free or busy.

Step 4: Creation of Task

This defines that how jobs are set to the resources which are assigned by different users. Once the jobs are assigned there are different ways for execution of jobs by the admin. The admin chooses options for performing execution of the jobs based on different algorithms. In this the available resources are assigned to the jobs once the resources are assigned the tasks are executed. If different resources are selected, jobs are executed in parallel way and if two same resources comes up it is executed in serial. Once the jobs are executed the results can be seen i.e. the details about the time taken for execution of jobs.

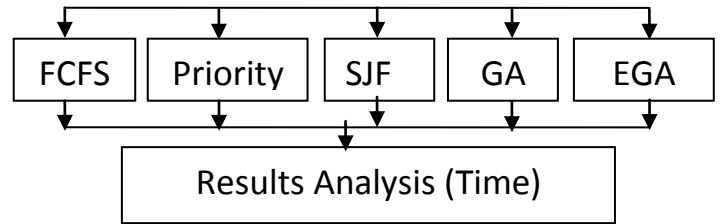
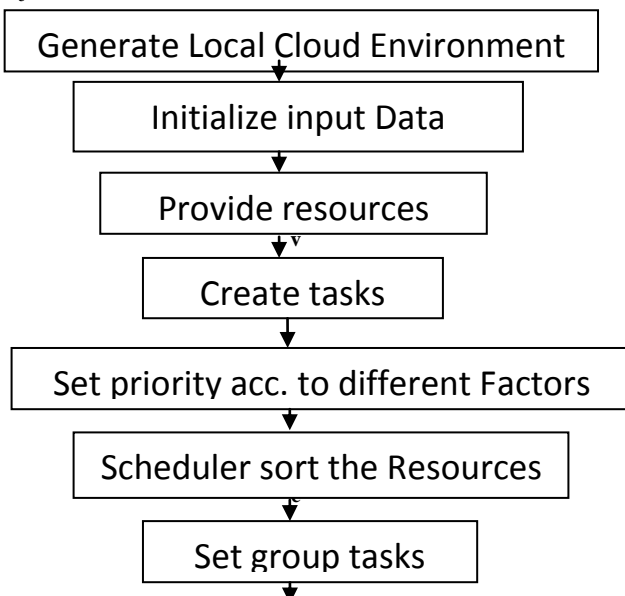


Figure 1: Proposed System

3.1 Enhanced GA based Algorithm

3.1.1 Jobs

In coming jobs are sorted according to priority. Priority is set using the concept of RBAC (Role based Access Control).

3.1.2 Resources

These are sorted according to capability which is calculated using GA algorithm.

Calculate Fit value using following equation:-

Fit value can be calculated using mean and grand mean. The formulas for calculating mean and grand mean is defined below:

Where, makespan = the time difference between the start and finish of a sequence of jobs.

$$M_0 = \text{mean}$$

N_i = No. of available resources

M_i = makespan of each resources

L = range of resources

M_i = grand mean

W_i = weight factor of resources

$$\text{mean} = \frac{\text{sum of makespan}}{\text{no. of resources}}$$

$$M_0 = \frac{1}{N} \sum_{i=1}^L N_i M_i$$

$$\text{grand mean} = \frac{\text{sum of weights}}{\text{no. of resources}}$$

$$M_i = \frac{1}{N} \sum_{j=1}^L W_j^i, \quad i = 1, 2, 3, \dots, L$$

$$\text{fit value} = \sqrt{\sum (M_i - M_0)(M_i - M_0)}$$

The algorithm level design for enhanced GA is defined below:

```

Algorithm Enhanced GA
{
    For (i==1; i<=n; i++)
    {
        Job_i = job selected by calling method
        Priority
        Resource_i = sorted according to
        capability
        For job_mapping
        {
            Calculate fitness = new_fit;
            Perform crossover;
            Perform mutation;
        }
    }
}
Else
{
    Call other method
}
    
```

4. RESULTS & COMPARISON

To analyze the performance of the proposed enhanced GA algorithm it can be compared with previous existed algorithms. For this, the present system select 5 same jobs for the execution by all defined existed methods and proposed method.

4.1 Execution by SJF Scheduling

For this 5 jobs can be submitted by different users to the cloud system. The execution can be done using SJF method. The results can be shown in the graphs.

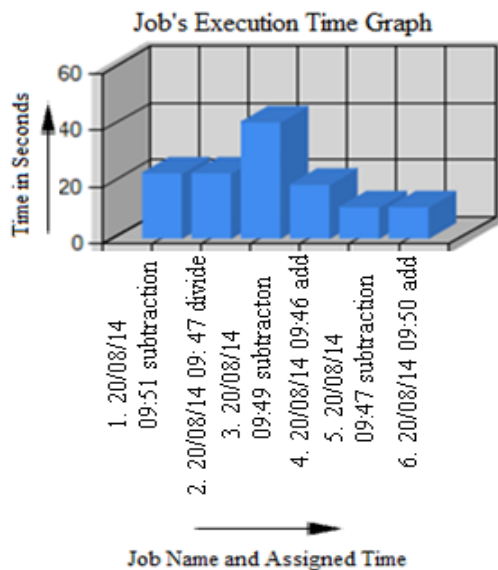


Figure 2: SJF Results

First job is executed on date 20-08-14 and time 09.51o'clock and job type is subtraction i.e. subtract the values. Time taken by this job is 23seconds and so on. Proposed work is based on total time of total no. of jobs. Total time taken by SJF algorithm for executed five jobs is 2.5 minutes

4.2 Execution by FCFS Scheduling

For this 5 jobs can be submitted by different users to the cloud system. The execution can be done using FCFS method. The results can be shown in the graphs.

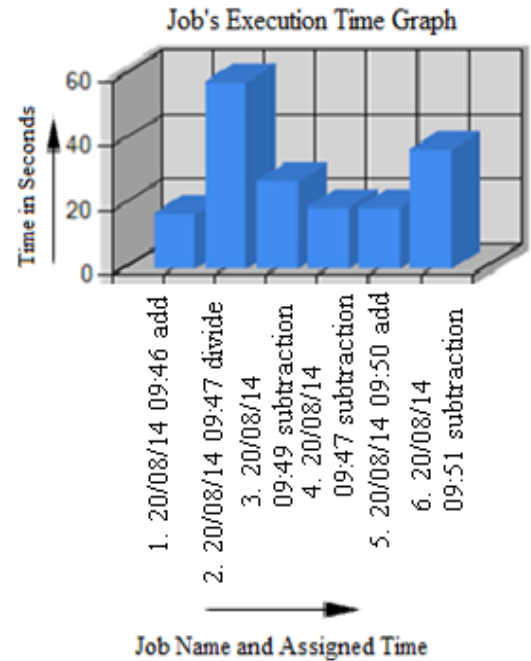


Figure 3: FCFS Results

First job is executed on date 20-08-14 and time 09.46 o'clock and job type is add i.e. add the values. Time taken by this job is 17 seconds and so on. Proposed work is based on total time of total no. of jobs. Total time taken by FCFS algorithm for executed five jobs is 2.95 minutes

4.3 Execution by Priority Scheduling

For this 5 jobs can be submitted by different users to the cloud system. The execution can be done using Priority method. The results can be shown in the graphs.

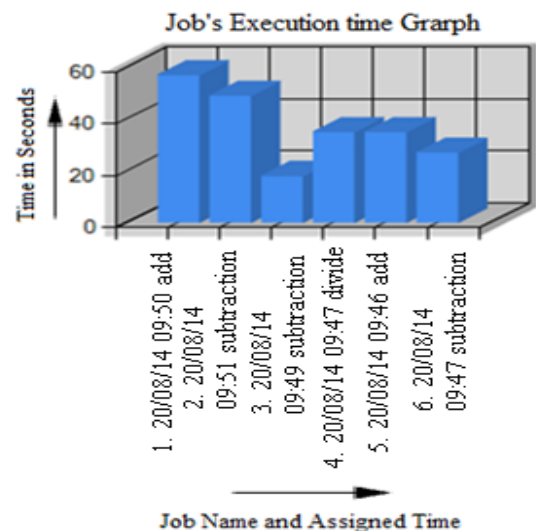


Figure 4: Priority Results

First job is executed on date 20-08-14 and time 09.50o'clock and job type is add i.e. add the values. Time taken by this job is 57 seconds and so on. Proposed work is based on total time of total no. of jobs. Total time taken by SJF algorithm for executed five jobs is 3.68 minutes

4.4 Execution by GA Scheduling

For this 5 jobs can be submitted by different users to the cloud system. The execution can be done using GA method. The results can be shown in the graphs.

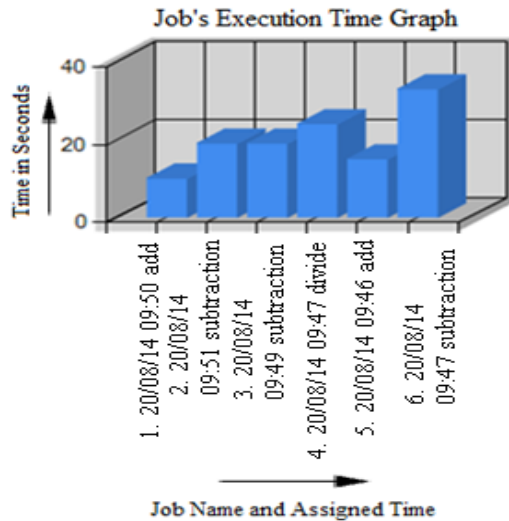


Figure 5: GA Results

First job is executed on date 20-08-14 and time 09.50o'clock and job type is add i.e. add the values. Time taken by this job is 10seconds and so on. Proposed work is based on total time of total no. of jobs. Total time taken by SJF algorithm for executed five jobs is 2 minutes

5.5 Execution by Enhanced GA Scheduling

For this 5 jobs can be submitted by different users to the cloud system. The execution can be done using Enhanced GA method. The results can be shown in the graphs.

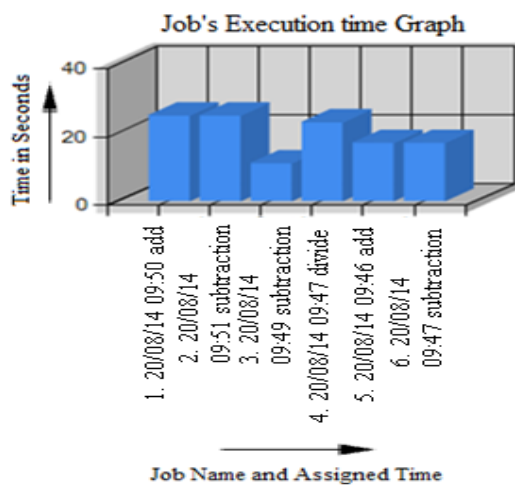


Figure 6: Enhanced GA Results

First job is executed on date 20-08-14 and time 09.50o'clock and job type is add i.e. add the values. Time taken by this job is 25 seconds and so on. Proposed work is based on total time of total no. of jobs. Total time taken by SJF algorithm for executed five jobs is 1.9 minutes

To analyze the performance of this cloud system, a comparative analysis can be done by using different algorithms. Total time can be calculated to compare the performance of different algorithm. The table 1 shows the total time taken by different execution algorithms to execute the job.

Table 1: Execution time while executing 6 Jobs

Job/ Method	FCFS	Priority	SJF	GA	Enhanced GA
Job_1	2.7	2.8	3.2	1.9	1.56
Job_2	3.4	2.6	1.8	1.53	1.45
Job_3	2.28	3.6	3.5	1.56	1.45
Job_4	3.58	1.9	2.43	1.6	1.55
Job_5	3.3	3.4	3.3	2	1.78
Job_6	3.6	3.6	2.6	1.8	1.6

The comparative analysis of all these methods can be shown in figure 7.

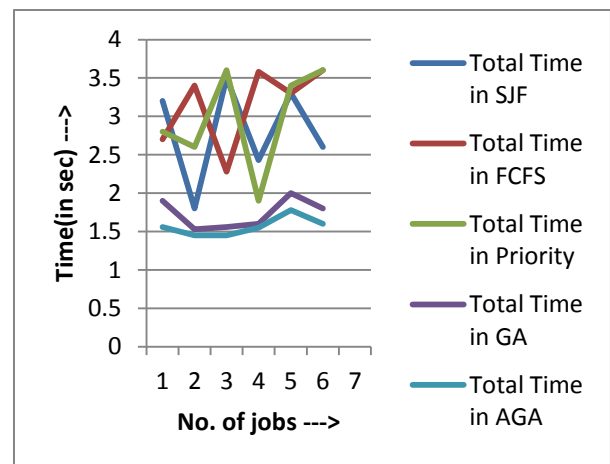


Figure 7: Comparative analysis

The performance of this proposed system can be analyzed by using different algorithms. These algorithms are FCFS, Priority, SJF, GA and Advanced GA which been implemented on local cloud environment. The differences between these are:

- The first algorithm is SJF means 'Shortest Job First'. This will execute the smallest job as the first job. So resources can be free fast and queue length will be very less. But if there are number of shorter jobs than the prior job then it will get affected with some sort of problems
- The second algorithm is FCFS means 'First come First serve'. This will execute the tasks as in the same manner they are being submitted by users. Which will firstly assign job that will executed first. It affects the system with wastage of time while any other prior task needs to be executed first.
- The third is priority; here in this work priority can be assigned based on the roles of different users. User belongs to higher authority have the first

priority to solve their task. This will get the first priority to solve the prior task first but it also have a problem like if prior task has a very big computations then the others shorter jobs have to wait for such long time.

- The fourth is GA means Genetic algorithm. In this resource mapping and job mapping both will be done by using this algorithm. This will select the best resource to being available for the optimized tasks. This results as better than the others but there is a need to modify their fitness function to select the best values which will be done in the next method.
- The last is enhanced GA. In this a new method is being introduced to calculate fitness in GA and this result is better than the all above defined algorithms.

5. CONCLUSION & FUTURE WORK

As cloud computing is the current field of research and task scheduling is one of the focused areas in it so this work is also based on the above defined areas to improve the performance of cloud in this. This work defines, that every person wants to solve their task first. But system responds in the way that can be defined in the algorithms. This cloud system works on different algorithms that are FCFS, Priority, SJF, GA and Enhanced GA. To analyze the performance each algorithm used by same defined scenario and calculate time of execution by using it. The performance of each algorithm can be defined individually and comparative analysis can also be performed by using this system. From the comparative analysis it is clear that the proposed algorithm which is enhanced GA performs better than the other methods. This can be proved by calculating time analysis as a result analysis.

The results given by the proposed work is much better than the other methods used previously. But the performance analysis can be made only by using limited number of jobs and resources. So, future work could go in the direction to test this system by increasing number of resources and number of jobs & analyze the performance of the proposed algorithm.

6. REFERENCES

- [1] Federico Etro, "Introducing Cloud Computing", London Conference on Cloud Computing For the Public Sector, November 4, 2010, pp. 01-20.
- [2] Ling Qian et al, "Cloud Computing: An Overview", 1st International Conference, December, 2009, pp. 626-63.
- [3] Dimpi Rani, Rajiv Kumar Ranjan, "A Comparative Study of SaaS, PaaS and IaaS in Cloud Computing", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 6, June 2014, pp. 158-161.
- [4] Ram Kumar Sharma and Nagesh Sharma, "A Dynamic Optimization Algorithm for Task Scheduling in Cloud Computing With Resource Utilization", International Journal of Scientific Engineering and Technology, Volume No.2, Issue No.10, pp. 1062-1068.
- [5] Lu Huang, Hai-shan Chen and Ting-ting Hu, "Survey on Resource Allocation Policy and Job Scheduling Algorithms of Cloud Computing", Journal of Software, Vol. 8, No. 2, February 2013, pp. 480-487.
- [6] Shu-Ching, Wang Kuo-Qin, Yan, "A Three-Phases Scheduling in a Hierarchical Cloud Computing Network", 2011 Third International Conference on Communications and Mobile Computing, 978-0-7695-4357-4/11, 2011 IEEE DOI 10.1109/CMC.2011.28
- [7] Hitech A. Ravani et al, "Genetic Algorithm Based Resource Scheduling Technique in Cloud Computing," Volume 1, Issue 7, December 2013.
- [8] Nimisha Singla et al, "Review of Efficient Resource Scheduling Algorithms in Cloud Computing", Volume 3, Issue 8, August 2013.
- [9] 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.
- [10] Tarun Goyal & Aakanksha Agrawal, "Host Scheduling Algorithm Using Genetic Algorithm In Cloud Computing Environment," International Journal of Research in Engineering & Technology (IJRET) Vol. 1, Issue 1, June 2013.